

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

THE INTERACTION OF
SELECTED INSTRUCTIONAL PROCEDURES
WITH SOME PUPIL VARIABLES:

A PILOT STUDY

A thesis presented in partial fulfilment
of the requirements for the degree of Master
of Arts in Education at Massey University.

Eric Lester Archer
1970

ABSTRACT

This investigation reports an experimental study of the interaction between children manifesting different cognitive styles (descriptive, categorical and relational) and two instructional methods (rule-explained, and rule-derived). The subjects for the experiment were 120 high-scorers (stanine 7+) on the specially-constructed cognitive style instrument. Equal numbers of Form I girls and boys were randomly assigned to four groups in two experimental conditions. With sex, cognitive style and method the major independent factors, the basic cell in the factorial design comprised five pupils. Teachers were added as a control, and with objectives and occasions of testing being measured across all pupils, the full design became a seven-variable one, pupils being the doubly-crossed nested factor.

The concepts and principles of stability were taught to the four groups of thirty children, over two class periods totalling one hundred minutes. The specially-trained experimental teachers taught two classes each, one by Method R.E., the other by Method R.D. Control methods included random assignment, the crossing of time, order and place of teaching, and the equating of time and content. Four dependent measures were constructed to assess pupil performance at two levels of objectives, knowledge and understanding, and application-transfer. Administered by the one tester in the school hall to all subjects, on two occasions (the day following the experimental teaching and fourteen days later), the four tests

provided measures of initial learning, transfer, retention and delayed-application-transfer.

The major interaction hypotheses postulated a higher mean score for "descriptive" children after Treatment R.D., and a higher mean score for "relational" children after Treatment R.E. Neither hypothesis was supported by the data. However, significant sex differences in cognitive style were observed. Boys tended to make more descriptive responses than girls at this age, while girls tended to make more relational responses than did the boys. Treatment R.E., an expository procedure, led to higher initial learning and retention scores than did Treatment R.D., but scores on the application-transfer tests did not differ significantly. Relative scores, however, displayed a contrasting pattern between the two method groups. The mean scores of the R.E. group for the three tests following the first test administered, were all below the measure of initial learning, whereas the reverse pattern was evident for the R.D. group.

Further examination of the data for each of the dependent measures by means of four-way analyses of variance and of covariance, was carried out. While these procedures provided additional evidence, certain limitations in the experiment and in the instruments used qualified the findings.

ACKNOWLEDGEMENT

I should like to acknowledge with gratitude the assistance given me by many people during the course of this study. In particular, I would thank the following:

The headmasters, staff and children of the schools co-operating in the experiment, and in its preparatory phases;

The ten Teachers College students who trained and acted as experimental teachers and observers;

Several staff colleagues of Hamilton Teachers College, especially Mr J. Dickie, who constructed the general models of Appendix C and who acted as a consultant on statistical problems, and Mr R. Katterns, who trained the team of students who coded the Amidon-Flanders Interaction-Analysis;

Professor C.C.N. Hill, who provided the original impetus for the study, and whose guidance and constructive comment I have valued;

and, finally,

My wife and family, the former for typing the report, the latter for assisting with many clerical tasks, all for their tolerance and encouragement.

E.L. Archer

Hamilton.

7th March, 1970.

TABLE OF CONTENTS

	<u>Page</u>	
CHAPTER I.	THE RESEARCH ISSUE	1
	INTRODUCTION	1
	THE PROBLEM	2
	EDUCATIONAL SIGNIFICANCE	6
CHAPTER II.	REVIEW OF LITERATURE: COGNITIVE STYLE	9
CHAPTER III.	REVIEW OF METHODS: DISCOVERY	26
	SOME ISSUES	26
	CONCEPTUAL PROBLEMS	29
	METHODOLOGICAL PROBLEMS	37
	SOME TYPICAL FINDINGS	43
CHAPTER IV.	REVIEW OF METHODS: EXPOSITORY	51
	RATIONALE AND THEORY	51
	PEDAGOGICAL IMPLICATIONS	57
	EMPIRICAL EVIDENCE	63
CHAPTER V.	DEFINITIONS, CRITERIA AND HYPOTHESES	72
	RESEARCH CONTEXT	72
	COGNITIVE STYLE	73
	Descriptive	74
	Inferential-Categorical	74
	Relational-Contextual	74
	TREATMENTS	76
	Treatment Criteria	77
	Additional Comments on Treatment Criteria	78
	HYPOTHESES	79
CHAPTER VI.	EXPERIMENTAL PROCEDURES	82
	SUBJECTS AND THEIR ASSIGNMENT TO CONDITIONS	82
	Assignment to Groups	82
	Assignment of Groups to Conditions	87
	THE GENERAL FEATURES OF THE STUDY	90
	INDEPENDENT VARIABLES	92
	Cognitive Style and Sex	92
	Treatment and Teachers	92
	Objectives and Occasions	96
	FURTHER EXPERIMENTAL CONTROLS	99
	DEPENDENT VARIABLES	100
	DATA PROCESSING	101

	<u>Page</u>	
CHAPTER VII.	THE INSTRUMENTS	108
	COGNITIVE STYLE MEASURES	109
	Triad One	109
	Reliability	112
	Validity	113
	Triad Two	116
	SCIENCE CONTENT MEASURES	120
	The Pre-Test	122
	The Post-Tests	125
	Reliability	128
	THE ATTITUDE SCALES	131
CHAPTER VIII.	RESEARCH FINDINGS	137
	SEX DIFFERENCES AND COGNITIVE STYLE	137
	THE MAJOR HYPOTHESES	140
	Hypotheses 2.1, 2.2 and 2.3	144
	Comparison of Score Patterns	150
	Hypotheses 3.1 and 3.2	151
	An unexpected finding: sex differences	153
	CONTROL GROUP PERFORMANCE	155
	TESTS FOR INDEPENDENCE OF TREATMENTS	156
	TEACHER EFFECTS	161
CHAPTER IX.	SUMMARY, DISCUSSION AND GENERAL CONCLUSIONS	163
	SUMMARY OF THE STUDY	163
	SUMMARY OF FINDINGS	166
	DISCUSSION AND GENERAL CONCLUSIONS	168
	EDUCATIONAL IMPLICATIONS	175
	FURTHER RESEARCH	176
APPENDICES.	A. INSTRUMENTS	180
	B. ADDITIONAL DETAILS OF TRAINING PROGRAMME	233
	C. THE GENERAL MODELS OF THE ANALYSES OF VARIANCE AND OF COVARIANCE	244
	D. TABLES OF RESULTS FROM THE VARIOUS ANALYSES	251
	E. WORK MATRICES FOR TEACHERS AND TREATMENTS	267
BIBLIOGRAPHY		272

LIST OF TABLES

		<u>Page</u>
TABLE I	Stanine levels for assignment of subjects.	86
TABLE II	Cell means for the five groups.	86
TABLE III	Balancing pattern for the experimental lessons.	88
TABLE IV	Sex response pattern.	110
TABLE V	Stability coefficients for cognitive style (Triad One).	113
TABLE VI	Between-sex differences in cognitive style.	114
TABLE VII	Extreme response set and cognitive style.	115
TABLE VIII	Discrimination indices (Item 5, Triad Two).	117
TABLE IX	Stability coefficients (Triad Two)	118
TABLE X	Correlations between Triad One and Triad Two.	118
TABLE XI	Item analysis data for embedded items.	123
TABLE XII	Post-test data for the experimental groups.	130
TABLE XIII	Correlation of Attitude Scales with I.Q. and post-tests.	136
TABLE XIV	Cognitive style - correlation coefficients between cognitive style and Otis I.Q. scores.	138
TABLE XV	Sex differences in cognitive styles among Form I pupils.	140
TABLE XVI	Selected summary of hypothesis-related results from the main analysis of variance.	143-144
TABLE XVII	Summaries of main effects and first-order interactions for analyses of covariance and analyses of variance	148-149

LIST OF FIGURES

		<u>Page</u>
Figure 1	Predicted interaction B x E based on hypothetical data	145
Figure 2	Obtained interaction B x E	145
Figure 3	Relative scores on the four post- tests in order of administration across time (adjusted means)	150
Figure 4	Interaction of sex of pupils and two testing occasions	153
Figure 5	Adjusted mean scores for boys and girls on the four dependent measures	154

CHAPTER I

THE RESEARCH ISSUEINTRODUCTION.

This investigation reports an experimental study of the interaction between children manifesting different cognitive styles, and two instructional methods.

The cognitive style variable was assumed to be an individual difference factor which would interact differentially with teaching procedures to generate different levels of learning, and different outcomes. The two instructional methods chosen as interacting independent variables possess a substantial research and theoretical literature, claims and counter-claims for the advantages of each having been a feature of the dialogues in educational psychology for many years. The teaching methods and the cognitive styles studied are cognate to the extent that each is concerned with the conceptual activities of children. The content of the experimental lessons, related to but not part of the current New Zealand science curriculum for Forms I to IV, involved concepts and principles.

The interactions between the three cognitive modes and the two instructional methods were expected to be complex. Not only were differential interactions anticipated between methods and styles, and with the sex of pupils, but, furthermore, the two methods were expected to promote different outcomes irrespective of style.

THE PROBLEM

"I suggest that we set out to invent interactions. Specifically, we ought to take a differential variable we think promising and design alternative treatments to interact with that variable." (Cronbach, 1967, p.32).

"The question of how people differ in the rate, extent, style, and quality of their learning is one which has concerned psychologists for a great many years," observed R.M. Gagné (1967, p.xi) in his introduction to a conference on Learning and Individual Differences. In further introductory comments, he stated "It appears that for many years the tradition of intelligence testing seems to have cast an obscuring shadow over the whole enterprise." While claiming that considerable advances have been made in our knowledge of learning, its varieties and conditions, over the last decade, Gagné doubted whether similar gains in our understanding (and presumably treatment) of individual differences have been achieved.

Cronbach (1967, p.23) later during the Conference, almost in direct response to Gagné, proposed a research strategy to examine the problem of instructional methods and individual differences, and pointed out that his approach called for "a new psychological theory of aptitude."

"I presume that an individual has a greater aptitude for learning.....from one method of teaching than from another method that is equally good on average. Aptitude, pragmatically, includes whatever promotes the child's survival in a particular educational environment, and it may have as much to do with styles of thought and with personality variables as with the abilities covered in conventional tests." (Cronbach, *ibid.*)

Cognitive style may well be a "promising differential variable" or "aptitude". De Cecco (1968, p.75) regards some of these styles as "learning styles.....personal ways in which individuals process information in the course of learning new concepts and principles", and cites conceptual tempo (Kagan and associates, 1964, 1965 and 1966) and selection strategies (Bruner et al., 1956) as two examples of learning style which will interact with teaching procedures whose "objectives involve the learning of concepts and principles, and problem solving." These two patterns of functioning are considered by De Cecco as having many similarities in the way they influence individual learning.

Harvey, Hunt and Schroder (1961) postulate that training conditions will interact differentially with the conceptual systems of individual learners to produce either progression or arrestation. Furthermore, their definition of concept (ibid., pp.1-8) is such that cognitive style may reasonably be subsumed within such "subject-object ties", "experiential filters through which impinging events are screened, gauged and evaluated", and "organisational properties that are not restricted to any particular referent object, but might be directed toward any object." Their particular "theoretical formulation seems especially useful because it gives emphasis to cognitive factors (clearly of relevance to a formal learning situation and its outcomes) but also because it does not in the usual way retain the dichotomy between affect and cognition." (Garrett, 1969, p.12). Nevertheless, the very eclecticism and comprehensive

generality of this position with its integrating of the findings and theories of Piaget and Werner, Kelly, Heider, Berlyne, Gesell, Festinger, Erikson, Goldstein and Scheerer, necessitates a restriction to a particular cognitive construct.

The Kagan (1963, p.76) construct with its tripartite classification, "analytic-descriptive", "inferential-categorical", "relational", has a body of literature indicating increasing research support. Its orientation differs from a number of well-known style constructs.

"While Witkin (1962) and Gardner (1959) have stressed perceptual tasks, Kagan has concentrated mostly on conceptual activities, particularly in children."
(Holtzman, 1966, p.12)

Consequently, it would appear that the cognitive style here discussed is an especially appropriate individual variable to consider in interaction with instructional methods directed towards cognitive objectives. It has further been at least partially demonstrated that the style construct has correlates with reflective-impulsive behaviour (Kagan et al., 1963, 1964, 1966; Sigel et al., 1967), with sex differences (Kagan et al., 1963; Sigel et al., 1967), with attending behaviour (Blum and Adcock, 1968; Sigel et al., 1967; Kagan and Rosman, 1964), with concept-learning tasks (Kagan et al., 1966; Lee, Kagan and Rabson, 1963), and with reading performance (Kagan, 1965), among other school-relevant and learning behaviour. Lawrence's (1967, p.107) cautionary note and comments seem well worth citing here:

"As yet the evidence is strongly suggestive rather than compelling, but it points to the possibility that there are identifiable cognitive styles or strategies which distinguish the modes of approach

of individuals to a variety of cognitive tasks Such styles are manifestations of consistent modes of cognitive organisation and appear to be related to other aspects of the individual's functioning, such as impulsiveness, and to the interpersonal relationships of the child's earliest years."

Gage (1964^a, pp.268-285) argues persuasively that cognitive learning tasks and objectives imply the use of strategies (Smith et al., 1967, p.1) based on cognitive learning theory. While there may be a danger in assuming definitive one-to-one relationships between objectives and methods, and making, in an extended Rylean sense, category mistakes, there is a plausibility in Gage's case. At the same time it is apparent that cognitive learning theory, somewhat like "teaching", is a polymorphous concept, including within the one generic term diverse referents. For example, meaningful reception learning (Ausubel, 1963^b; Carroll, 1964) is sometimes assumed to be the antithesis of discovery learning (Bruner, 1961; Kersh and Wittrock, 1962; Suchman, 1966; Wittrock, 1963^c) by teachers and, on occasion, by some writers (Hendrix, 1961). Rather than representing a Dewey-decried "either-or" dichotomy, it may well be that each theory provides an explanation of different learning outcomes for different children. Because of the practical and theoretical educational significance, and the research interest generated by each, the alternative treatments employed in this study will be forms of expository teaching (Ausubel, 1963^d, p.19) and guided discovery (Wittrock, 1963^e; Gagné and Brown, 1961).

Cronbach (1967, p.23) has stated the essence of the problem. More directly in the terms of this investigation

the problem can be posed as the following question. "Will girls and boys manifesting different cognitive styles achieve different outcomes as a consequence of two different teaching strategies (Taba, 1963), 'guided discovery' (principle derived) and 'expository' (principle given) teaching?" A number of sub-problems emerge in considering the educational significance of the problem.

EDUCATIONAL SIGNIFICANCE.

The educational significance of individual differences is well-nigh axiomatic. Mussen (1965, p.129) argues that important characteristics of learners are children's "available cognitive structures" (Ausubel, 1968), or what Cronbach (1965) has termed "symbolic systems at command", and that research is needed on cognitive style, perceptual modes and abilities, typical categories, characteristic concepts and modes of thought of children at different ages. "While personalized teaching" Mussen continues, "may be impracticable, it is at least theoretically possible that the most efficient teaching methods are those congruent with learners' cognitive styles." (Underlining not in original).

Bruner (1966) suggests that "predispositions for learning" should be a priority variable in any theory of instruction, and while he does not deal specifically in this context with cognitive style, much of his work implies that he would include it. (Bruner et al., 1956; Bruner and Olver, 1963; Bruner and Tajfel, 1961). It seems likely that Piaget's theory of cognitive adaptation (Gagné, 1968) would imply con-

sideration of a "match" (Hunt, 1961) between the stage of intellectual functioning and the task, and as Werner (Harris, 1956) suggests, "stage" implies both qualitative and quantitative change.

Methods-research outcomes have often been inconclusive, and for many possible reasons: operational definitions vary from study to study; criterion measures vary with experimental treatments; assumptions about subjects vary; so do the tasks, research designs and theoretical orientations. A basic factor contributing to the inconclusiveness is suggested by McKeachie (1961, p.111):

"Students who profit from one method may do poorly in another, while other students may do poorly in the first method and well in the second. When we average them together we find little overall difference between methods."

While it is not clear that procedures were held constant, and that it was not a teaching effect that made the difference, Heil, Powell and Feifer (1960, 1961) demonstrated that different kinds of teachers had different effects with different kinds of children.

It is felt, therefore, that this study may provide some clues to "learning styles", and to teaching methods which may be appropriate to them. Further consideration will be given to a dialogue of established educational importance, by employing variants of discovery and expositional methods in relationship to instructional objectives. A collateral product of the study will be data on the cognitive style of a group of New Zealand school children as distinct from children in the North American culture. Another claim to educational worthwhileness in this study lies in the juxtapositioning of

the chosen cognitive style dimension and these two instructional methods in particular; a survey of the literature indicates that this is a new venture.

While investigation in near naturalistic settings are inevitably compromised by the multitude of variables (Campbell and Stanley, in Gage, 1963), and while we possess as yet no theory of teaching (Gage, 1964^a), nor does the educational psychology of learning provide unqualified guidance (Gagné, 1962^a; Hill, 1963^b; Hilgard, 1964), the practical exigencies of schooling demand that some attempt be made to locate important factors in teaching. (Tilton, 1951, p.7; Scandura, 1964, p.149).

CHAPTER II

REVIEW OF LITERATURE: COGNITIVE STYLE

"The general view that cognitive functioning can be understood qualitatively as well as quantitatively," writes Cropley (1969, p.5), "is now well established in the notion of 'cognitive style' (Ausubel and Ausubel, 1966; Witkin et al., 1954) and patterning of styles (Gardner, 1964)."

The Annual Review of Psychology (Klein et al., 1967, pp.508-518) provides an imposing survey, and Ausubel (1968, p.171) enumerates almost two dozen named modes of cognitive functioning without listing any of those subsumed in studies of creativity. It is obviously premature and presumptuous to attempt any synoptic sweep in the hope of establishing relationships among those which have been studied to date. Nevertheless, many writers will agree to a considerable extent on a broad description of cognitive style as referring to "self consistent and enduring individual differences in cognitive organisation and functioning." (Ausubel, 1968, p.170).

Holtzman points out that the style dimensions "tend to be rather elusive, turning up in different configurations", that "their appearance as stable factors with the possible exception of Witkin's field-dependency construct, has not been clearly demonstrated", and that operational definitions and measures vary from investigator to investigator.

"And yet at the same time one cannot deny their existence as potentially important aspects of mental functioning that merit further, more extensive study." (Holtzman, 1966, p.4).

It is apparent, Holtzman further argues, that there are similarities in the theoretical concepts and research strategies of Witkin, Gardner and Kagan. Wallach (1962) has commented, in similar vein, that all three stress the general concept of active analytic versus passive global cognitive functioning.

However, even such integrative generalisations require qualification. Holtzman himself wrote "While Witkin and Gardner have stressed perceptual tasks, Kagan has concentrated mostly on conceptual activities, particularly in children." (1966, p.12). And Wallach, in association with Kagan (1965), demonstrated that global or "thematic" functioning need not be passive at all, that it may, indeed did in their investigations, correlate with some measures of creativity. It nevertheless does seem, in spite of quite real differences, that there is a considerable degree of overlap between the theoretical constructs. The overlap is continued in such behavioural correlates as response delay, impulse control, and cue dependency, as well as the sex differences reported in the accounts of the studies of Witkin, Gardner and Kagan. (cf. Wright and Kagan, 1963).

There appears to be considerable justification, then, in concentrating on the Kagan analytic-categorical-relational dimension, with its associated reflective-impulsive syndrome, and considering other constructs only where they seem relevant.

This is not to assume any final definitiveness in the findings, or that the dimension is necessarily unitary or a single continuum; the evidence of sex differences alone would preclude such an assertion. Some confidence in this course of action is provided by the fact that a number of educational psychologists representing somewhat different schools of thought regard the Kagan cognitive style construct as having direct relevance to teaching and learning (Mussen, 1965; Cronbach, 1965; Gordon, 1966; Ausubel, 1968; de Cecco, 1968).

The origins of cognitive style remain unresolved, and, as might be expected, support for both genic and experiential factors may be found. (Witkin et al., 1962; Kagan and Garn, 1963; Bing, 1963; Kagan and Lewis, 1965; Hess and Shipman, 1965; Maccoby, 1967). In the present state of the art an interaction thesis appears reasonable. While dispositional factors undoubtedly operate, a number of psychosocio-cultural factors seem important.

Bing (1963), studying the differences in the child-rearing practices of sixty mothers, and in their behaviour with their children, in relation to differences in the children's cognitive development, found that:

"discrepant verbal ability appears to be fostered by a close relationship with a demanding and somewhat intrusive mother, while discrepant non verbal abilities are enhanced by allowing the child a considerable degree of freedom to experiment on his own." (p.647).

This conclusion is interesting when sex differences in verbal and numerical abilities are considered. However, there are other interesting pointers in Bing's study:

"The verbal interchanges between child and mother seem to produce some intervening conditions which enhance verbal more than number or spatial performance." (p.646).

The role of language in cognitive development remains a field of speculation, and the conceptual mode under discussion is no exception. Hess and Shipman (1965) found status differences in concept utilisation using a figure sort with the Kagan mode of classification. These paralleled the social-status differences found in the same sample in regard to verbal codes categorised after Bernstein, as well as person-oriented versus status-oriented control systems. Further status differences were found in maternal teaching style. The marked relationship between these social status factors and the cognitive styles of the children of the one hundred and sixty-three Negro mothers is explained by the investigators' conclusion:

"The gross differences appeared in the verbal and cognitive environments that they (the mothers) presented." (p.59).

Tentatively one might conjecture that the relatively lower reliability figures for the inferential-categorical mode, which appears to depend in part upon language conventions, might be related to the respondent's verbal code. It is also likely that boys' preferences for science and mathematics are linked with child-rearing practices as well as cultural-social expectations and sex-typing consequences.

A consistent finding from the studies of Jerome Kagan and his colleagues, although not unique to them (Witkin, 1962), is that there are significant sex differences in conceptual style (Kagan et al., 1963; 1964; Sigel, 1967). Sigel's (1967,

p.3) summary is appropriate and succinct:

"Developmental trends from cross-sections and longitudinal studies with elementary school children indicate that descriptive part-whole and categorical-inferential responses increase in frequency with age, while relational-contextual responses tend to decrease with age. Boys show a greater rate of increase than girls in use of descriptive part-whole responses, while relational-contextual responses decrease for both sexes. Boys and girls increase in the use of categorical-inferential labelling, but girls show greater increases than boys. (Kagan et al., 1963; Sigel, 1965)".

In the early reports (Kagan et al., 1963, p.89), confirming evidence was noted in the sex-different stability coefficients based on a test-retest of the conceptual styles test, with a twelve month interval as the children moved through grade three to grade four. The analytic responses showed little change for girls ($r = .70$; $p < .001$), while the nonanalytic mode for boys reflected more than modest stability. ($r = .64$; $p < .001$). The differences reported are not confined to the test instruments, and were observed in a number of behaviour patterns; for example, error scores in a learning task with geometric stimuli were more frequent for girls than for boys. Kagan claims (p.97):

"The greater number of figure errors for girls, in contrast to boys, is concordant with their tendency to give fewer analytic responses on our tests and with Witkin's findings that girls are more field-dependent than boys."

There may be in this claim some confounding of two slightly different uses of the word "analytic", and of spatial ability with preferred mode. Nevertheless, the evidence for different patterns of cognitive style between male and female is convincing in toto. At the same time, it seems likely that the

distinctions are not simple, as Kagan (1963) indicated in reply to Wallach:

"The analytic attitude appears to be multidimensional and multiply-determined. For example, analytic responses are of different significance in girls and boys." (p.123).

A variable which has been associated with the particular conceptual mode being considered has been that of delay. Analytic children, also considered reflective in conceptual tempo, suggests Kagan, tend to withhold response until they have evaluated the risk in alternative answers. Others (impulsives) who make relational responses on conceptual styles tests, choose quickly and with less thorough evaluation of the various possibilities. Ward (1968, pp.867-868) wrote:

"For children in the early school years, this response style (i.e. to delay) predicts a number of errors on tests of inductive reasoning (Kagan, Pearson and Welch, 1966) and of word reading (Kagan, 1965) as well as preference for analytic as opposed to thematic groupings on the Conceptual Style Test (Kagan, Moss and Sigel, 1963)."

Ward tested the postulation (Kagan et al., 1964, p.34) that the impulsive child was avoiding anxiety stress by responding quickly, and found that impulsive children, after being informed of error, responded subsequently with greater delay. "Impulsive children were thus more responsive to evaluational cues", Ward concluded (1968, p.873), in reference to the five-year-old children studied. Ward agreed with Kagan (Kagan et al., 1964, pp.33-34) that longer response times might be "due to greater involvement and desire to do well on intellectual tasks, and that cognitive impulsivity is one

instance of a broader syndrome which includes high motor activity and short attention span." In this investigation, no effects were found of the child's age, sex or intelligence. The final statement Ward provides on his data, questions one earlier Kagan interpretation.

"They make it clear, however, that situational variables as well as factors intrinsic to the child play a role in effecting reflective or impulsive behaviour." (p.873).

This is not entirely surprising, for as Sigel (1967) points out, style is a "preferential mode of categorisation expressed in a situation where alternatives are possible." (p.2). Wachtel (1968) in an empirical investigation of the relation between the dimensions studied by Kagan and Witkin, as a replication of Messick and Fritsky's (1963) study, employed a group embedded-figures test and a figure sorting task with a design variation task as a criterion measure. The field-independent male psychology students showed general superiority over field-dependent students in identifying part aspects, were superior in verbal intelligence measures ($t = 2.48$; $p < .02$) and on the criterion task ($F = 23.21$; $p < .001$). It would appear, Wachtel reports (p.207), that a common capacity, independent of verbal intelligence, which was covaried in the analysis of variance, underlies performance on the embedded-figures test and on the design variations task. However, there was little evidence for any relation between the style of categorisation and performance on the earlier criterion task, or on the embedded-figures test. Witkin et al. (1962) indicated that the analytic character-

istic of field-independents is a specific superiority in extracting items from an embedding context, rather than a general responding to discrete portions of a stimulus. Factor analytic studies by Karp (1963) and others have lent support to this interpretation. Wachtel's conclusions (p.209) are that the two measures represent different aspects of analytic functioning, that while Witkin's tests "measure a subject's capacity", Kagan's tests measure "a stylistic preference". (Italics in original). If this is so, it provides additional support to Sigel's comments and to Ward's summation. It would also account for some of the variations reported in other studies where instruments of a demand nature rather than those presenting alternative possibilities have been used. Where overlaps between responses on Witkin's tests and the conceptual styles test are found, they probably reflect the relationship between perceptual and cognitive processes discussed by Ausubel (1965).

While situational factors can be expected to interact with stylistic preferences for information processing, these conceptual modes do appear to become increasingly stable, and are not always susceptible to much modification (cf. Yando and Kagan, 1966). Two particular accounts in the literature are relevant to this point. Kagan, Pearson and Welch (1966b) attempted to answer the two questions:

Can an impulsive attitude be modified through direct training in reflection?

Is there any training advantage when the child initially perceives some basis of similarity to the trainer?

While the training procedures did produce larger response latencies among impulsive children, and the tendency to delay did generalise to an adult other than the trainer, no significant change was observed in the number of errors made, nor was generalisation to an inductive reasoning test found. Edwards (1969) tried to evaluate the effects of instruction in categorising on a modified conceptual styles test. The first grade boys involved in the investigation were assigned by three I.Q. levels (Pintner-Cunningham) to one of three conditions - prompted, unprompted or control. It is noted that Edwards' interest seems to be focussed mainly on the number of responses. The subjects in both treatment groups made significantly more categorisations than the control group, and also made more analytic and inferential-categorical responses. However, change scores are not reported, nor is it stated whether the modified response pattern was stable over time. The unprompted group, for example, performed as well as the prompted insofar as total categorisations and use of analytic and categorical-inferential responses were concerned. Edwards states (p.142):

"From these data, then, children who are required to categorise objects on their own do as well or better when classifying new objects than children who are given intensive cues suggesting reasons for matching objects."

An inspection of the results indicates that the unprompted group made more relational responses than did the prompted

group, maintaining a similar pattern across categories to the control group. It is in the categorical-inferential mode that the prompted group seems to have gained. If this is indeed so, it may reflect a language effect as suggested earlier in this review.

Two other conclusions drawn by Edwards seem relevant to this discussion of Kagan's hypothetical construct. A significant positive correlation ($r = .36$; $p < .01$) found between inferential-categorical and relational responses with these subjects, and this, together with Wallach's and Kogan's (1965) report that boys high in both intelligence and creativity exhibited a balanced usage of relational and inferential-categorical reasons when grouping, Edwards interprets as throwing doubt on the unidimensionality of relational concept preferences as posited by Kagan et al. (1963). It is possible that the categorical-inferential mode, however, is the more complex. Edwards reports that intellectual ability was related to the frequency of categorical-inferential concepts employed to classify objects on his criterion test, a finding supported in Wallach and Kogan's (1965) study, whereas I.Q. was not related to the number of relational or analytic concepts employed in categorising objects in either study. In each case it must be noted the subjects were boys. A different pattern emerges when the relationship between categorising preferences and general intellectual ability for girls is considered.

Sigel (1967, p.3) summarises some of the data:

"Relationship between styles of categorisation and I.Q. (California Mental Maturity Scale) has been assessed in a number of studies. For the descriptive part-whole approach, significant relationships have been found with performance I.Q. for boys only; relational-contextual labelling was negatively related to verbal and performance I.Q. for the boys and girls, but significantly negative only for performance and full-scale I.Q. for boys, but showed strong trends in relation to I.Q. for girls (Kagan et al., 1963). Categorical-inferential labels were significantly associated with verbal and performance I.Q. for boys, but only trends are evident for the girls. The relationships with I.Q., then, are contingent on sex of the respondent, as well as whether the I.Q. is verbal or performance."

It seems likely that some of the discrepancies in the literature in regard to general abilities and conceptual style are associated with somewhat differing measures of both categorising preferences and intelligence, confounded by sex and age, or developmental level. Although Wachtel (1968) argues for a distinction between capacity and style, there is no necessary bifurcation between level of intellectual functioning and mode; both Cropley (1969) and Holtzman (1966) postulate a developmental interaction. Nevertheless, the research reports and the theoretical literature incline the balance of evidence towards the view that the cognitive style preferences under discussion have shown temporal stability in individuals and inter-task generality.

The generality is also noted in association with a number of personality characteristics. For example, descriptive part-whole responses have been "found to be related to cautiousness, controlled impulse expression, reflectiveness, independence and achievement orientation in boys, while relational-contextual responses related to impulsivity,

dependence and anxiety." (Sigel, 1967, p.4). The results for girls and women over the same domains are equivocal.

This study is concerned with the extent to which categorising preferences are variables affecting school learning. Accordingly, attention is now turned to school-task-relevant investigations. Kagan, Pearson and Welch (1966^a) showed that:

"a reflective tendency generalises to inductive-reasoning tasks that contain certain response uncertainty..... The impulsive child responds quickly in situations where inferences are required; he seems to report the first idea that occurs to him." (p.594).

In discussing their results the authors offer advice to teachers, remarking on the requirement of inference from the child in arithmetic, social studies and science, and, in fact, any programme which emphasises discovery learning methods of instruction. A similar conclusion as far as response uncertainty is concerned is reported in Kagan's (1965) studies of reflection-impulsivity and reading ability in primary grade children. Hypotheses that reflective children, in contrast to impulsive ones, would make fewer errors in reading English words presented singly or in a prose section, were confirmed. The degree of response uncertainty seems to be a factor in the situation influencing the extent to which a preferred mode is employed, for relationships between fast decision times and reading errors were more clearly apparent with high verbal subjects than low verbal subjects, for whom the tasks may have been too difficult to offer alternatives. Kagan suggests (1965^b) that discovery approaches, especially where

they involve greater than usual cognitive cost, may be inappropriate, since "impulsive children are apt to settle on the wrong conclusions in the inferential method, and become vulnerable to developing feelings of inadequacy." (p.561). Lee et al.(1963) produce further data from the responses of thirty third-grade boys, identified by means of the Conceptual Styles Test, to show that those who have a preference for analytic groupings learn analytic concepts more readily than do non-analytic boys, who learn relational concepts with greater ease. The researchers state (p.442):

"These results suggest that the final outcome of a conceptualisation task is not merely a function of the ability to form associations between stimuli and responses, or the availability of mediational labels that are relevant to the content of the class being formed. Conceptual products are also influenced by the individual's preferred focus of attention during the initial stages of learning."

"Evidence has been accruing," write Blum and Adcock (1967, p.31) in their selective review of attention and early learning, "that children attend differently, not only on the basis of age and sex, but also on the basis of what has become known as 'cognitive style'." Attempts have been made to differentiate between analytic and non-analytic styles with indices of attention. Kagan and Rosman (1964) found that cardiac deceleration and respiration variability of analytic and non-analytic boys differ, as do these rates during periods of attention and rest. Blum and Adcock continue their review commenting on attention as a two-faceted phenomenon involving "both scan and field articulation, depending on the task involved", and pointing out that "we see from Kagan et al.'s work that a tendency towards one or other type attentive

behaviour exists relatively independent of the task requirements." The reviewers note that research by Blum and Broverman (1967) supported this position, and that with the additional confirmation of Gardner and Long's (1962) investigation of scan and field articulation in sixty housewives, it might "point to the enduring nature of such response or cognitive styles." (p.32)

An area still lacking the fine edge of precision concerns the measures which have been used to define cognitive styles. For some studies an array of discrete figures provide the stimuli, the subject being asked to select groups from the array; for other studies, a set of triads comprises the stimulus situation, with the subject selecting two from each group of three on the basis of perceived similarity. The reasons provided for grouping or selecting are scored according to criteria; the scores are then coded in the three categories descriptive-analytic, categorical-inferential and relational-contextual. Referring to the triad system, the senior author (Kagan et al., 1963) expressed the judgement:

"We regard the ...conceptual style test as a much better measure of analytic attitude in children than the free sorting of human figures." (p.110).

The criteria for scoring have varied somewhat in detail in the reports (Kagan et al., 1963; Wallach and Kogan, 1965; Sigel, 1967), but have maintained a general consistency. The first conceptual styles test was the subject of criticism on the grounds (Wright and Kagan, 1963) that:

- (a) negative correlations between analytic and relational modes could be test artifacts, and that

- (b) the test was constructed in such a way as to limit the number of categorical-inferential responses.

Subsequent studies suggest that the negative correlations are more likely to reflect actual differences in cognitive style than to be artifacts. In addition, the conceptual styles test permits subjects to select the same two stimulus figures from the triad but for reasons which may be coded under different categories from the tripartite classification. Furthermore, the number of inferential-categorical responses may be increased through item construction and selection following item analysis. Thus, it is suggested, cognitive style may be defined by a conceptual styles test which presents triads from which the subject will select pairs according to perceived similarity, the reason for each pairing being stated by the subject and coded by the examiner on the basis of established criteria. The criteria already established (Kagan et al., 1963, 1964; Wallach and Kogan, 1965; Sigel et al., 1967) will guide the coding in this study.

While necessarily selective, the literature and research studies surveyed provide support for asserting the stability, generality and educational relevance of Jerome Kagan's cognitive style construct. The discrepancies between investigations seem to a large extent to be explicable on the grounds of defining measures, the age and sex of the subjects, and the nature of the task, be it demand or one that permitted preference in choosing among alternatives. The conceptual styles are complex

and multidimensional, but there is a coherence about the patterns of responses and their correlates which provide shape for the construct, even if definitive boundaries are not possible at this stage. The developmental trends observed in longitudinal studies (Kagan and Moss, 1962), together with many theoretical postulations (e.g. Piaget, Werner, Vernon, Witkin, Gardner), imply a progressive differentiation in cognitive functioning such that at any particular "stage", a general pattern of preferred categorising might be anticipated with younger children demonstrating a tendency towards the global, and older persons demonstrating a more differentiated mode. To the extent that research endorses this likelihood, one might argue for the stochastic (Berlyne, 1965) nature of cognitive style. Certainly, conceptual modes in this sense have parallels with factor analytic studies of intelligence, and at the same time with Vernon's (1961) hierarchical view of "the infinitely varied thinking processes" termed intelligence.

The argument thus far has tended to emphasise group trends, but this has not been to deny individual differences and associated variability. Strong motivational factors (e.g. a need for approval from a demanding teacher) interacting with stylistic preferences in different stimulus situations are likely to generate some variation from the probabilistic predictions based on defining measures. The interrelations of style and motivation are still unresolved. While situational factors, for example, undoubtedly play a part in deciding the extent to which preferred modes are employed,

there seems little doubt that individual consistencies in cognitive style, in turn, influence which stimulus aspects are selected as a basis for grouping. Those questions which persist (e.g. concerning the determinants, the role of language, the extent to which the styles are modifiable, the stage at which they become stabilised and resist change, the judgements about the long-term consequences of one rather than another mode) are the very questions which persist in regard to the whole range of human intellectual functioning.

The assumptions underlying a theory of cognitive style are the assumptions that Baldwin (cited in Kohlberg, 1968, p.1017) expands in connection with any "cognitive" theory. This study is concerned less with error score or delay in near-laboratory types of situation than it is with the interaction between cognitive style and instructional methods in a near-naturalistic school situation. Kagan (1965, pp.561-562) suggests part of the reason for this:

"There is not only interaction between the psychological organisation of the child and the method of presentation, but also between the substantive content and the method of presentation. It would seem that contents that are tailored for discovery learning are disciplines containing principles that are usually induced through the proliferation of hypotheses."

CHAPTER III

REVIEW OF METHODS: DISCOVERYSOME ISSUES.

Discussing "Issues Current in Educational Psychology", Cronbach (1965) exposed to critical examination some of the slogans in regard to discovery and didactic teaching methods, and revealed the programmatic (Scheffler, 1960) and emotive nature of many of the definitions of the two terms. The case for (Bruner, 1961; Hendrix, 1961; Taba, 1962; Suchman, 1964) and against (Ausubel, 1961, 1963², 1968) discovery procedures has become part of the corpus literarum of educational psychology, particularly since 1955. The educational and psychological rationale of discovery learning has been most eloquently stated by Bruner (1961), who sees the benefits in terms of increased intellectual potency, intrinsic motivation, improved techniques or heuristics of problem solving and enquiry, more personal meaning through involvement, more generic learning to promote retention and transfer power and in its contribution to such major educational goals as problem-solving ability, autonomy and learning-how-to-learn.

After a close scrutiny of the research and writing, Ausubel (1968, pp.472-473) summarised what he considered "the legitimate claims, the defensible uses, and the palpable advantages of the discovery method". Among these warranted claims were counted advantages for children in the concrete-operational stage of development; for learning tasks where

concrete-empirical experience is a pre-requisite to meaning, or where the task is difficult and unfamiliar; as a means of evaluating hypothesis-making, problem-solving skills and desirable attitudes towards inquiry and the scientific method.

Ausubel concludes:

"Finally it still seems plausible to suppose that the greater effort, motivation, excitement and vividness associated with independent discovery lead to somewhat greater learning and retention. One might expect the advantages conferred by discovery techniques to be even greater with respect to transferrability....." (sic)

Ausubel's main criticisms of the discovery hypothesis (Wittrock, 1966), apart from his rejection of many extravagant claims (e.g. Hendrix, 1961), are twofold. He questions the appropriateness of the method for transmitting the substantive content of a discipline to cognitively mature students, and he questions the relative efficiency of the method in achieving cognitive goals (concepts and principles) which might be as effectively learned through expository teaching in less time and with less cognitive cost.

The very words 'discovery' and 'exposition' tempt proponents to raise positional banners, and Ausubel (1963^a, p.19) found himself impelled to provide a cogent defence of expository teaching. Exposition is regarded as a creative teaching enterprise, demanding ability in "selecting, organising, presenting and translating" of subject matter content in a developmentally appropriate manner, so that a direct grasp of higher-order relationships among abstractions might be meaningfully achieved by learners. For Carroll (1964), too, apart from economy and efficiency in use of time,

expositional procedures when combined with practice are particularly successful in teaching concepts and principles, and have special strength in the presentation of an organising framework in which to learn the unifying principles of a discipline. As Cronbach (1965) incisively argued, didactic teaching need not be authoritarian, identified with rote learning, or (1966) condemned to occupy a strategically-weak position vis-à-vis discovery approaches.

Fewer tendencies to polarise the two strategies are apparent in the more recent literature (Cronbach, 1965, 1966; Johnson and Stratton, 1966; Ausubel, 1968). Kersh (1964, p. 230) commented after a survey of several studies "..... learning by discovery is not necessarily the most effective learning procedure for all teaching objectives...." Taba (1963), Ausubel (1961), Kersh and Wittrock (1962) and Bibergall (1966) concur in seeing rote learning as distinguishable from meaningful reception learning and discovery learning, and each recognises the potential meaningfulness inherent in both discovery and expository teaching. Either type, Bibergall (1966) agrees with Ausubel, can be made meaningful or rote. The dialogue in the past has been overlaid with an excess of static arising from different views about the role of verbalisation (Hendrix, 1961; Taba, 1963; Wittrock, 1963; Gagné and Smith, 1962); about priorities in educational goals (Cronbach, 1966) and consequently criteria; about the deductive-inductive debate; and by the remnants of the authoritarian-democratic dialectic (Anderson, 1959).

CONCEPTUAL PROBLEMS.

The appropriateness of either method appears to depend on the objectives pursued (Kersh, 1964; Cronbach, 1966), the nature of the task (Kersh, 1964), and pupil characteristics including, inter alia, age (Ausubel, 1964^d), prior learning (Gagné and Brown, 1961), set (Wittrock, 1963^b, 1964), and general ability (Corman, 1957). Dearden's (1967) philosophical analysis of discovery teaching throws a probing light on such unstated assumptions as "abstractionism", and points out that discovery learning is no unitary construct, but one which differs in degree as well as kind at different developmental levels. Scandura (1964, p.149) endorses the educational psychologist's awareness of the complexity when he observes "..... exposition and discovery refer to classes of methodology - not uniquely defined methods...."

Conceptual and research problems characterise much of the discourse concerning discovery. For example, the question "What is discovery?" raises epistemological, linguistic and psychological issues. Wittrock (1966) has analysed some of the issues, exposing the tautological problem resulting from attributing the same name to independent, dependent and intervening variables, and the problems resulting from semantic inconsistencies in labelling different treatments (independent variables) in the same way. The semantic problem appears to require resolution through clearly stipulated operational definitions, since, at this stage, there is no taxonomy of treatments. Wittrock (1963^a) proposed that descriptive terms be used to label the treatments, and that

the amount of guidance or direction in providing rules and/or solutions be specified. (Wittrock and Kersh, 1962). Wittrock (1963^a, p.184) exemplifies the problem in referring to the inconsistent use of labels for independent variables in empirical studies:

"Kittell's (1957) intermediate direction group, Kersh's (1962) rote learning group, Craig's (1956) directed group, Katona's (1940, p.36) understanding group, and Forgas and Schwartz's (1957) observer groups are all examples of giving of rules."

Later studies have tended to be more precise, by either specifying the treatments in such a way as to indicate the provision (or not) of rules and solutions (Guthrie, 1967; Roughead and Scandura, 1968), or by stating clear criteria by which to define the treatments (Oliver and Shaver, 1966; Worthen, 1968^b). A further feature of more recent studies has been the tendency to avoid the tautological problem by restricting the term discovery to one class of major variables (Worthen, 1968^b). In addition to more precise operational definitions and the restriction of the term discovery, a number of writers (see e.g. Shulman and Keislar, 1966) have suggested that where the term is used to refer to intervening variables, the criteria on which the inference is based should be clearly indicated.

For example, a distinction can be drawn between learning to discover, and learning by discovery (Glaser, 1966). Although a close examination of criterion measures (cf. Kersh, 1958, 1962, and Gagné and Brown, 1961) will help to identify in which sense the term is being used, the distinction is more honoured in the breach than the observance. Even Shulman

and Keislar (1966, p.192), discussing the Conference on Discovery from their editorial perspective, fail to maintain the distinction. Commenting on the intervening variable as "an inference based on several kinds of evidence," they suggest as minimal criteria:

- "1. Although the learner may have access to a good deal of information prior to solving a problem, the identification of the solutions themselves must never be part of the information he is given.
2. The learner must be able to generalize the solution to other situations. If no such transfer is evident, the successful first solution is considered an accident and not a discovery."

These criteria, in spite of the inferences which can be drawn from the first, appear to belong to the learning by discovery category.

That the uses of the word discovery are manifold became obvious at the Stanford-sponsored conference. One group of participants wished to discard the notion of covert mediational mechanisms, and to concentrate on input and output variables only. The majority preferred to accept the concept of internal mediators as a useful construct for any investigation of discovery learning, and of these some, like Kendler, felt that language was of central importance in the mediating processes. Gagné, representing a third point of view, saw the term as having broad application to any task, be it motor skill, association, discrimination or principle learning and problem solving, where the learner is engaged in search and selection processes. The fourth viewpoint was provided by those who limited the use of the word to cognitive areas of

learning; these participants emphasised the act of discovery per se (Bruner, 1961).

The research literature reflects similar variety. Kersh (1958,1962) attributed to motivation a central place among the intervening variables; motivational factors accounting for the "travel effect" of discovery learning (cf. Baddely's 1963 study of zeigarnik effects). Hendrix (1961) postulated pre-verbal intuiting of relationships as the crucial internal mechanism. Kornreich (1969) suggested that problem-solving strategies rather than specific response choices were the mediators of importance if transfer were the objective. Each tends to use the word primarily in the learning to discover sense, as well as the learning by discovery sense, but the categories overlap and merge until the boundaries are lost. The diverse uses of the term may be reflections of the perennial process-product debate.

If the question "What is discovery?" has elicited a host of answers, it is no surprise to find a similar yield to the question "What is discovered?". Kersh and Wittrock (1962), after surveying a number of studies, conclude that the answer can be parsimoniously provided by deciding what it is that has been practised and reinforced. Scandura (1968) suggests that the answer depends largely on the set provided by instructions, and that the set may be, in Set-Function Language, to seek derivation-rules for deriving a class of more specific rules. Hendrix (1947) argued, with theoretical vigour if not empirical rigour (Ausubel, 1968), that the essential discovery is a sub-verbal awareness, a pre-requisite

to meaning. Kornreich (1969) suggests that what is prompted is what is learned, but his work and the investigations of Wittrock indicate that the degree of specificity of the cues and prompts may be the critical factors in determining what is discovered.

Another question underlying the research concerns the stage of learning when discovery, as a covert process, occurs. This question is relevant to claims for sub-verbal awareness (Hendrix, 1961), and is answered in part by Kersh and Wittrock (1962, p.462), who state that it is "likely to occur prior to the learner's first acceptable response." Jacobson et al., (1969), discussing the relationship of intelligence and mediating processes to conceptual learning, suggest that these influences are likely to have most effect during the response generation stage rather than the recognition stage. Though the direction of the answers is consistent, the explanations, rather like those referring to problem solving (Duncan, 1956), seem still rather tentative.

The answers provided in the literature appear to have been influenced by the criterion measures used, by the theoretical orientation of the researcher, and by the nature of the learning task. Reports have stressed motivation (Kersh, 1958,1962; Bruner, 1961); attitudes (Kagan, 1966^a; Cronbach, 1966; Worthen, 1968); self-perceptions (Cronbach, 1966; Kagan, 1966^b); and affectivity generally (Wittrock, 1966), apart from cognitive outcomes. Cognitive objectives most frequently stressed include retention (Craig, 1956; Ray, 1961; Kersh, 1962; Guthrie, 1967); ability to transfer (Haslerud

and Meyers, 1958; Gagné and Brown, 1961; Guthrie, 1967); problem-solving techniques (Wittrock et al., 1964; Worthen, 1968); principle learning (Forgus and Schwartz, 1957; Haslerud and Meyers, 1958); as well as concept identification and discrimination (Kornreich, 1969). The most recent studies have used multiple criteria. For example, Worthen (1968) used tests of initial learning, retention, transfer and heuristics, supplemented by measures of attitudes and pupil perception of the teaching approach.

Learning tasks have been somewhat restricted in range. Cryptogram tasks (Haslerud and Meyers, 1958); word tasks (Craig, 1956; Kittell, 1957); number series (Kersh, 1958, 1962; Gagné and Brown, 1961; Roughead and Scandura, 1968); and codes of various kinds (Forgus and Schwartz, 1957; Wittrock, 1963; Guthrie, 1967) represent the major types. Later studies show trends towards school learning tasks. Ray's subjects (1961) studied the use and principles of operation of the micrometer; Worthen's pupils (1968) were involved in the usual Junior High School mathematics curriculum; Oliver and Shaver (1966) were concerned with controversial social studies material at the High School level.

Subjects have varied across studies, from College students (e.g. Craig, 1956; Kersh, 1958; Haslerud and Meyers, 1958) through Grade Nine (e.g. Ray, 1961) and Grade Six (e.g. Kittell, 1957) to Kindergarten children (e.g. Wittrock, Keislar and Stern, 1964). In some cases the subjects have been volunteers (Roughead and Scandura, 1968); of one sex only (Forgus and Schwartz, 1957); or stratified by intelli-

gence (Ray, 1961). Individual difference variables have infrequently been a focus. It is difficult, for example, to find a study where sex of learner has been used as a contrast basis when comparing treatment effects. It is understandable that Cronbach (1966) should propose major changes in research strategies.

Thus, what is discovered depends not only on what is practised or reinforced, but also on what is prompted, or what is measured, or the nature of the task, and on those learner characteristics relevant to the task and to the method used. In controlled experimental situations with individuals, with very small groups, or where feedback is carefully circumscribed, it may be possible to answer the question "who discovers?", but in naturalistic settings it is doubtful that any certain answers can be supplied. Not only are behavioural clues susceptible to different interpretations, but where they can be recorded on videotape, it soon becomes apparent that some children are "target" pupils, while others are "audiential" (Biddle and Adams, 1967). The interactions of teacher and target pupils may well have "ripple effects" (Kounin and Gump, 1961), but inferences about the covert mental processes of the audiential pupils can only be rather speculative at this stage. Of those children presumed to discover, some may have learned from actual discoverers by observation learning (Bandura and Walters, 1963), by meaningful reception learning (Ausubel, 1968), or as a consequence of reinforcement of acceptable approximations (cf. Jackson, 1966; Holt, 1964; Bugelski, 1964). A number of investigators (e.g. Flanders)

imply that frequent opportunities for independent problem solving may be the exception rather than the rule in many classrooms.

Wittrock (1966, pp.42-43), in summarising the conceptual issues, makes several recommendations:

- (a) The ends which discovery is presumed to serve, be they subject matter, learning, or problem solving procedures, should be clearly specified; and within the subject matter, distinctions should be attempted between the ends of rule learning, the learning of generalisations, and of specific information.
- (b) The results should be carefully related to well-specified dependent variables, and contrasts between findings should be limited to the same types of indices of discovery.
- (c) Discovery should be viewed as a set of very complex procedures, not ingenuously confused with induction, nor depreciating verbal learning, nor disregarding the sequential pattern and presentation of stimuli.

Bibergall (1966) makes similar points as she distinguishes between approaches to discovery learning according to the degree of autonomy (pure, guided, directed) given each learner. That the conceptual problems have practical consequences becomes apparent as methodological problems are considered.

METHODOLOGICAL PROBLEMS.

Both Cronbach (1966) and Wittrock (1966) explicate a number of methodological issues in reviewing the research literature. Each refers to treatment factors and design factors. Difficulties associated with stating either precise operational definitions, or, perhaps more justifiably because of the complexity of teaching methods, clear criteria to delineate the teaching procedures, have made replication studies rare. The tendency has been to hold tasks constant, vary treatments, and often subjects, and then to generalise conclusions to different tasks. As Cronbach implies, this seems peculiarly illogical. Furthermore, where discovery has been compared with alternative didactic teaching strategies, the didactic group has often suffered disadvantages in time, in kind of goal set, in limited rationalisation of solutions, in the way the task has been structured, or in the way the results have been analysed (Cronbach, 1966, pp.76-97). Clearly, each treatment must be given a "fair chance to show what it can do."

A number of within-treatment variables have been regarded as significant antecedents by investigators. Gagné and Smith (1962) have demonstrated the value of requiring pupils to verbalise during practice in problem solving. Instructions (Underwood and Richardson, 1956; Kersh, 1965; Scandura, 1966); provision of information (Corman, 1957); timing (Scandura, 1964); cues (Wickelgren, 1964; Wittrock et al., 1964; Wittrock and Keislar, 1965); set (Wittrock, 1962, 1963b, 1963c; Yonge, 1966); and sequence (Newton and

Hickey, 1965; Worthen, 1968), have all been shown to be influential factors in guiding children to achieve criterion performance. Cronbach, after remarking on the arbitrariness of many experimental tasks, made a cogent case for tasks where rationality is at a maximum, where answers are empirically confirmable, logically necessary, or are readily discerned consequences. To date the optimum role of each of the within-treatment variables is unresolved. While the current trend is analytical, with increasing attention being given to sequence, set, and cues to mathemagenic behaviours (Rothkopf, 1965; McDonald, 1968), there remains the danger exemplified in the writings of Locke and Hume, wherein analysis to ultimate components creates a new problem in synthesising. It may well be the only course left open to experimenters is to conduct multi-level analyses ranging from the instructional gestalt level (Siegel and Siegel, 1967) through to micro-level analysis (cf. Gage, 1966; Smith, 1967), in studying a particular teaching strategy. (See Scandura, 1964, pp.155-156). Perhaps this was an additional reason for Meux's (1967) remark on Shulman and Keislar's (1966) volume:

"The reviewer was considerably puzzled, however, in finding not one classroom observation system referred to in the book." (p.551).

Another area receiving surprisingly little emphasis in the research designs has been that of individual differences. Gagné and Brown (1961) took account of pre-requisite learning as a factor in determining transfer. Bruner (1961, p.22) has indicated the importance of prior learning in his remark "Discovery, like surprise, favors the well-prepared mind."

Kersh and Kagan both suggest that motivational factors may be important predispositions for discovery learning, and along with other writers (e.g. Hunt, 1960; Suchman, 1961), make reference to cognitive (e.g. Festinger, 1957; Berlyne, 1965) and competence (White, 1959) motivational constructs as having explanatory power. Paradowski (1967, p.50) reports that curiosity "significantly increased both intentional learning and incidental learning", a conclusion which has suggestive overtones in regard to discovery and the " sleeper effect" found by Kersh (1962).

Ray (1961), using a treatment-by-levels design, states:

"The complete absence in this experiment of significant interaction between teaching method and learning ability was an unexpected result." (p.278).

This finding, Ray suggests, would justify the use of discovery procedures with pupils of low mental ability. A number of studies of concept attainment and transfer provide a lead in treating such variables as age, mental age and sex in relation to method of training (Hilgard et al., 1954; Osler and Fivel, 1961; Osler and Trautman, 1961), but this lead has not been a feature of discovery research. (However, cf. Corman, 1957). Cronbach (1966, p.90) hypothesises that "the interacting variables may have more to do with personality than with ability." He continues:

"I am tempted by the notion that pupils who are negativistic may blossom under discovery training, whereas pupils who are anxiously dependent may be paralyzed by demands for self-reliance."

It is interesting to relate such comments to the theoretical arguments of Harvey, Hunt and Schroder (1961),

to the empirical work of Heil et al. (1961), and to the study undertaken by Tallmudge and Shearer (1969). The latter summarise their findings (p.228) as follows:

"The study reported here produced results which strongly supported the existence of learning styles. These results also suggested that the correlates of learning styles are noncognitive, rather than cognitive, individual characteristics. Finally, they supported the notion that the nature of 'content' of the learning experience is a critical factor affecting the magnitude and direction of relationships existing between learner characteristics and instructional methods."

At this stage, individual characteristics have been neglected in the research on discovery. A major problem exists in selecting the most salient pupil characteristics for further investigation, for some characteristics may be task-specific, and others method-specific. Carroll's (1963) model of school learning would appear to allow for both categories, task and method, and for both cognitive and non-cognitive pupil attributes.

Design problems in the literature concern randomisation (or the lack of it) in assigning subjects, the question of relative time taken to reach criterion performance between treatments, and the legitimacy of the procedures used in both data and statistical analysis. Two general weaknesses noted are that of controls in naturalistic settings, and that resulting from overgeneralisation of findings. Cronbach and Wittrock both deal with these issues, citing Olson's (1965) critical note on Haslerud and Meyers' transfer-of-principles experiment as one demonstration of the experimenters' illogical analysis of the data. The problem of time in training

under different treatments has been a most persistent one. If criterion performance is to be reached with all pupils, it is likely that some individual learners in the discovery group will require a longer time than anyone in the directed group (but not always, as Gagné and Brown demonstrate). If time is held constant, the discovery group may fail to reach a criterion level which parallels that of the non-discovery group. Attempts to resolve this problem have led to doubtful uses of "difference scores" (see Cronbach's comments (1966, pp.82-83) on the Haslerud-Meyers study) to tasks which restrict time differences, and hence tend to become arbitrary, to different goals for the contrasted groups, and to such tight treatment controls that to extrapolate to typical school-learning situations becomes suspect. Cronbach (1966, p.84), while recognising that no recommendation will be appropriate for all studies, indicates his solution:

".....my inclination is to fall back on optimization within a fixed time."

Such a procedure seems to be consonant with what happens in the classroom, insofar as teachers feel bound to 'cover' a given syllabus of instruction in a given time.

Siegel and Siegel (1967) sustain a coherent and convincing case for a multivariate paradigm for research on instruction, and argue, in presenting their instructional-gestalt paradigm, for analysis of variance procedures. Change or difference scores are confounded by regression effects, measurement error problems, assumptions of equivalence of situation, mood and tasks, as well as scores.

Correlational studies

"would not provide information about interactions within the other independent variable clusters or between, say, learner and instructor variables." (Siegel and Siegel, 1967, p.275).

While earlier studies tended to use difference scores, comparisons of means, Chi squares (e.g. Kittell, 1957; Haslerud and Meyers, 1958; Kersh, 1962), later studies have tended to use analysis of variance (e.g. Ray, 1961; Guthrie, 1967) and analysis of covariance procedures. (Worthen, 1968).

In sum, the evolution of discovery research has been characterised by increasing precision and sophistication in definitions, design, and data and statistical processing. Conclusions are more cautiously drawn, in that few reports attempt to extrapolate beyond the subject area providing the experimental context. Covariance procedures permit intact groups to be used with some degree of control for group differences, and thus facilitate investigations in naturalistic settings. A further feature has been the increased range of dependent variables measured. At the same time, difficulties remain. The complexity of the treatments is such that the potent factors, or the combination of factors, have not been finally identified. The significant interactive individual difference factors require further delineation. Decisions about objectives, time, instructional strategies, appropriate criteria and measures confront each experimenter anew. At this stage somewhat molar research may still be necessary.

SOME TYPICAL FINDINGS.

Because of the thorough reviews (Ausubel, 1964; Kersh and Wittrock, 1962; Wittrock, 1963^a, 1966; Cronbach, 1966) already available, and because the extent of the literature makes an exhaustive survey of the research well-nigh impossible, this section will be confined in the main to studies reported since the mid-fifties, and especially to those subsequent to the publication of Shulman and Keislar's volume. Attention will be given to studies where one of the treatments involved a measure of guidance designed to lead the learners to generalise a rule from examples.

One rather general set of trends is evident in many of the conclusions reached by different investigators. Where the criterion is initial learning of a limited number of more or less specific answers, rather than the transfer or retention of what has been learned, the most highly directed groups (i.e. rule and answer given, expository) do as well as, or better than the other groups (Kittell, 1957; Kersh, 1958, 1962; Haslerud and Meyers, 1958; Worthen, 1968^a). Where the criterion is transfer, those groups receiving an intermediate amount of guidance, or those who derived principles from examples, perform as well as, or better than the groups given both rule and answer (Forgus and Schwartz, 1957; Kersh, 1958, 1962; Haslerud and Meyers, 1958; Gagné and Brown, 1961; Wittrock, 1963^a; Guthrie, 1967; Worthen, 1968^a). The results for retention measures are less clear, some investigators reporting advantages for rule-given groups (Craig, 1956; Guthrie, 1967), and others for example-given rule-derived

groups (Corman, 1957; Wittrock, 1963; Guthrie, 1967; Worthen, 1968). However, one must be cautious in regarding such generalisations as having any final research mandate, for, while an attempt has been made here to adjust for variable treatment labels, tasks, subjects and measures are not equivalent. Nevertheless, the trend is suggestive, as is the "post-experimental" gain reported for minimally directed groups whose retention scores exceed their initial learning scores. (Haslerud and Meyers, 1958; Kersh, 1958; Wittrock, 1963).

On the other hand, the literature in regard to affective outcomes is more difficult to locate, and the results are consequently more tentatively reported. Interest is represented in Kersh's studies, and Cronbach's comments (1966, p.88) are pertinent, and refer to Kersh's 1964 experiment.

"After sixteen training sessions there was no difference between didactic and inductive groups in tendency to use the information out of class."

Worthen (1968) administered a semantic differential and a statement attitude scale to discovery and expository groups who had been exposed to six weeks of training. No significant differences were found between the groups. This result gives some support to Cronbach (1966, p.88), who questioned Kersh's interpretation of the " sleeper-effect " of discovery learning as being attributable to motivation, and counter-proposed "I have long felt that this result can be attributed to novelty." Novelty, Hawthorne effects, deliberate structuring of one treatment to be a non-rational drill (cf. McConnell,

1934; Thiele, 1938; as cited by Cronbach, 1966, p.81) would contribute to the positive affective outcomes claimed for discovery approaches. The research to date provides little support for such claims. For example, Scandura (1964) implies Hawthorne effects when he reports (p.51):

"The E and D Ss apparently were highly motivated; only on the N-problems, and more frequently in the E-class, did any of the Ss appear to 'give up'."

Kersh and Wittrock (1962, p.468), concluding their interpretation of the learning by discovery research, decide that guided discovery may provide a most useful strategy.

"Guided discovery seems to offer a happy medium between independent discovery and highly directed learning. Some of the efficiency of directed learning is maintained along with the benefits of the discovery process, specifically, motivation and problem-solving skill."

There are a number of investigations which are, in part at least, concerned with guided discovery. Of these, the most frequently cited is the Gagné-Brown (1961) experiment, which led the learners in the guided discovery (G.D.) treatment by means of a small-step programme to the point where they were required to state a general rule. After the statement of the rule, the subjects were given practice in applying the rule to a number of specific examples. The G.D. treatment was compared with a rule-and-example (R. and E.) treatment, also using a small-step programme, and a discovery (D.) treatment using a large-step programme. Gains between learning sessions were significant for all conditions; and on criterion measures, transfer, time required, number of hints needed, and final performance, the results indicate

best performance for Condition G.D., worst for R. and E., intermediate for D. No retention measure was used. The researchers interpret their results as emphasising the importance of 'what is learned' as opposed to 'how it is learned' for problem solving, and suggest that the small-step G.D. programme requires "subjects to actively produce certain concepts, a feature which may be lacking in the R. and E. programme."

Roughead and Scandura (1968) utilised what was basically the same task with four programmes - specific rule given (R.), discovery (D.), guided-discovery (G.) and exposition of derivation rule (E.); in seven conditions - R. alone; R.D., D.R.; R.G., G.R.; R.E., E.R. Groups R. and R.D. performed at one level which was reliably ($P < .001$) below the common level of the other five groups on time-required, hints-needed and weighted scores on the within-scope and extra-scope transfer tests, although all groups had performed at an equivalent level on a criterion learning test. The investigators conclude:

"First, 'what is learned' during guided discovery can at least sometimes be identified and taught by exposition - with equivalent results..... The second point to be emphasized concerns the sequence effect - if a person already knows the desired responses, then he is not likely to discover another rule by which such responses may be derived, even if he has all the prerequisites and is given an opportunity to do so." (p.288).

The term sequence is gaining higher density as interest in instructional variables increases, and it is not always certain whether its referent is content-order; the position of rule relative to instances and examples (Werthen, 1968);

the placement of theory in regard to data (Leith and McHugh, 1967); 'timing' as an instructional tactic (Scandura, 1964; Taba and Elzey, 1964); or the path through learning set (Gagné, 1962). As far as the discovery research is concerned, the referent seems to be to the position of the rule or principle in relation to examples. What is learned, on the other hand, appears to relate to both task analysis (Gagné, 1962^b) and to such instructional manoeuvres as reinforcement, practice opportunities, attention focussing, feedback, set, cues and instructions. That Roughead and Scandura recognise the intricacies is demonstrated by two comments (p.288):

"..... there are undoubtedly a large number of situations where, because of the complexity of the situation, 'what is learned' by discovery may be difficult, if not impossible, to define."

"Why and how sequence affects 'what is learned' is still open to speculation (e.g. Guthrie, 1967; Yonge, 1966)."

Neither of the studies referred to above employed a retention test. In a number of experiments where retention has been reported (e.g. Craig, 1956; Forgas and Schwartz, 1957; Ray, 1961; Wittrock, 1962, cited by Kersh and Wittrock, 1962), the guided discovery groups have shown superiority to discovery and rote learning groups. Wittrock and Keislar (1965) demonstrated that either specific cues or concept cues produced significantly better performance on tests of learning, transfer to new instances, retention and a delayed test of transfer to new instances, than did very general cues. Guided discovery procedures may provide for cues, hints and questions, along with feedback, which influence both what is learned and how it is learned.

For example, Kornreich (1969, p.384), who like Kersh was interested in the strategies learned, found that:

"Significantly more Ss in Group G.D. acquired the strategy (focusing) than the other two groups ($p < .05$ for Group P; $p < .001$ for Group D, two-tailed test). Following up on observation, 20 more Ss were run. It was found that the effectiveness of the Group G.D. procedure was that it prompted S to reread the instructions."

Kornreich, in discussing his findings, made some pertinent observations which are relevant to this discussion. The programmed procedure appeared to change the subjects' choice of responses but not the strategy, and subjects showed dependence on the experimenter's comments. Kornreich adds:

"Optimal prompting may differ for different criterion performances, so that a different technique may be optimal for transfer tests than for retention tests." (p.388).

Anderson, Faust and Roderick (1968) report that overprompting produced significantly lower post-test results than did a standard programme with adult subjects. Wittrock has reported in several studies that class cues "can produce attention and also transfer to other instances within the class." (Wittrock et al., 1964). Wickelgren (1964), too, noted that cues in the form of questions and hints could direct students into methods or strategies which generalise to transfer problems. Corman (1957) was also of the opinion that appropriate guidance was beneficial for learning and transfer, although he differs from Ray in considering that there may be an interaction between method (amount and kind of information) and learner ability.

Wittrock (1963a) provided some coherence and order amongst the diversity of experimentation:

"When the criterion is initial learning of a few responses, explicit and detailed direction seems to be the most effective and efficient. When the criteria are retention and transfer, some intermediate amount of direction seems to produce the best results." (p.189).

In general, more recent studies mirror, as one would expect, research interests and issues current in educational psychology. The treatments described reflect interest in instructional theory (e.g. Gage, 1964^a; Gagné, 1965^d; Carroll, 1965; Stolurow, 1965; Bruner, 1966), and may be considered in terms of strategies and tactics, with increasing attention being paid to such manipulable variables (see p. 37 above) as task form, sequence and nature, and instructions, cues, set, feedback and timing. Attention (White, 1963) remains a variable awaiting a formal place in the studies. The intervening variables similarly echo theoretical interests in mediational processes - information processing; association, conceptual and principle learning; strategies; assimilation and accommodation; cognitive cost. Not so evident are earlier dichotomies such as meaningful versus rote learning. Dependent variables remain bounded by requirements of measurement, although multiple criteria are more common and reflect the influence of writers like Tyler and Bloom. As Cronbach (1966, p.90) points out, many educationally-valuable objectives await evaluation; the matching of pupil characteristics and instructional techniques awaits further investigation.

Learning by discovery remains an hypothesis. Cronbach

(1966) suggests a need for more complex experiments,
involving

"a five-fold interaction - subject matter, with
type of instruction, with timing of instruction,
with type of pupil, with outcome." (pp.91-92).

CHAPTER IV

REVIEW OF METHODS: EXPOSITORYRATIONALE AND THEORY.

Although expository teaching is a widespread teaching practice in schools and colleges, it has attracted comparatively little research (Carroll, 1964). Ausubel (1963a) believes that the reason for the lack of research is that expository teaching has been identified with rote learning; a point Cronbach (1966) appears to support with his remark that most experiments on discovery have "stacked the cards against" expository teaching. A number of terms have been used for similar but not identical approaches: deductive teaching, didactic teaching, Ruleg sequence, and, pejoratively, authoritarian teaching. De Cecco (1968, p.468) defines expository teaching, using Wittrock's 1963 classification of teaching methods, as "the situation in which the teacher gives both the principles and the problem solutions." Ausubel (1967,1968) defines it, after conceding that a conceptual distinction can be made between learning and teaching (Smith, 1960), in terms of those practically-useful, reciprocal relationships between learning and teaching. Thus expository teaching can be viewed as the reciprocal of meaningful, reception learning.

"That is, the principal content of what is to be learned is typically presented to the learner in more or less final form. Under these circumstances, the learner is simply required to comprehend the material and to incorporate it into his cognitive structure so that it is available for either repro-

duction, related learning or problem solving at some future date." (Ausubel, 1968, p.83).

Before considering the theoretical network providing the justification for didactic teaching in Ausubel's sense, it seems appropriate to review briefly the rationale for his position. First, Ausubel conceives of man as a rational and symbolising entity, characterised by language use.

".... the human capacity for representational symbolism and verbalization make possible both:
 (a) the original (discovery) of ideas at a uniquely high level of abstraction, generality and precision, and
 (b) the cumulation and transmission of these ideas during the course of cultural history." (1968, p.82).

Second, formal education is distinctively concerned with intellectual training, with fostering intellectual growth, and with transmitting worthwhile subject-matter knowledge which is meaningful, consonant with contemporary scholarship, and developed to differential levels of individual excellence.

"Hence in setting our academic goals, we must be concerned with the ultimate intellectual objectives of schooling, namely, with the long-term acquisition of valid and usable bodies of knowledge and intellectual skills, and with the development of ability to think critically, systematically and independently." (1968, p.31).

Third, such knowledge has value in its own right.

"Meaningfully organized subject matter taught by competent teachers can generate considerable drive for learning as an end in itself. The value of much school learning, after all, can be defended only on the grounds that it enhances pupils' understanding of important ideas in their culture - not because it has, even remotely, any practical use or implications. Nevertheless, some aspects of academic training do constitute, in a general way, just as important a preparation for adult living as education that is explicitly directed toward vocational and family adjustment." (1968, p.31).

Fourth, educators and teachers cannot abdicate their responsibility to stimulate the development of interests, motiv-

ations and intellectual needs to structure curriculum content and to teach children to learn by themselves. This involves a matching of instruction to pupils' prior experiences, level of cognitive maturity, and body of meanings, by means of the selection, organisation, interpretation and sequential arrangement of learning materials and guidance of learning experiences.

"Such needs, however, are not endogenous, but acquired - and largely through exposure to provocative, meaningful and developmentally appropriate instruction." (1968, p.33).

From this perspective, Ausubel developed a theory of meaningful reception learning which has guided an increasing amount of research. As might be anticipated from the rationale outlined above, the theory assumes a number of distinctions (e.g. between reception and discovery learning, between rote and meaningful learning, and between logical, potential and psychological meaning). The research revealed some similar general features in that the outcomes most frequently measured have been learning and retention, the subjects have been of junior high school age and older, and the content has been largely propositional and conceptual.

Gagné's review (1969; pp.287-290) of Ausubel's 1968 volume includes a concise summary of the general theory.

"The core of his theory is the meaningful learning of ideas, which is, according to this author, what the vast preponderance of school learning is all about. The learning of ideas takes place when a novel idea, usually stated as a verbal proposition, is assimilated into an existing cognitive structure. One variable influencing learning and retention is the availability in cognitive structure of specifically relevant anchoring ideas. Such ideas may operate to bring about subsumption of a new idea, either derivative

subsumption (when the new idea is a specific example of the anchoring idea), or correlative subsumption (when the new idea is an elaboration or modification of the anchoring idea). A factor influencing the retention of the newly subsumed idea is the extent to which it is discriminable from the anchoring ideas that assimilate it; in other words, its degree of dissociability. Immediately following the learning dissociability is high, but at this stage it is followed by increasing obliteration as the assimilation process proceeds. Finally, learning and retention are functions of the stability and clarity of the anchoring ideas. Within such a theory, rote learning becomes a highly special form of learning that relates to cognitive structure only in arbitrary, verbatim fashion, without the kind of assimilation that applies to meaningful learning. Learning of meaningful material, however, may be acquired by discovery or by reception....."

While such a summary is useful, the import of Ausubel's message is clarified further as one takes account of some of the distinctions and presuppositions. As far as the discovery-reception distinction is concerned, Ausubel regards each as capable of being a rote or a meaningful form of learning, but each differs in usefulness according to the objectives pursued, time requirements, the type of learning task and the cognitive maturity of the learner. A temporal distinction is also made:

"The distinctive and prior learning task, in other words, is to discover something - which of two maze paths leads to the goal, the precise nature of a relationship between two variables, the common attributes of a number of diverse instances etc. After this phase is completed, the discovered content is internalized just as in reception learning." (Ausubel, 1966, p.158).

Meaningful learning presupposes that the learner manifests a set to learn meaningfully rather than to internalise in arbitrary and verbatim fashion (i.e. rote), and that the learning task itself is potentially meaningful.

Thus optimal conditions involve a match between a particular learner's cognitive structure, and the extent to which there inheres in the learning task both some reasonable basis for relating the new material to the learner's ideational scaffolding (nonarbitrariness), and some symbolic equivalence to the learner's organisation of established meanings (substantiveness). Psychological meaning so conceived is a consequence of an idiosyncratic and phenomenological transaction. Clearly, Ausubel is interested in conceptual learning. But, unlike many educational psychologists with similar interests, Ausubel has displayed a research concern with verbally rather than ostensively defined concepts, and with the process of concept assimilation rather than concept formation. (Ausubel, 1966, pp.163-167).

This concern is consistent with his discussion of developmental considerations to be taken into account when teaching concepts. The argument is twofold insofar as it is based on verbal ability and Piaget's stage theory. The pre-operational child, limited in ability to deal with symbolic representations in acquiring concepts, can only discover their criterial attributes by overtly manipulating diverse instances of objects or events, using subverbally the necessary conceptualising operations of abstraction, hypothesis testing, differentiation, and generalisation. During the stage of concrete operations

"the learner is able to manipulate new relationships between verbally expressed ideas, and hence can assimilate concepts, providing that he has some recently prior concrete exposure to particular exemplars of the concept in question." (Ausubel, 1966, p.166).

At this stage, discovery probably enhances the intuitive meaningfulness of new concepts as a result of bringing the learner into more intimate contact with the concreteness and specificity of the experience upon which such meaningfulness depends. To a lesser degree, it is claimed, the same situation holds for older learners when initially exposed to difficult, unfamiliar, new concepts. However, at the formal-operational stage of concept development, relationships between abstract and general ideas can be assimilated (or discovered) without dependence on overt manipulation, or concrete and particularised experience. It is at this final stage that Ausubel believes concept assimilation is most appropriate, and that meaningful reception-learning constitutes an efficient primary means of acquiring large bodies of subject matter knowledge. This stage is equated with that reached by Junior High School students. While there are no doubt real differences in interpretation, some parallels may be glimpsed between the Piagetian model, Bruner's enactive, iconic and symbolic modes of representation, and Ausubel's manipulative, verbal-concrete and abstract sequences.

For Ausubel, knowledge is an ideational phenomenon, not a capability of performing different classes of tasks (e.g. Gagné, 1965^o), nor sets of associations, or habit-family hierarchies (e.g. Berlyne, 1965). An outcome of this viewpoint is Ausubel's contention that the principles of proactive and retroactive interference are inapplicable as explanations of the forgetting of meaningful verbal material. This is better explained as a cognitive process of subsumption, or

reduction to an ideational common denominator. From a similar perspective, Ausubel (1965) distinguishes and relates perceptual and cognitive processes in meaningful verbal learning, attributing to repetition a telescoping effect on the two processes, and a consolidating effect on retention. It is, then, not surprising to find the originator of this theory deploring the separation of educational psychology and classroom learning, to find that he regards the simple laboratory studies of learning as having doubtful relevance to the qualitatively different learning in schools, and to find his expressed belief that a theory of classroom learning is the most feasible point of departure for discovering general principles of teaching.

PEDAGOGICAL IMPLICATIONS.

It follows from the general theory that two major classes of variables are significant in influencing the learner's capability of acquiring new knowledge in a given field. First is cognitive structure itself - the substantive relevant content of the learner's knowledge in a particular domain at any given time, and its organisation, stability and clarity, which provide anchorage and discriminability. Second is the programming of the new material, its selection, sequential arrangement, organisation and internal logic, presentation and the associated practice, feedback and over-learning provisions.

In this context, specific readiness is reinterpreted

in terms of cognitive structure variables. Where the learner has few anchoring subsumers available in cognitive structure, or where these lack discriminability, advance organisers become the major compensating strategy.

"The principal function of the organizer is to bridge the gap between what the learner already knows and what he needs to know before he can successfully learn the task at hand." (Ausubel, 1968, p.148).

With completely unfamiliar material, expository organisers provide the relevant proximate subsumers giving ideational anchorage in terms already familiar to the learner. With relatively familiar material, comparative organisers are used to integrate new concepts with basically similar concepts in cognitive structure, as well as to increase discriminability between new and existing ideas. Organisers serve several functions, providing set and ideational scaffolding, and enhancing discriminability. Following the principle of progressive differentiation, organisers are introduced in advance of the learning material itself, and are presented at a higher level of abstraction, generality and inclusiveness than the new subject-matter.

A further principle applying to the programming of instruction is that of integrative reconciliation,

"best described as antithetical in spirit and approach to the ubiquitous practice of textbook writers of compartmentalizing and segregating particular ideas or topics within respective chapters or sub-chapters." (Ausubel, 1967, p.240).

The principle involves a serious teaching effort to explore explicitly relationships among related ideas, to point out significant similarities and differences, and to reconcile real or apparent inconsistencies, so that new ideas are

comprehended and interpreted (integrated) in terms of existing understandings and paradigms provided by analogous, familiar, previously learned and already established ideas in cognitive structure. Organisers may further the principle of integrative reconciliation.

Learning is enhanced when the task manifests internal logic, defined by Ausubel (1967, p.243) as "a function of the plausibility, lucidity and nonarbitrariness of the material, rather than of its logical or substantive validity."

He elaborates:

"At least four aspects of the internal logic of material affect the extent to which it is endowed with potential meaning:

1. Adequacy of definition and diction, including precise, consistent and unambiguous use of terms; the definition of all new terms prior to use; and the simplest least technical language that is compatible with conveying precise meaning.
2. Use of concrete-empirical data and relevant analogies when developmentally warranted or otherwise helpful in the acquisition, clarification, or dramatization of meaning.
3. Encouragement of an active, critical, reflective, and analytic approach on the part of the learner by requiring him to reformulate the material in terms of his own vocabulary, experiential background and structure of ideas.
4. Explicit delineation of the distinctive logic and philosophy of each subject-matter discipline - that is, its implicit epistemological assumptions; the general problems of causality, categorization, inquiry and measurement that are specific to the discipline; and the distinctive strategy of learning how to learn the particular subject-matter of the discipline."

All the factors referred to above - advance organisers, consolidation, progressive differentiation, integrative reconciliation, internal logic, and potential meaningfulness - are assumed to have facilitating effects on learning and retention, largely through their enhancement of the major,

cognitive-structure variables of the availability, discriminability, stability and clarity of the appropriate, relevant subsumers. The task and practice variable postulated to influence the efficiency of learning is structured practice which involves careful sequencing, pacing and gradation of difficulty; differential application; skilled presentation and organisation of material. The effects of frequency are direct (successive trials cumulatively modify cognitive structure by summing the influences of contiguity and feedback, enhancing the dissociability strength of meanings) and indirect (alterations in cognitive structure resulting from earlier trials affect learning and retention processes during subsequent trials). Motivation, not in this theoretical context an indispensable for learning, is regarded as a cognitive drive reciprocally related to meaningful learning.

"Typically, however, motivational and attitudinal variables are not directly involved in the cognitive interactional process. They energize and expedite this process during learning by enhancing effort, attention and immediate readiness for learning, and thereby facilitate dissociability strength catalytically and nonspecifically." (Ausubel, 1967^a, p.253).

On the grounds of their appropriateness for education, Ausubel questions achievement-oriented, ego-enhancing and anxiety-reducing motivational constructs. For him grades are intellectually stultifying; the important goal is to know and to understand.

So far, Ausubel's theory of meaningful reception learning has been described somewhat uncritically. He frankly admits the necessity for much more supporting research evidence, and for more studies to test the theory. His comments, in

regard to guided discovery, are indicative (1968, p.304):

"Numerous short-term studies have demonstrated that guided discovery is more efficacious for learning, retention, and transfer than is either completely autonomous discovery or the provision of complete guidance. However, these findings do not necessarily indicate that guided discovery is more effective for teaching subject-matter content than is simple didactic exposition."

While recognising the plausibility of the theory insofar as it refers to some categories of human learning, it does seem that the benefits arising from expository teaching may apply more to the learning and retention of substantive content than to problem solving and lateral transfer.

"Actually, the principal effect of existing cognitive structure on new cognitive performance is on the learning and retention of newly presented materials where potential meanings are given - not on the solution of problems requiring the application and reorganization of cognitive structure to new ends." (1968, p.130).

In view of the outline to this point, it is to be expected that Ausubel would acknowledge the probability of types of learning (Melton, 1964; Gagné, 1965), for his whole argument is premised on the assumption that substantive cognitive outcomes are the sine qua non of the school, and economically (Carroll, 1967) and educationally, such goals are most effectively achieved through meaningful reception learning, particularly with cognitively mature learners. Some difficulties arise when exploring any such theory - recognition of the field-of-reference boundaries, interpretation of intentions (sometimes disguised in system-specific vocabulary), clarifying assumptions and criteria, finding legitimate bases for comparison with other viewpoints, and

grasping the practicability of the theory for prescription as well as explanation. The theory appears to meet criteria of coherence and consistency, but it does raise problems, for example, in identifying subsumers in individual cognitive structures, in the requirements of pedagogical skills (cf. Rudin's 1961 study of the teacher's effectiveness as a lecturer), in regard to the everpresent assumption, in paradigms, of linearity, and to questions about collateral objectives.

"As in the case of discovery learning" remarks de Cecco (1968, p.468), "it is probably difficult to find pure examples of expository teaching." It is also difficult to find many clear examples of a range of organisers as described by Ausubel. While he provides definitions of important constructs in the system-specific terms of his theory, operational definitions are not so frequent. The Ruleg sequence (Glaser, 1966) in programmed instruction, meets some of the requirements Ausubel stipulates for expository teaching, as long as there is compatibility between the learner's cognitive structure and the potential meaningfulness of the rule. Not too dissimilar, claims de Cecco (1968), is the general strategy of deductive teaching in which the teacher often begins with a definition of the concepts or principles, illustrates them, and unfolds their implications.

Bibergall (1966, p.230), concluding her survey of discovery learning, suggests:

".... the expositional technique would be most profitable in the following cases
 (i) for specific (subject) matter aspects,
 (ii) when the student is at the abstract level of development, and
 (iii) for quick initial learning."

EMPIRICAL EVIDENCE.

Most of the research to date on expository teaching, as discussed, has been undertaken by Ausubel and his associates, although a number of studies relating to discovery (e.g. Roughead and Scandura, 1968; Worthen, 1968;^{3b} Wittrock, 1963) and to instructional theory and methodology (e.g. Nuthall, 1966; Leith and McHugh, 1967; Collis, 1969), apart from those designed to test aspects of programmed instruction, bear more or less directly on didactic instruction. The value of advanced organisers in facilitating meaningful verbal learning has been demonstrated in several experiments (Ausubel, 1960; Ausubel and Fitzgerald, 1961, 1962; Ausubel and Youssef, 1963; Wittrock, 1963b).

Of the Ausubel studies, four (Ausubel, 1960; Ausubel and Fitzgerald, 1961, 1962; Ausubel and Youssef, 1963) are representative, and continue the earlier (Ausubel, Robbins and Blake, 1957; Ausubel and Blake, 1958) research programme, employing similar strategies. Later studies (Ausubel, Stager and Gaito, 1968, 1969) are similar to those cited, in materials used, in that the subjects (Grade 13 students as against senior undergraduates) are 'cognitively mature', and in the general strategy used for testing the facilitating effects of advance organisers. The advance organisers, introduced (in the case of expository organisers) after establishing an essentially zero baseline for the content in cognitive structure, were specially prepared passages. It has been a feature of the research to use passages which subjects read, rather than the more typical classroom situation where the teacher

presents introductory materials orally. Another feature of the studies has been in the kinds of dependent measures used - initial learning and retention have been tested almost exclusively.

Two of the studies (Ausubel, 1960; Ausubel and Fitzgerald, 1961) indicated that, when the presented material had no relation to known concepts (e.g. a text on metallurgical properties of carbon steel), and when the presented material (e.g. a text on Buddhism) can be related to known concepts (e.g. concepts in Christianity), the provision of an advance organiser (expository and comparative respectively) can be similarly facilitating. Further evidence (Ausubel and Fitzgerald, 1962) indicates that the effect of the introductory passage is most marked for students with low verbal ability (S.C.A.T. scores). The conclusion is suggested that

"subjects of average and better ability are evidently capable of spontaneously organizing new material around relevant, more inclusive concepts, and hence derive little or no benefit from introduced advance organizers." (Ausubel and Fitzgerald, 1962, p.247).

Wittrock's (1963c) study is provocative in reference to the one just cited. The subjects, Grades 10, 11 and 12 students, were assigned on a stratified random basis according to sex, I.Q. and grade level to one of three experimental conditions, sets to notice differences, to notice similarities or to read and remember the passage (neutral). The above-median-I.Q. subjects under the neutral condition

"tended to recognize more differences than did the less intelligent subjects. Further, the differences instructions, but not the similarities instructions, produced a decided effect with the less intelligent subjects." (p.74).

If, as indicated, less intelligent pupils tend to recognise similarities, and more intelligent pupils, while being more flexible in response to directions, tend to note differences, the set-providing function of suitable advance organisers may be expected to increase discriminability. Unfortunately, results by sex are not reported.

Gagné, in a recent study (1969), derived further supporting evidence for the efficacy of organisers in regard to retention. He states (p.413):

"The second major conclusion relates to the effects of different kinds of context on the recall of facts. A context which is superordinate, introducing context facts with a general topic sentence, facilitates recall of facts to a greater extent than does a coordinate or unrelated context. The coordinate context, containing facts related to the fact to be remembered, also leads to greater recall than does an unrelated context. Both these findings are consistent with the theory of Ausubel (1968) to the effect that retention of meaningful facts is improved when efforts are made to mobilize anchoring ideas within a pre-existing cognitive structure. Somehow, even in the face of interference, the organization of facts into topics operates to facilitate retention of particular facts."

It may be that some general qualities of organisers will provide different kinds of set. Tuckman et al. (1968, pp.66-67) report an unexpected inference from their results:

"the strategy of search could be made more readily to transfer than the skill of search, as the result of limited prior experience."

Reynolds (1968, pp.133-138), dealing with verbal-perceptual material, found that

"pre-familiarization with a single integrated map structure provided greater transfer to sentence learning than did pre-familiarization with the same map which had been fragmented into separate and discrete pictures."

Additional work suggestive of support for an Ausubelian

mode of expository teaching has been reported by Collis (1969) and by Leith and McHugh (1967). The former, whose subjects were "at the early secondary school level", used mathematics textbook material which had been organised according to Ausubel's themes of progressive differentiation and integrative reconciliation. While the mathematics criterion tests revealed no significant advantage for the experimental group over the control group after one year of the programme, there were significant gains in attitude for the children using the experimental textbook. Leith and McHugh investigated the effect of varying the order of the main passages, one of which, the theoretical passage, was designed to act as an organiser. The organiser (T) occupied the introductory, mediating and summing positions relative to two passages, one of which, the patrilineal (P), was designed to be readily assimilated; the other, the matrilineal (M) being much more difficult. Six treatments evolved from the ordering of the main passages, and, with the remaining two conditions omitting the theory passage, a total of eight groups were included in the design. The group having the order M.T.P. performed best on all tests of both P and M; however, it is noted that without facilitation the recall of the M passage was poorer than with facilitation. The investigators consider (p.116):

"Ausubel's recommendation of the use of an advance organiser has been justified in the case of material which does not conflict with established cognitive patterns. But the advance organizer is no better than either a mediator between passages or a summary in this case."

Their conclusion in regard to the place of theory is, then:

"In the conditions of this experiment it would seem to be after the learning of particular concepts rather than before the, especially when the more difficult exemplar is given first."

Thus, while facilitating effects were found for the theory passage, the researchers raise questions related to sequencing.

Another set of studies giving varying degrees of support to the effectiveness of expository teaching have been those which have also had reference to discovery learning. Scandura and his colleagues investigating mathematical learning, have provided some evidence, particularly insofar as rule provision is an aspect of didactic teaching. One study (Scandura, Woodward and Lee, 1967) "demonstrated rather conclusively the behavioral relevance of rule generality", in that the subjects given the most general form of the rule achieved greater within-scope transfer. Roughead and Scandura (1967, already cited) emphasise two conclusions. First, what is learned during guided discovery learning can at least sometimes be taught by exposition, and with equivalent results.

"As we identify just what it is that is learned by discovery in more and more situations, we shall be in an increasingly better position to impart that same knowledge by exposition." (Scandura, 1968, p.339).

Second, where the sequence is to provide the rule before giving opportunities to discover, pupils perform no better than if they had been given the rule only.

"In effect, prior knowledge may actually interfere in a very substantial way with later opportunities to discover." (Scandura, 1968, p.339).

Scandura's basic argument appears to be that expository teaching can be much more effective than it typically is, if it can be specified clearly, through structural analysis, what it is that is to be learned. Nevertheless, Scandura sounds a cautionary note (1964, p.155):

"Comparisons between exposition and discovery necessarily involve the simultaneous variation of several variables and as such cannot be taken as definitive, no matter what the results."

A number of experiments have produced results in which the learners receiving expository instruction have achieved significantly more highly in initial learning tasks than have contrast groups. (e.g. Kersh, 1958,1962; Wittrock, 1963^c; Worthen, 1968^d). In many studies the expository group has also manifested superiority when time or trials to master learning has been the criterion. (e.g. Wittrock, 1963^c; Guthrie, 1967). Ausubel (1968, p.215), discussing the extent to which intellectual development can be accelerated, expresses the opinion that didactic procedures together with concrete-empirical props can be helpful.

"Considerable evidence, however, indicates that the use of various verbal didactic procedures (prior verbalization of principles, the use of verbal rules, filmed verbal explanations, confronting the child verbally with his own contradictions), in conjunction with concrete-empirical props, can accelerate the acquisition of conservation and probability theory (Frank, in Bruner, 1964; Kohnstamm, 1966; Ojemann and Pritchett, 1963; Ojemann, Maxey and Snider, 1966; Sullivan, 1966)."

An examination of Ausubel's (1968, pp.147-162) summary of transfer and its pedagogical facilitation reveals, unexpectedly, the number of times research related to retention rather than transfer is cited. This suggests an interpretation

of transfer in a within-scope sense more than an extra-scope sense, and may have some bearing on the typical use of retention as a criterion measure in research on exposition. Although much evidence can be found for the transfer value of understood general principles (e.g. Hilgard et al., 1953, 1954), Scandura (1968, p.336) expresses some doubt insofar as exposition is concerned.

"When rules are presented in an expository fashion, it is normally too much to expect generalization to problems to which the principle does not immediately apply."

The evidence tends to provide strong support for the benefits of expository procedures in regard to initial learning scores and time to criteria, less consistent results for retention, and ambiguous results for transfer (cf. Worthen, 1968^a, p.1). Different criteria make results hard to compare; nor is it clear just what aspects of expository teaching have greatest potency for given objectives. Smith and Smith (1966, pp.324-325), reviewing research on transfer from programmed learning, make some interesting observations which have, at least, tangential pertinence to this discussion. Gagné and Dick (1962) found transfer test scores very low indeed following programmed instruction on equations, although the same subjects had scored reasonably well on verbal and performance post-tests.

"Cartwright used systematic and unsystematic sequences of frames to teach fractions to mentally retarded adolescents (Stolurow, 1963). The groups using the systematic sequence remembered more of the specific facts taught, but the group using the unsystematic sequence transferred more of their knowledge to other problems. Krumboltz and Bonawitz (1962) tested the effect of presenting confir-

ming responses in a program as complete thoughts, usually be repeating the relevant parts of frames with the desired responses inserted, instead of as isolated words or phrases. Their groups showed no difference in knowledge of terminology, but the context group excelled in ability to apply principles." (Smith and Smith, 1966, p.325).

Pertinent, too, is a recent study by Rowell, Simon and Wiseman (1969), in which verbal reception and guided discovery techniques were compared, in a classroom situation, for their effectiveness in promoting the formation of stable, usable, cognitive schemata in comparable groups of university students. The results indicated that explicit direction can produce performances superior to those resulting from a guided discovery approach for immediate recall, delayed retention and transfer ($p = .05$ or less). The investigators, in their discussion, point out the relatively slow but steady decline in performance on successive tests of the verbal reception group. This decline is one which might be explained as obliterative subsumption, as there seemed to be, after ten weeks, retention of "the overall structure" and loss of detailed information. Two factors, it is suggested, could have contributed to the better performance of the directed group, the imposed time limit on the experiments, and the learning history of the subjects. One further result is given special attention by the experimenters, viz. the relative results of the two groups on Application Test 3. In terms of absolute scores on this test, the guided discovery group actually improved, whereas the scores of the verbal reception group continued to decline. While the latter's scores were still superior to the former's, the former showed marked relative improvement. The authors

conclude that, subject to the experimental limitations, the results of their study, taken as a whole, suggest that instruction techniques of an expository or verbal reception nature can have greater pedagogic merit than has sometimes been assumed.

"This, of course, is no longer a unique conclusion: a number of writers have expressed basically similar views." (Rowell et al., 1969, p.242).

CHAPTER V

DEFINITIONS, CRITERIA AND HYPOTHESESRESEARCH CONTEXT.

The respective reviews of literature and research have referred to theoretical orientations, and the discussion of the general research problem carried overtones of theory. As so many educational psychologists (e.g. Gage, 1964^a; Bruner, 1966) have pointed out, there is no theory of teaching to provide an integrating network synthesising the relationship between instruction, learning, and individual and developmental characteristics. Bruner (1966, pp.39-72), in delineating his prolegomenon to a theory of instruction, indicates the preliminary nature of his attempts. Tentatively, then, and at a very general level, it is here suggested that the teacher's task, as far as cognitive learning goals are concerned, is that of manipulating the cognitive gap, arranging the match, between the learner's present status and the substance of the experiences to be provided for him to achieve the goals. Instructional strategies are the means intended to facilitate the achieving of goals, and, no doubt, different strategies set different task requirements, different data processing conditions. Ausubel argues for assimilation following a contiguity of meanings; Suchman for accommodation following a discrepant event, a discontinuity of meanings. Each assumes as given, cognitive structure, active data processing and a resolution either by subsumption or reconcil-

iation. Kagan posits preferred modes of processing. All share a general cognitive perspective. It is in this broad context that this present study is undertaken.

COGNITIVE STYLE.

For the purposes of this study, cognitive style is defined at the general level, as a preferential mode of categorisation expressed in a situation where alternatives are possible. (Sigel et al., 1967). The situation is that provided by a Kagan-style Triads test, in which groups of three stimulus-pictures are presented in booklets to the subjects, whose task is to select two from each group of three pictures, on the basis of perceived cue properties which suggest to the subject the way in which the chosen two "are alike or go together in some way." Alternatives are provided insofar as each triad can suggest several pair groupings. The mode of categorising is indicated by the reason the subject is required to state for each such grouping, the written reasons being classified according to criteria set out below, and the resulting three scores being separately summed to give a measure of cognitive style in the three previously identified categories: descriptive, inferential-categorical, and relational-contextual. (Kagan et al., 1963). The criteria outlined below are adapted from those given by Kagan, with some simplification and refinement to increase

clarity and consistency of application (cf. Sigel et al., 1967; Wallach and Kogan, 1965; Gordon, 1966; Edwards, 1968).

Scoring Criteria:

Descriptive.

Groupings in this category are based on similarity in objective, physical characteristics among a group of stimuli. The concrete attribute shared by objects can be perceived through any of the senses, although visually perceived facets of the stimuli are expected to dominate. All descriptive labels must contain a reference to a commonly-shared, physical attribute of the grouped stimuli. Two sub-classes have been identified (Kagan et al., 1963) - descriptive-analytic, where labels denote observable parts of the figure, and descriptive-global, in which labels refer to similarities based on the total objective manifestations of the stimuli. Examples of descriptive-analytic categorisations referring to animal stimuli are: 'they both have tails, hair, four legs, one ear drooping'; human stimuli are: 'both have guns, wear shoes, have hands on hips, wear skirts'; object stimuli are: 'both contain liquid, have a leg missing, have a lense, have handles'. Examples of descriptive-global labelling (which were actually infrequent) are: 'both have the same shape, are in the same position, are the same size'.

Inferential-categorical.

These concepts classify objects because of some characteristic shared by all, but what they share is not inherent in the physical nature of the stimuli grouped. The labels do not contain a direct reference to an objective, physical attribute of the group of stimuli (unlike "descriptive" labels), and yet each object is an individual instance of the label (unlike "relational" concepts). The classifications frequently fall into the area of location, usage or superordinate class which subsumes the particular instances. Location examples would include: 'both are found near water', 'both live in water'; usage examples are: 'for eating', 'both are for travelling on or in'; superordinate classification is exemplified by: 'both are animate', 'both are forms of footwear'.

Relational-contextual.

The distinguishing characteristic of these conceptual labels is that they grow out of the relationship between or among the stimuli grouped together, and then serve as a kind of umbrella over the collection. Because of this, no stimulus is an independent instance of the concept, and some have greater weight in determining what the concept became than did others. These relationships among the stimuli grouped together are functional, and

build on connections of a temporal, spatial or complementary nature. The temporal-spatial complexes are concrete or situational in nature, generally thematic. Typical examples are: 'a dog lives in a kennel', 'the father and mother had a baby', 'the thread goes through the needle'. Sigel et al. (1967, p.2) sum up by defining these responses as "indicating interdependence in that particular situation expressed in functional or thematic labels."

Thus, operationally, cognitive style is defined by subjects' responses on a twenty-five item triads measure, based on Kagan's conceptual styles test (Kagan et al., 1963; Sigel et al., 1967). In the interests of consistency, the three modes will be more briefly termed descriptive, categorical and relational in the remainder of the report. Because so few descriptive-global responses were made, it would not have been inappropriate to use the label 'analytical' for this style. However, 'descriptive' is probably more accurate overall. It was noted that the children's statements used in formulating reasons for choices of pair-groupings, revealed language patterns related to the mode. These patterns showed typical characteristics. For example, categorical choices were usually phrased "they both are...."; descriptive choices usually appeared as "they both have....", or "they have both got...."; and the relational responses manifested such verbal forms as ".... wears", or ".... killed". Three of the frequent responses to Item One are illustrative: "A man wears a watch" (relational); "They both have numbers" (descriptive); "They are both measures" (categorical).

TREATMENTS.

The treatments, labelled in the literature Guided Discovery and Expository, or those which are modifications of Example-Rule and Rule-Example approaches, are sufficiently complex for no brief definition - operational, constitutive or stipulative - to be completely adequate. Already, problems of naming the treatment independent variables have been discussed with reference to semantic consistency and tautology. As long as assumptions about intervening variables are not made because of inferences possible from the naming of the independent variable, the name given to the treatments may be less important than the defining criteria. Furthermore, few of the recent investigations have avoided naming the treatments as discovery or expository modes (e.g. Roughead and Scandura, 1968; Guthrie, 1967; Worthen, 1968; Rowell et al., 1969). However, greater care has been exemplified in providing clear descriptions of the treatments. Here, an attempt is made to define the treatments by criteria guiding teacher responsibility and task requirements of pupils, an approach similar to that employed by Worthen. There are, nevertheless, some differences intended to make the procedures, while still independent, more typical of life in the classroom than those described by Worthen. At the same time, serious effort has been made to make each treatment as effective a teaching procedure as possible. To ensure individual rule derivation, Worthen's discovery pupils were prevented from sharing ideas; if errors were made, the teacher said nothing about the rule or why the answer was incorrect. In most classrooms in the

writer's experience, these are not common features. Glick (1968) in fact suggests that some pupils (target) mediate the instruction for others.

As it is probably impossible to specify all the distinguishing characteristics of each treatment, and as it is likely that any particular component has potency in association with and in the context of all other components, the defining criteria given are those which have been given some significance in research and theoretical writings. The labels to be given the treatments, following Wittrock, are descriptive, and conform to their functions as independent variables. These labels, then, are Rule-Explained-and-Demonstrated (R.E.), and Rule-Derived-with-Guidance (R.D.).

Treatment Criteria.

	<u>R.E.</u>	<u>R.D.</u>
Sequence (in general and for each <u>phase</u>).	Principles (Rules), generalisations and concepts presented by teacher. Examples and attributes explained and demonstrated with apparatus. Terms presented concurrently with explanation and demonstration. Pupils practise, apply principle to type situations, to apparatus. i.e. (a) Teacher explanation of principles. Terms given in context, with demonstration and explanation. (b) Pupils practise with examples.	Children observe, manipulate and discuss; and through examples, directions to note attributes, problems and use of apparatus, formulate generalisations, principles (rules) and concepts. Labels and terms given as pupils begin to verbalise. Hints and cues given to assist derivation of rules etc. i.e. (a) Pupil activity to derive principles, rules etc. Teacher guides activity. (b) Pupils formulate rules etc. Teacher provides terms in context.

R.E.R.D.

Cues, provision of information.	Teacher is the source of information, explains in detail and as clearly and simply as precision permits. Uses concrete-empirical props to demonstrate, of same kind as used by pupils in treatment R.D.	Teacher avoids telling directly, but by questions, hints and cues, leads children to formulate and provide answers. Indirect rather than direct strategy. Cues, as far as possible, not <u>too</u> specific.
Teacher response to pupil questions	Refers children to rules explained earlier. Ensures 'meaning' by examples or analogy where necessary. Suggests recall of explanatory principles, or, if needed, re-explains.	Reflects back, gives hints or further leading questions, provides another instance or simple problem. Avoids direct explanation. Suggests recall of earlier problem and derived rule.
Role of teacher questions	To assist recall of inclusive concepts, principles. To aid discriminability and relatability. To evaluate pupil understanding of concepts and principles. To assist recall of relevant prior learning.	To guide thinking. To provide direction and hints to assist children in deriving and formulating rules etc. To focus attention on criterial attributes, instances. To assist recall of relevant prior learning.
General Set.	To know and understand.	To find out, to understand.

Additional comments on treatment criteria.

As far as possible, a warm, supportive approach should be evident in each treatment, with reinforcement employed in giving children a feeling of success (in problem solving or gaining meaning), providing knowledge of progress, feedback or correctness of response, approval of effort, and so on. The apparatus and equipment used to demonstrate and for practice in treatment R.E. will be the same type as pupils will use in treatment R.D. in the derivation of rules. Complete

control of all treatment variables is impossible, but it is felt that the treatments are sufficiently distinctive to be acknowledged as independent.

Operationally, the independence of the treatments will be defined by coded scores resulting from analysis of audio-taped samples of each lesson. The form of analysis is the Interaction Analysis System described by Amidon and Flanders (1967). The treatments are anticipated to differ in directness-indirectness, proportion of time in teacher lecturing, in pupil initiated talk, in teacher questioning, and, consequently, in pupil responding.

HYPOTHESES.

The surveys of the literature and previous investigations (Chapters II-IV) contributed to the formulation of the problem, leading in turn to the generation of the hypotheses below, which are stated in the research form. The null form is regarded as a procedure, a convention or formula for testing probabilistic inference rather than a means of indicating the practical or educational significance of any findings. (See discussion by Sax, 1968, pp.107, 418, 419). The latter kind of significance can be evaluated on the basis of theory-based argument, supported by such evidence as can reasonably be derived from the data, within the limits of experimental restrictions, the internal coherence of the case, and its external consistency with other findings and

explanations. The hypotheses are given in three groups: those referring to cognitive style, to treatments and to interactions.

- 1.1 The girls in the population studied will make significantly more relational responses on the cognitive style measure than will the boys in the population.
- 1.2 The boys in the population studied will make significantly more descriptive responses on the cognitive style measure than will the girls in the population.
- 2.1 Pupils receiving Treatment R.E. will demonstrate higher levels of initial learning than those receiving Treatment R.D.
- 2.2. Pupils receiving Treatment R.D. will gain higher scores on measures of application and transfer than will those receiving Treatment R.E.
- 2.3 There will be no significant differences in scores on the measure of retention between pupils receiving Treatment R.E. and those receiving Treatment R.D.
- 2.4 There will be no significant differences between teachers as measured by pupils' scores on post-tests of initial learning, retention and application-transfer.
- 2.5 Treatment R.E. will demonstrate more Teacher-lecturing, and less Teacher-questioning and Pupil-talk than Treatment R.D., as indicated by Categories 4, 5, 8 and 9 of an Amidon-Flanders Interaction-Analysis.

- 3.1 Pupils whose preferred cognitive style is descriptive will demonstrate more effective learning as measured by performance on post-tests following Treatment R.D. than will pupils receiving Treatment R.E.
- 3.2 Pupils whose preferred cognitive style is relational will demonstrate more effective learning as measured by performance on post-tests following Treatment R.E. than will pupils receiving Treatment R.D.

Specifically, these hypotheses predict significant F-ratios at the .05 level, in a multiple analysis of variance, for Main Effects for Factors B (Treatments), C (Cognitive Style) and D (Sex); for First Order Interactions B x C (Treatments x Cognitive Style), B x E (Treatments x Objectives), B x D (Treatments x Sex), and C x D (Style x Sex); and for Second Order Interactions B x C x D, B x C x E and B x C x F (Treatments x Style x Occasions). While it is recognised that the .05 level of significance does, as an accepted probability level increase the risk of Type I errors, this level is deemed more appropriate for a multiple interaction study in a naturalistic setting, where within-group variance can be expected to be quite large, than would the more stringent .01 level. For an exploratory study, locating trends seems as important as the possibility of reaching conclusions.

CHAPTER VI

EXPERIMENTAL PROCEDURESSUBJECTS AND THEIR ASSIGNMENT TO CONDITIONS.Assignment to groups.

The subjects for the experimental and control groups were drawn from among the Form I children at a large Normal Intermediate School in Hamilton city, only those few children classified as 'special class' pupils being excluded. A modified form of streaming is used in the school, there being one top-stream class (made up of pupils selected from each year intake on the basis of general ability as indicated by scores on the Otis Intermediate, attainment, and recommendations from the contributing schools), and nine parallel classes each with similar distributions of pupils as regards age, sex, general ability and attainment. The means of the ten Form I classes for age and I.Q. (Otis Intermediate, 1936 norms) are 146.24 months and 109.82 respectively. The average age of the one hundred and fifty children in the experiment was 143.62 months; the average I.Q. was 112.55 (not significantly different from the more general mean). An examination of school records of parents' occupations revealed no reason for the school to be considered as atypical as far as socio-economic status is concerned.

The choice of Form I children, and of the particular school, was influenced by several factors. First, the children

were approaching the Junior High School level proposed by Ausubel as an appropriate stage for the use of expository methods. Although not all Form I children in this school could be expected to be cognitively mature in the sense of manifesting formal operational thinking, it appeared reasonable, following Ausubel's argument, to expect that, with the use of supporting concrete-empirical props, they would be capable of assimilating potentially meaningful concepts and principles. The developmental stage theories, also, provide no reasons to suspect that discovery-type approaches would be any less suitable for these children. Secondly, the children were, in the main, accustomed to observers and 'special lessons', so that neither extreme of docility through being overawed, nor excitability due to novelty, was likely, particularly when student teachers and Teachers College staff were the observers and teachers. Thirdly, the school was currently implementing the new science curriculum for Forms I to IV, and, as a result, the anticipated co-operation and support of the school staff was willingly given. Furthermore, the science content of the experimental lessons was such that both principles-explained-and-demonstrated, and principles-derived-through-guidance approaches were deemed equally appropriate methods to teach the content. Again, the introduction of a new curriculum taught by a team of teachers, suggested some similarity of experiences in science would be shared by all the Form I pupils.

The children were assigned to the experimental and

control conditions on a stratified random basis according to the procedures outlined below. The cognitive style measures, Triad One and Triad Two, were administered by the writer in the ten classrooms over a two-day period, with no testing being undertaken before 9.30 a.m. or after 2.30 p.m. on either day. As the writer had previously visited all the classrooms on other occasions, he was known to the children, and no rapport difficulties were experienced in administering the measures in a relatively 'non-test' atmosphere (cf. Wallach and Kogan, 1965, pp.20-24). The answer forms were scored in accord with the criteria (pp.74-75), and all scores, for boys and girls separately, were converted into Stanine units, in order to render the different distributions for boys and girls, and in each of the modes, equivalent. The stanine scale, a normalised nine-point standard scale in half-sigma units, with equivalent standard score limits of -2.75 to $+2.75$, percentile rank limits 0.3 to 99.7, and having a mean of 5.0 with the two limiting classes being 1 and 9, has been used extensively by the E.T.S. (see Durost in Gronlund, 1968).

All pupils had been allotted a code number in order to facilitate assignment to groups by use of a table of random numbers (Cochrane and Cox, 1957). The pupils were ranked, high to low, in Stanine units derived from Triad One scores, in six categories - the three cognitive styles and two sexes - identification now being by code number. Where pupils shared the same stanine unit, two additional data were used in determining rank order: (a) the difference between the dominant mode and less dominant modes in stanine units, and (b) scores

on Triad Two as an indication of consistency of preference (i.e. generalisability). The experimental design required four experimental groups and one control group, each group containing thirty pupils (fifteen boys and fifteen girls), with five of each sex representing each of the cognitive style dimensions, so that the basic cells contained five children. A cut-off point on the ranked stanine scores, for each cognitive mode and for both boys and girls, was set at 7, a lower limit equivalent to the standard score of +0.75. Thus, stanine units 9, 8 and 7 in each of the six categories, defined the population group, from which the subjects were assigned at random using a table of random numbers (Cochrane and Cox, 1957) and pupils' code numbers. The rationale for the particular procedure is provided in Kerlinger's (1964) discussion of the maxmincon principle; in this case maximising the between-cells variance (style), and at the same time minimising the within-cells variance (style). Similar methods have been employed in investigations where the individual characteristics studied have had a relatively short psychometric and research history (e.g. divergent thinking study of Gallagher and Jenne, 1967).

Although the design requirements were for one hundred and fifty children, the numbers in each of the ranked, sex-differentiated, cognitive style categories permitted the assignment of one hundred and eighty children to six groups, with a reserve pool of children who had demonstrated similar categorising preferences to those assigned. It was, therefore, decided to assign six groups, maintaining one group (the sixth)

as a reserve so that, should any pupil be absent for the experimental period, it would be possible to allocate a replacement, at random, from an equivalent cell. The minimal requirements, for five groups only, are shown in Table I.

TABLE I: STANINE LEVELS FOR ASSIGNMENT OF SUBJECTS

Cognitive Style	Sex of Subjects	Raw Score Equivalents of Stanine Scale			Maximum Number Available	Minimum Number Required
		9	8	7		
Descriptive	Boys	14+	13,12,	11,10, 9,	39	25
	Girls	11+	10, 9,	8, 7,	34	25
Categorical	Boys	17+	16,15,14,	13,12,	36	25
	Girls	16+	15,14,	13,12,11,	32	25
Relational	Boys	21+	20,19,18,	17,16,15,	37	25
	Girls	22+	21,20,	19,18,	32	25

The means for cells for the five groups and six categories were very similar, and revealed no differences which could not be accounted for by chance variations alone. (The largest $t = 1.20$, 8 d.f., N.S.). They are shown below in Table II.

TABLE II: CELL MEANS FOR THE FIVE GROUPS

Cognitive Style	Sex of Subjects	Group 1	Group 2	Group 3	Group 4	Control Group
Descriptive	Boys	8.0	8.0	8.0	8.0	8.0
	Girls	7.8	7.6	7.8	7.8	7.8
Categorical	Boys	8.0	7.8	8.0	7.6	8.0
	Girls	7.4	7.6	7.8	7.6	7.6
Relational	Boys	7.8	8.0	7.8	8.0	8.2
	Girls	8.0	8.0	7.8	8.0	8.2

Some restriction on the generalisability of any findings from this investigation may follow from the fact that this sample is from a school in which the average Otis score is above the standard mean of 100. However, Elley's (1969, pp.140-155) re-standardisation of the Otis Test provides data which would suggest that the mean I.Q. in the school approximates the national mean for this age group. A more serious restriction is that the subjects assigned to the experimental and control groups were high scorers on the cognitive style instruments. At the same time, the children manifested no exceptional characteristics which would make them markedly different from other children of their age.

Assignment of Groups to Conditions.

With the aid of a playing die, the groups were assigned to control and experimental conditions. Four groups were required for the experimental conditions, and one for control. Although the sixth group was not required, it too was given the control lessons, but the results of this group were not processed. The function of the sixth group was to permit the necessary crossing of rooms, time of day for teaching, and general balancing of procedures for the control group. The four experimental groups were then assigned randomly by coin tossing, to the experimental treatments - two to R.E. and two to R.D. The teachers, one male and one female, were similarly assigned, so that each teacher was allotted two groups, one to be taught by Treatment R.E., the other to be taught by Treatment R.D. Four classrooms only were available for the

two afternoons on which the experiment was to be conducted - the science room, the art room, and two large classrooms. The experimental groups drew the science and art rooms, the control group the classrooms. It was decided to equate physical conditions in the experiment by conducting the lessons so that the experimental groups had one of their two lessons (either first or second) in each of the rooms. Thus, rooms and time of day were also assigned by random, using a coin. The general balancing pattern is shown in the table below.

TABLE III: BALANCING PATTERN FOR THE EXPERIMENTAL LESSONS

Day	Group	Teacher One			Teacher Two			Teacher Three		
		Treat-ment	Room	Time	Treat-ment	Room	Time	Treat-ment	Room	Time
1	1	R.E.	X	1						
	2	R.D.	X	2						
	3				R.E.	Y	2			
	4				R.D.	Y	1			
	5							C	P	1
2	1	R.E.	Y	2						
	2	R.D.	Y	1						
	3				R.E.	X	1			
	4				R.D.	X	2			
	5							C	Q	2

Rooms were crossed to minimise differences in physical environment; time of day was crossed to equate fatigue; and

for the teachers, the order of teaching the lessons was crossed to balance unsuspected influences of strain, preference or initial nervousness at the beginning of the lessons. Rather than use pupils as their own controls and confounding the treatments, the two teachers were trained to teach each strategy. Worthen (1968) adopted this practice in his study.

The final randomising procedure was related to the pupils' classroom positions. The experimental rooms, of similar size, lighting and ventilation, were equipped with similar furniture, arranged to permit groups of six to eight children to sit together. The effects of established pupil-pupil relationships were unknown, and could have varying effects. Therefore cards were prepared (5 inches x 3 inches, folded to stand on the desks) on which were printed in bold, black letters one and a half inches tall, the Christian name and the initial of the surname of each child in the study. Under the fold, in small pencilled lettering, were coded symbols for the cell (e.g. R/B1 = relational boy, group one), and for the pupil (e.g. 25 R.5 = twenty-fifth boy on the Form I school roll, home room 5). The cards were randomly distributed to seating positions before the children entered the rooms, with two limits. Girls and boys were placed in separate groups (a practice in the school), and, as far as possible, pupils from the same classroom were placed in different groups^x. The control group was also treated in the manner just described.

In sum, then, randomisation procedures were adopted in assigning pupils to groups, groups to conditions, teachers

* In fact, the randomisation made the second limit unnecessary.

to groups and conditions, groups to classrooms and periods of the day. The initial order of teaching the lessons was also randomised. For lessons following the initial one, those factors which could be interchanged (order, room, periods) were interchanged. However, seating positions were held constant for the lessons, teachers taught the same groups throughout, and the same arrangement of furniture in each room was maintained.

THE GENERAL FEATURES OF THE STUDY.

Cronbach (1966) has emphasised the need for more complex experiments, and Siegel (1967), in similar vein, has presented a cogent case for analysis of variance studies which investigate more variables than are included in those typically reported. This investigation is focussed on interactions between variables which are believed to be important in teaching situations. The major independent variables were three cognitive styles (descriptive, inferential-categorical and relational); two teaching strategies (rule-explained and rule-derived); two teachers; and the sex of pupils. The effects of two further independent variables (objectives and occasions) were examined by using four different science-content post-tests. The test items sampled "objectives" at two levels (knowledge-understanding, and application-transfer), on two occasions (the day after the final lesson, and fourteen days later). All tests were given at the same time, and under the same conditions for all subjects.

The basic experiment, then, resembled a $2 \times 3 \times 2$ fac-

torial design, with method, cognitive style and sex being the three factors. For control purposes, an additional factor, "teachers", was included in the design, while the factors "objectives" and "occasions" were introduced to all multiple dependent measures to be used in the experiment.

Scores from "different pupils" were obtained for each level of teachers, method, cognitive style and sex. Scores from "same pupils" were obtained for each level of objectives and occasions. Following Lewis (1968) and Winer (1962), "pupils" were then regarded as a supplementary factor, one which was nested within teachers, method, cognitive style and sex, and yet at the same time was crossed by objectives and occasions. Such a design can be classified as a seven-way experiment, containing a double crossing of a nested factor. The nested factor "pupils" was regarded as a random effect, while teachers, method, cognitive style, objectives and occasions were taken to be fixed effects. (See Appendix for further details of the design, for choice of error term, and for general models).

The control group was considered apart from the basic analysis of variance design. Because the control group shared some of the post-measures, but not all, it was impracticable to enter control group results into an already complex design. However, the functions of the control group, amplified earlier, indicate why it was not intended to enter this group into the analysis of variance. Further explication of data processing and statistical features appears in later sections of this chapter.

INDEPENDENT VARIABLES.

Cognitive Style and Sex.

The cognitive style variable has been considered in some detail in Chapters II and V, and the manipulation by stratified random sampling described. The fundamental cell entry was five for each conceptual mode and each sex, giving ten entries across cells for treatment by style contrasts. Class group size is thirty, treatment groups size is sixty. The size of the minimal cognitive style cell could not be increased further without making class group size too large for observation required to test the independence of the treatments. Unfortunately, the possibility that the restriction on cell numbers could act to diminish the statistical significance of any findings had to be accepted as a concomitant.

Treatment and Teachers.

The treatments have been defined, and their criteria delineated in earlier parts of this report (pp.76-79). It is now necessary to report the means by which the teachers were trained to apply each treatment, and to describe the techniques employed to evaluate the independence of each treatment.

Ten second-year student teachers, whose academic study in the Teachers College was science, were selected by members of the science department. The criteria for selection were that they showed superior ability in their science studies, and had demonstrated above average competence in practice teaching. Student teachers were selected for the experimental

teaching because of the anticipated sensitivity to close observation, method reversion tendencies and differential experience of primary school teachers already established in the profession. Of the ten students, four (one man and three women), who showed themselves to be confident, of similar intellectual ability (based on AI/AQ scores), of similar age, and similar teaching skill (based on College supervisors' ratings), were selected to be the teachers, while the remaining six were to be the observers.

The nomination of which students were to fulfil which of the roles, teacher or observer, was made after the group had worked with the writer over a period of four weeks. The total training programme extended over ten weeks, the meetings and discussions, practice teaching and observations, averaging more than five hours a week, giving a total in excess of fifty hours. The training programme, supplementary details of which are in the Appendix, included lectures, discussion and reading related to:

- (a) a study of learning theory and the relationship of this to emerging theories of instruction,
- (b) an examination of task analysis and statements of objectives following Gagné's models,
- (c) a detailed study of the conceptual basis of the science content - stability, its concepts and attributes, and major principles,
- (d) a study of cognitive styles, particularly the Kagan construct, and some of the instruments used to define them, and

- (e) the development of a teaching paradigm to act as a synthesis of the foregoing.

The paradigm thus developed was then applied to the experimental tasks in formulating:

- (a) objectives for Form I children displaying characteristics believed typical of the experimental population in regard to prior learning, attitudes to science, academic achievement, and range of abilities;
- (b) task analysis, sequences and appropriate forms (enactive, iconic, symbolic - see Waterhouse, 1968) of presentation of conceptual learning on the topic of stability;
- (c) teacher functions in the two treatments, R.E. and R.D.; and
- (d) generally detailed plans of the lessons.

The treatment criteria already outlined in Chapter V were built up and amplified, with the student group, so that both teachers and observers were thoroughly familiar with them. Apparatus and equipment was made, collected, and duplicated, so that there were sufficient identical materials available for each treatment. Practice teaching was arranged with two random samples of Form I Intermediate school children from a non-experimental school, and with four groups of senior primary pupils at a country model school. In this way, provision was made for each teacher to teach each treatment, for the observers to practise their skills, and, incidentally, for some of the tests and test items to be tried out. Each practice teaching session was followed by an evaluation and discussion period, aiming at refining the treatments and

increasing their independence. By the time the experiment was to be conducted, all students were completely au fait with the requirements.

One limitation was imposed. No student teacher or observer was informed about, or had access to information about the cognitive style or intellectual ability of any individual pupil. They knew only that the experimental and control groups had a certain general composition, and that they could anticipate certain prior learning and positive attitudes towards the lessons, i.e. information relevant to good instruction.

To ensure that the treatments were those intended, two observers remained in each classroom for the lessons, one recording teacher statements, the other pupil responses. Combined, these gave a full transcript to check against portable Phillips cassette tape-recordings made of each lesson. The lessons were equated in time - each lasting fifty minutes - with a total time for each group of one hundred minutes. The task analysis and prior planning ensured that each treatment dealt with the same content, although in different sequences and strategies. Each of the experimental teachers taught two lessons on each of the two afternoons during which the experiment was run. To allow for re-orientation to a change of instructional method, to give a period of relative relaxation, to permit apparatus, equipment and materials to be checked and moved to the second classroom, and to allow time for pupil name cards to be collected and issued in their random order, a twenty-minute break was made between the two lessons on any day. The first lesson each afternoon began at 1 p.m., the

second at 2.10 p.m.

Because the groups were random samples, some children came from each Form I classroom for each of the lessons. To avoid confusion over lesson times and rooms, each member of the school staff having pupils taking part in the lessons was given a typed list, showing pupils' names, room numbers and times for the lessons. To be doubly sure, each pupil engaged in the experiment was also given an individual slip, on each day, telling where to go and when to attend. This method was valuable in helping to identify any absentees. On the morning of the first day's experimental teaching, it was found that six children were absent. Replacements were assigned at random from the appropriate cells in group six, the vacancies in this group being filled from the reserve pool. By good fortune, the only absentees on the second day were from group six. Thus, for the duration of the experiment, the experimental groups and the control group remained intact. The writer visited each of the rooms, which were in close proximity, near the beginning and end of each lesson, staying the same time in each room and alternating his order of visiting. No difficulties were apparent, the lessons appearing to be as planned and rehearsed.

Objectives and Occasions.

Objectives occupy a somewhat paradoxical position in methods research, a paradox evinced by frequent synonymous use of the terms goals and outcomes, objectives and products.

None of the teaching models published in various journals and texts (e.g. de Cecco, 1968; Verduin, 1967) place objectives in other than an antecedent position in directing teaching strategies. The primacy of objectives is not only in directing the manoeuvres and tactics employed within any teaching strategy, but is also in guiding the selection of appropriate methods. Implicit in the two instructional approaches employed in this study, as the review of the literature suggests, are different objectives. Furthermore, in the evaluation process, the criteria by which judgements about pupil performance are made, are two groupings of objectives, both specific behavioural objectives and broader curriculum objectives. From performance, learning and achievement of objectives are subsequently inferred (Gagné, 1965). The tendency in much of the reported research is to regard pupil performance, and thus "the objectives" as dependent. In one sense it might be said that objectives are both independent and dependent. However, it seems legitimate when conducting methods research studies, to regard the treatment and the objectives directing the treatment as independent, and to consider performance as dependent.

There must be, it seems axiomatic, an interaction between objectives and strategies, between pupils and objectives. Education is a process in a time dimension, permitting a most intricate and manifold matrix of interactions. Carroll (in Krumboltz, 1965) defines three of the variables in his model for school learning in terms of time (i.e. aptitude, perserverance and opportunity to learn). The objectives in

this study are measured in performances after a relatively short time span. However, they are measured on two occasions in the belief that school learning is "over the long haul". Thus, objectives and occasions are built into the statistical design. The time dimension, of course, brings with it the measurement problems related to changes in scores, reliability practice and regression effects. Cronbach (1966) has claimed a need for longer term studies in classrooms, but, while he is no doubt justified in making the claim, such studies would require an entirely different research strategy from this one.

The broad objectives of the experimental unit are set at two levels, as indicated by the selected exemplars below:

Example 1 - knowledge and understanding.

The children will demonstrate their understanding of the concept and principles of stability by the following performances:

- Identify objects which are stable.
- Recognise and state the attributes of such objects.
- Discriminate between examples and non-examples.
- Suggest where the centre of gravity is located in various objects.
- Demonstrate the effect of tilting an object until its centre of gravity falls outside the base....

These behaviours will be demonstrated orally, through manipulation, and, especially, through the selection and construction of correct and appropriate answers in paper and pencil tests.

Example 2 - application and transfer.

The children will manifest, through performances such as those listed, skill in applying and transferring their knowledge:

- Decide why newly presented or portrayed objects balance in stable positions.
- Select the most stable of several human balancing positions.
- Solve problems related to balance and stability.
- Suggest ways to make a hanging object more stable....

Further examples of objectives are given in the appendix, where they are used to exemplify the content validity of one of the tests.

FURTHER EXPERIMENTAL CONTROLS.

Some degree of control was exerted over the prior learning of the subjects. The school's science teaching team had agreed to teach, for the fortnight immediately preceding the experimental unit, Unit 4 of the Forms I to IV Science Curriculum, Force and Energy. The Departmental Science Guides were followed, test items were prepared and discussed with the teachers, who were thus fully aware of the evaluative criteria to be employed at the end of the gravity unit. The five embedded items were not, however, shown to the school's science teachers, lest they inadvertently taught the specific content of these items.

It was assumed that the randomisation procedures adopted in assigning subjects to groups would have controlled (McGuigan, 1960, pp.107-120) for a number of extraneous variables, including general ability, reading ability, level of motivation, socio-economic status, and general attainment. However, as a further check, measures of academic motivation were constructed and administered. Their scores, together with Otis Intermediate I.Q. scores resulting from tests administered by the first assistant of the school some nine months earlier, were recorded, to be used in statistical controls should tests show they were necessary. It was also assumed that Hawthorne effects would act equally on each of the experimental groups, and thus be balanced out by the experimental design itself. Content, time, physical conditions were equated and held constant. The teachers for all groups were matched, but for the experimental conditions one

was male, the other female. The techniques employed were those classified by McGuigan (ibid., 1960) as balancing, maintaining constancy of conditions, randomising and equating. Elimination and counter-balancing were not employed, because of practical difficulties and possible confounding of the treatments.

Obviously it is impossible to control all conditions in near naturalistic settings, and the assumption that subjects will regard the situation as not too dissimilar from normal conditions underlies the whole experimental enterprise in education. Observations of the children revealed general signs of high interest, co-operation and involvement in all conditions. Novelty factors may have contributed to this state of affairs.

DEPENDENT VARIABLES.

The dependent variables are pupil performance on each of the four post-tests, described in the chapter on experimental instruments. The items in the tests are conceptually rather than factually oriented, in keeping with the whole research perspective. Five items embedded in a test preceding the experiment were designed to establish entry level, and later, with answer alternatives re-randomised, were included in the test of initial learning and the retention test. In order to maintain comparability between the measures, all test scores were converted into C-scale units.

"The C-scale has many of the advantages of the T-scale. It refers obtained scores to a common scale that is related to the normal distribution. If the population distribution on a measured trait is normal, then the distribution of C-scores properly represents that population, and the units of measurement may be regarded as equal. It lacks the refinement of a small unit such as that provided by the T-scale. On the other hand, it probably more nearly represents the accuracy of discrimination actually made by means of tests" (Guilford, 1950, p.305).

As the range of scores on all four post-tests approximated fifteen raw score units, and the distributions were not, by inspection, different from a normal pattern, the eleven-point scale represented little loss in precision.

It was also intended to use attitude scale data from pre- and post-tests to evaluate motivational changes under each treatment, but with only one hundred minutes of teaching directed towards the achievement of cognitive objectives, it was realised that little change could be anticipated. The stability coefficients (0.789 and 0.910) between the two administrations (fortnight interval, $N = 120$) provide justification for the position taken. The scores from these measures, corrected for extreme response set, were also converted into C-scale units in preparation for regression analysis and multiple correlations, in order to determine whether or not attitudes to academic achievement and science contributed sufficient variance for them to be employed as covariates in an analysis of covariance.

DATA PROCESSING.

All scores on the dependent measures were converted into C-scale units. This permitted assumptions about

normality to be met, and increased the likelihood that the assumptions of homogeneity would be met before analysis of variance computations were begun. A number of studies (see McGuigan,¹⁹⁶⁰ pp.285-286) have shown that analysis of variance is a particularly robust statistical procedure, standing up to all but gross violations of the homogeneity assumption. It will be recalled, also, that the cognitive style scores for each sex had been converted into stanine units, thus normalising the distributions and permitting comparisons. The high scorers in each mode (stanine 7 plus) were the population from whom the sample was drawn. Thus, for all the measures referred to, frequency distributions were drawn up, means and standard deviations were calculated, and normalisation by transformation into stanine and C-scales completed. Details of the instruments, construction, item analysis information, reliability and validity data are reported elsewhere. All marking was done by the writer, two checks on the consistency of marking being made with other raters. One pupil from the experimental groups was absent for the post-tests, and as she had scored on the mean (stanine 5, C-scale 5) on pre-test measures, the mean of post-test scores was entered as the best prediction of her scores. (Travers, 1969, pp.357-358).

The analysis of variance model is based on assumptions about the additivity of contributory variance, and the assumptions of the additive nature of variations are, in turn, dependent on other assumptions. These include the mutual independence of observations within sets (random sampling) and experimental homogeneity and normality of distributions within

sets (Guilford, 1950, pp.257-258). The assumptions concerning mutual independence of observations within sets, and the normality of distributions within sets, were judged to have been satisfied by the experimental procedures used. Kerlinger (1966), in evaluating the factorial designs with three or more variables, points out the disadvantages in complexity, and the consequent difficulty of interpreting multiple interactions. This study, however, is concerned with interaction rather than with linearity.

"In the opinion of some educational thinkers, the study of interactions in educational research is becoming increasingly important and should become a central preoccupation of educational research workers". (Kerlinger, 1966, p.228).

Analysis of variance also permits the control of some variables (e.g. teachers) to be built into the statistical design, hence allowing their influence to be gauged and compared. It seemed to the writer that analysis of variance, then, was consonant with the research enterprise, and that there was no logical incompatibility between the statistical processing and the experimental strategies. The particular form of statistical design used (i.e. with crossed and balanced factors, and nested factors) is justified and exemplified by Lewis (1968).

However, there are certain consequences which follow from the use of the statistical model and the data processing procedures described. It has been argued that the "logic" of the teaching paradigms gives priority to objectives, and so places them among the primary independent variables. On the other hand, the action of separately normalising each set of

scores derived from the dependent measures, removed the chance of gaining significant main effects for objectives, while still permitting the possible interactions to be shown. The core of the investigation was the interaction between different levels of objectives and of treatments and cognitive styles, in producing different achievements in science. At the same time, the procedures used reduced the chances of this occurring through spurious effects likely to be manifested by combining the scores from the four science tests, with their differing levels of difficulty. Another way of conceptualising the design would have been to regard the four science tests as sub-tests of the one measure of science attainment, establishing a common scale, and using this as the dependent variable. In this way, main effects for objectives would have been illuminated, and no doubt the reliability of the overall measure, being four times longer than any of the sub-tests, greatly increased in comparison with the science content tests.

Because general ability, defined by Otis scores, contributed marked variance to the post-tests (r 's 0.433 to 0.540), and had near-zero correlations with the cognitive style scores (-0.041 to +0.148 for girls, and -0.135 to +0.095 for boys, $N = 60$ in each case), it was considered appropriate and reasonable to apply analysis of covariance statistics to the post-tests, using I.Q. as the covariate. Multiple correlations with the dependent measures, previous science attainment, the two attitude scales, and I.Q. as the variables, were computed on an I.B.M. 1130 to give a check

on whether or not a multiple covariate was required (Lindquist, 1953, p.336). However, with standard errors of the R 's ranging from 0.06 to 0.07, and the multiple correlations between the two extremes 0.399 and 0.623, it appeared that little, if anything, would be gained in precision by using additional covariates. The assumption of independence of the covariate from the major variables, in particular cognitive style, was considered met on the basis of the data above. The remaining assumption, homogeneity of regression, was tested (Edwards, 1969, p.338) and sustained by the non-significant F ratios. (F 's, 0.597 - 1.251, d.f. 23/72).

The analysis of covariance used takes each set of scores on the four post-tests separately, and examines the four major independent variables of teachers (an experimental check), cognitive style, sex and treatments, with increased precision - the variance due to I.Q. being removed. In this procedure, the nested-factor error term explicated by Lewis (1968) is not needed.

While the statistics used are considered rigorous and robust, there are educationally-difficult problems involved. Analysis of variance methods assume within-group variance as the error measure, but systematic variance resulting from particular characteristics not specified in the design is then also included in the non-systematic error measure. Furthermore, the assumption of homogeneity of variance on a given independent variable, while fitting in with entry behaviour ideas, does not coincide with the possibility that effective teaching may increase within-group variability on given

achievements.

The biological research model with its associated statistical procedures may be restricting in that increased precision tends to be linked with decreased generality in applying findings. (Sax, 1968). Thus, the very restrictions and controls in the study may impose further limitations on the investigation here reported. The two major analyses were calculated, following the preparation of tables (Moroney, 1956; Lewis, 1968) on Facit and Monromatic electric desk calculators. Analyses were also processed by I.B.M. electronic computers, the data being recoded suitably. The analysis of variance run was based on cell means rather than individual cases. The analyses of covariance were four one-way computations. These provided a cross check with the manual calculations.

The control check on treatments and teachers was undertaken through the coded scores derived from Amidon-Flanders interaction analysis. Two observers had recorded separately the teacher statements and pupil responses, these observations being combined later into a transcript of the lessons, and used to assist the group of students listening to the tape-recordings of the lessons. Three five-minute samplings from each of the eight lessons taped were used as the samples. The tape for one lesson was inaudible, and for another was of shorter duration than the remainder. The samples, because the counters on the recorders did not calibrate exactly, were close approximations to five minutes each, with about one hundred three-second codings per sample being

made. Because the cassette tapes permitted only thirty minutes playing time, the first thirty minutes of the first four lessons, and the final thirty minutes of the second four lessons were recorded, the five-minute samples being alternate five minutes throughout. For the former, the samples began with the first five minutes, for the latter with the second five minutes.

Five student teachers, trained by a Teachers College staff member who had studied the Amidon-Flanders system in the United States, coded the lessons after repeated playings, until there was one hundred percent agreement on the codings. Nine of the ten categories were followed precisely from the manual, the tenth being modified to code practice, problem-solving and manipulation of apparatus in groups. Combined matrices for each teacher, and for each treatment, were drawn up, and total scores in each category converted into percentages. Tests for significance of differences were then carried out. (Garrett, 1958, p.235). Examples of the matrices, showing typical moves in the teaching strategies, as well as totals and percentages, are given in the Appendix.

CHAPTER VII

THE INSTRUMENTS

Apart from the Otis, no appropriate standardised tests were available, and consequently all tests, including the cognitive style measures, had to be constructed for the purposes of this study. Sample items from the Conceptual Styles Test, as available in several publications (Kagan et al., 1963,1964; Gordon, 1966), provided a starting point in the cognitive style area.

In all, ten measures were constructed, two of these being administered as one test. The tests were:

Two Cognitive Style measures: Triad One and Triad Two.

A Pre-test measure: Force, Work and Energy - the five Embedded Items.

An Attitude Scale (Pupil Opinion Survey), which was made up of two measures:

- (a) Academic Achievement Motivation, and
- (b) Interest in Science.

A Test of Initial Learning: Science Test - Part I: A.

An Application-Transfer test: Science Test - Part II: A.

A Retention test: Science Test - Part I: B.

A Delayed Application-Transfer test: Science Test - Part II: B.

A Control Group measure: Simple Machines.

Samples of all tests, instructions and other data, appear in the Appendix.

COGNITIVE STYLE MEASURES.Triad One.

Forty-two items (triadic groups of stimulus pictures), of which eleven were published samples of the original Kagan test, were sketched, photocopied, cyclostyled, and stapled into booklets. Instructions and answer sheets were also prepared for trial. At an Intermediate School similar to the experimental school, two parallel-stream Form I classes were selected for the trial run. Scoring criteria were refined as a result of the experience in marking the reasons given by seventy-six children (Otis I.Q. range 85-136, age range 126-157 months, thirty-nine girls, thirty-seven boys) in the preliminary testings. Two students, who were also engaged in individual studies of cognitive style, check-marked thirty-four scripts, and with discrepancies on only six out of the one thousand four hundred and twenty-eight items, the marking was considered quite adequate.

All items were then analysed three times, once for each of the three cognitive style categories, using upper and lower 27% groupings and Flanagan's tables (Garrett, 1958, pp.365-368). Consequently, four discrimination indices were estimated for each item, three for the three cognitive styles, and an overall average discrimination index. It was decided to use a cut-off level of .35 for average item discrimination, and .25 for any particular mode. Of the forty-two items, twenty-five were retained, ten of the original Kagan items, and fourteen newly constructed ones. Two of the items

retained did not meet all criteria. One, a Kagan item, had an overall discrimination of .37, a discrimination for descriptive mode of .63, for relational mode of .78, but failed to attract responses in the categorical mode. The other had an overall discrimination of .44, descriptive .37, relational .77, and categorical .23. The average discrimination of the twenty-five items making up the final test was .56, range .23 to .88.

While discrimination indices are indications of item validity rather than test validity, it was hoped that the newly constructed items would display discrimination indices of a similar pattern to those of the established Kagan items. In fact, the twenty-five items selected for the test fell within the range of average discriminations .37 to .76, the two limiting indices being those of the Kagan items. As another clue to test validity, it was also hoped that sex-different response patterns, similar to those reported in the Kagan studies, would be evident. This was quite apparent in the scores. For example, means and standard deviations (in raw scores) showed the pattern anticipated, with a significant difference between relational means ($t = 3.755$, d.f. 74, $p < .005$).

TABLE IV: SEX RESPONSE PATTERNS

	Descriptive		Categorical		Relational	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Boys	7.14	6.35	8.45	6.40	7.24	7.41
Girls	4.97	4.50	7.07	5.06	12.76	5.15

Two of the students assisting with the scoring were engaged in a statistics project under the guidance of two senior members of the mathematics department of the Teachers College, and had administered the trial instrument to thirty Standard 4 and thirty-six Standard 3 children respectively. The split-half reliability quotients for the descriptive mode were, in order, .94 and .91 (all calculations having been checked by the students' supervisors). It appeared likely, then, that an adequate instrument could be constructed from the retained items.

The black and white line drawings were deliberately kept as close in style to the Kagan items as the writer could manage, i.e. not too precise or exactly representational. The retained items were coded 1 - 25, and, with a table of random numbers, arranged in final order (two items to a page), and stapled into booklet form, with the instructions making the front sheet. As a result of observation of some errors in pupils' recording of pairings during the trial testing, the items were numbered consecutively. However, the stimulus figures were lettered, not a b c repeated throughout for each item, but in the following fashion: 1 a b c; 2 d e f; 3 a b c; 4 d e f; etc., which is the method adopted in the S.T.E.P. battery. The answer sheet was not changed in general format from that used in the trial testing. The refined test was administered to yet another group of children, randomly selected from the Form I classes at a third Intermediate School, in order to ensure that the instructions, booklet form and answer sheet were adequate. As only thirty children

(fifteen boys and fifteen girls) were in the group, statistical data were not analysed rigorously. It was noted, nevertheless, that the response pattern across styles and sexes approximated that found with the trial test and that reported by Kagan. The pattern, in median raw scores, was: for boys, descriptive 7, categorical 4 and relational 3; and for girls descriptive 4, categorical 5 and relational 5.

Reliability.

The refined form of the twenty-five item test was administered to all children in the ten Form I classes at the experimental school late in the second term, some six weeks before the experimental study was to be conducted. This allowed sufficient time for a test-retest study of the responses of a random sample of Form I children, with a four-week interval, two of the weeks being school vacation. This interval seemed desirable in order to counteract any effects of test novelty, and to inhibit pupils from remembering their original statement of reasons. With the aid of a table of random numbers (Cochran and Cox, 1957), thirty boys and thirty girls were selected across all ten classes, and the test was re-administered. A stability coefficient was considered to be necessary, as the cognitive style construct was a major independent variable in the study, because of Kagan's claim about the stability of categorising preferences, and because of Carroll's arguments about the importance of time as a factor for any theory of teaching.

The coefficients for the three modes and two sexes are

shown below. The N in each case is twenty-nine; two children (one boy and one girl) were absent for the re-test.

TABLE V: STABILITY COEFFICIENTS FOR COGNITIVE
STYLE (TRIAD ONE)

	Descriptive	Categorical	Relational
Boys	.839	.753	.857
Girls	.863	.778	.935

Validity.

The cognitive style measure appeared to have good face validity in that the items constructed were plainly of the same form as those designed and used by Kagan and his associates. A further indicator was furnished by the fact that the items defining the upper and lower limits of the discrimination range were two of the original Kagan items, suggesting that item validity was satisfactory. It has already been implied that one measure of construct validity would be the extent to which the test separated boys and girls, particularly on the descriptive and relational response modes. The medians reported above all demonstrate the pattern theoretically predicted, in which boys at this age make more descriptive responses than do girls, and girls at the same age make more relational responses than do boys.

In order to make a more precise evaluation of the hypothesised sex differences in categorising styles, means and standard deviations (based on the scores of the random sample

from the experimental school) were calculated on an Olivetti Programmer 101. The between-sex means were found to be significantly different for both descriptive and relational modes (t-tests, two-tailed).

TABLE VI: BETWEEN-SEX DIFFERENCES IN COGNITIVE STYLE

	Descriptive		Categorical		Relational	
	Boys	Girls	Boys	Girls	Boys	Girls
Mean	7.483	5.034	8.448	6.966	8.724	12.862
S.D.	4.918	3.718	5.494	4.200	6.584	6.390
N	29	29	29	29	29	29
d.f.	56		56		56	
t	2.412		1.166		2.432	
p	< .02		N.S.		< .02	

One further suggestive trend supporting the argument for construct validity was evident from an examination of response sets on the attitude scales. Differences between sexes and between styles were noted in the number of extreme responses made. Using the formulae from Garrett (1958, pp. 223-225) for small samples, pooled S.D., and S.E. were computed, and t-tests (two-tailed) run to test the null hypothesis that no real mean differences exist between boys and girls on each cognitive style measure, nor within the sexes in each mode.

Table VII summarises the results.

TABLE VII: EXTREME RESPONSE SET AND COGNITIVE STYLES

	Descriptive		Categorical		Relational	
	Boys	Girls	Boys	Girls	Boys	Girls
Mean	8.840	10.920	10.000	9.160	10.160	12.040
S.D.	6.479	4.636	6.589	5.072	7.565	5.256
N	25	25	25	25	25	25
Within-Sex t	D x C 1.199	D x C 2.448	C x R <1.00	C x R 3.770	R x D 1.266	R x D 1.530
p	N.S.	<.02	N.S.	<.01	N.S.	N.S.
Between-Sex t	2.494		< 1.00		1.952	
p	< .02		N.S.		< .10(N.S.)	

In interpreting the results it is well to remember that for the four experimental groups and one control group, the children who provided the new data had been assigned on a stratified random basis. They were high scorers on the styles-measure, and this would have increased differences between cognitive style groups. Thus, while the null hypothesis is not sustained in regard to the between-sex differences on the descriptive mode, nor in regard to the girls' differences in extreme response scores between the descriptive and categorical modes, and between the categorical and relational, the grounds for not supporting it may be contingent on assignment, rather than a product of more general differences.

A little more confidence in the validity of Triad One was gained following observation of children at work during one of the attainment tests. The first seven children to

finish had scored highly on the relational style; the last eight had scored highly on the descriptive style. In the latter case all were boys, of whom two had the highest descriptive scores in the school. The seven children completing the test first were girls. Unfortunately, it was practically impossible to record more precise data across the one hundred and twenty children in the testing situation. Overall, the evidence for validity appears to be strongly suggestive rather than definitive.

Triad Two.

Triad One differs from Triad Two in requiring only one selection for each group of stimulus pictures. They are similar in asking for a pair grouping, for a statement of reasons, in having no time restriction, and in their non-test orientation. Item discrimination data accompanies the test samples in the Appendix. The development and testing of Triad Two paralleled that of Triad One in item construction, trial population and testing, and item discrimination procedures, but there was one point of departure. The instructions for Triad Two, which comprised five triads, ask subjects to make three pair groupings from each set of three stimulus pictures. The same two stimuli may be chosen for different reasons on each occasion, different pairs of stimuli may be chosen for essentially the same reason or reasons, and pairings may differ for each of the three selections.

The reasons for constructing Triad Two in this way were because:

- (a) the power of any item to attract one response category dominantly has not been ascertained;
- (b) it provided, in this form, an indication of generalisability;
- (c) it supplied a compromise between individual object-sorts and a group conceptual styles test, and
- (d) it provided a check to balance against a single measure criterion of cognitive style.

Item discrimination (procedures as for Triad One) for this short test was surprisingly good, with twelve indices being estimated for each item, for the three styles, for the three choices and with three average discriminations across styles for the three choices. The mean discrimination based on the sixty indices was .61. The figures for Item 5, the item with the highest average discrimination index, are given below in tabular form.

TABLE VIII: DISCRIMINATION INDICES (ITEM 5, TRIAD TWO)

	Descriptive	Categorical	Relational	Across Styles
As 1st choice	.66	.71	.81	.73
As 2nd choice	.84	.50	.98	.77
As 3rd choice	.76	.38	.68	.61

Other features which Triad Two shared with Triad One were patterns of sex differentiation (and thus validity), and patterns of test-retest reliability figures. Like the major test, the short-form Triad Two showed lower reliability quotients for the categorical mode than for the other two modes.

TABLE IX: STABILITY COEFFICIENTS (TRIAD TWO)

	Descriptive	Categorical	Relational
Boys	.784	.589	.843
Girls	.808	.614	.812

(N = 29 for each sex, each style. Interval 4 weeks)

The scores from the two tests were correlated (Pearson product-moment), with the following results:

TABLE X: CORRELATIONS BETWEEN TRIAD ONE AND TRIAD TWO

	Descriptive	Categorical	Relational
Boys	.664	.404	.627
Girls	.618	.214	.554

(No correction for attenuation was made)

The moderate coefficients may be interpreted to indicate the change of task from a first choice only (Triad One) to a three-choice situation (Triad Two). If this is a reasonable interpretation, it would be consonant with some of the findings discussed in Chapter II, where it was suggested that stimulus and task factors make demands on a subject which may be incongruent with the subject's preferred mode of categorising. Thus, the low relationship some writers (e.g. Wachtel) report between embedded-figures tests, which demand analysis, and conceptual styles tests, which permit the expression of preference, can be understood in terms of task demand and ability to perform.

Another feature merits comment. Kagan has remarked

that his results are more striking for boys than for girls, but the data found in this study suggest that girls might be more stable in their preferences at this age; certainly the standard deviations for girls are generally less than those for boys, and, apart from the relational mode with Triad Two, the girls' scores show slightly higher stability coefficients over a four-week period than do those of the boys. The generally lower reliability figures for the inferential-categorical style may be accounted for by several possibilities: scoring inconsistencies, verbal facility, the possibility that this mode is a later developmental acquisition than either the descriptive or the relational, or the relative attractiveness of the items which tended to be less discriminating for this mode.

It was necessary to conduct one other check, a check on the reliability of the marking criteria. A staff colleague, whose own research had been in a similar field, was given a random sample of thirty answer sheets especially reproduced so that no marking clues were available. After a brief discussion, during which he was given written criteria for scoring the answer sheets, he marked the thirty papers. On an item-by-item count, a 96.84% agreement between the original scoring and the check marker's scoring was revealed. Correlations between the two markings and total scores, mode by mode, were very high (descriptive $r = 0.994$, categorical $r = 0.987$, relational $r = 0.992$). As the writer had performed all the original scoring, this confirmation of consistency was judged to be very satisfactory.

SCIENCE CONTENT MEASURES.

The general procedures adopted in constructing the science tests for this investigation were as follows:

(i) The objectives of each test were defined as clearly as possible, in order to secure maximum face validity of the items as they were devised. A content analysis of the unit, together with its relationship to the immediately preceding curriculum unit, was made. This was not as rigorous or as detailed as the Gagné (1965) models, but, in keeping with suggestions made by Scandura (1967), specified the major conceptual components of the unit. Bloom's taxonomy was similarly simplified to two main levels:

- (a) knowledge and understanding of attributes, concepts and principles, and
- (b) application and transfer to related and other situations, e.g. physical education.

Both diagrammatic and verbal components were used in constructing the items.

(ii) Because it was desirable not to exceed sixty minutes at any one testing period, and because on one occasion three tests were to be administered, it was decided to restrict the science tests to twenty items each.

Within these limits, the content of the tests was chosen to include as wide a range of relevant material as possible, consonant with the test objectives and known abilities of the 11 - 12+ year age groups.

(iii) Sufficient items were constructed for about one-third

to be discarded, and yet leave enough for adequate instruments.

- (iv) Because the content was situation-specific, and included a large proportion of material not usually covered in the school science syllabus, it was impossible to employ the conventional pretesting of all items. Nor was it possible to use children from the experimental population because of sensitising problems, practice effects and discussion among pupils subsequent to testing. The compromise adopted was to use some of the items as informal evaluation following practice-teaching with children from other schools; to try some of the items with Teachers College science students; and to submit the tests to five Teachers College staff members (two science lecturers, one physical education lecturer and two education lecturers), and to a senior member of the school staff, for examination of their quality in regard to item construction, conceptual accuracy and probable difficulty. This resulted in the elimination of many items and the rewriting of others.
- (v) The final version of each test was designed to begin with two or three easy items which could be answered successfully by most of the children, and to have a gradient of difficulty permitting discrimination over a fifteen-point range. The ranges for the five tests were, in fact, 17, 15, 16, 15 and 16.
- (vi) The item analysis procedures used for all science test

- items were based on upper and lower 27% groups (Connaught and Skurnik, 1969, pp.225-232). This method, with Flanagan's tables, has been reported to be highly effective and very accurate. (Ibid.)
- (vii) All tests were estimated to require a maximum of thirty minutes for all children to complete, and were not speeded. There was no pupil who failed to complete the test.
- (viii) Two of the tests, Science Tests IA and IIA, were administered on the day following the final lesson. The other two, Science Tests IB and IIB, were given fourteen days later.

The Pre-test.

Many authorities (e.g. Winch and Campbell, 1969) have pointed out the problems arising from pre-tests which sensitise pupils to the experimental lesson content, and so introduce a confounding effect to an unknown extent. This is especially true when one of the treatments requires subjects to derive principles with a moderate degree of guidance. Accordingly, it was decided to follow a strategy employed by Wittrock (1963), and to embed five multiple-choice concept items in a test which evaluated a recently-taught science unit on force, work and energy, this unit being taught immediately prior to the experimental unit. It was anticipated that the discrimination indices for these five items would be low, and that each item would be answered correctly by one pupil in four, approximating chance expectations with items having

four choices and dealing with completely new material. These five quite different items with choices re-randomised on each occasion, were to re-appear in the test of initial learning (Science Test - Part I: A), the retention test (Science Test - Part I: B), and in a retention test for the control group. The history of the items over the three occasions is tabulated below, calculations being based on scores from upper and lower 27% groups and Flanagan's tables.

TABLE XI: ITEM ANALYSIS DATA FOR EMBEDDED ITEMS

Items	Difficulty Levels			Discrimination Indices		
	Pretest	Initial Learning	Retention	Pretest	Initial Learning	Retention
1.	.27	.35	.38	.17	.57	.56
2.	.24	.22	.35	.27	.27	.48
3.	.16	.38	.46	.21	.47	.41
4.	.23	.50	.53	.36	.57	.71
5.	.08	.18	.30	.19	.58	.73

Confirmation of near chance scores at entry into the experimental conditions was gained both from the difficulty and discrimination indices, and from analysis of covariance data, which revealed near zero regression coefficients (.019 - .163) with the four post-tests. It therefore seemed legitimate to regard the entry level for the experimental content to be a zero base line. The history of the five items was interpreted as supporting the general procedure adopted in test construction, since items could not be pre-tested and analysed where pupils had not been given the opportunity to learn the material.

The material was specially selected for the study because the science concepts and principles involved were novel to the children, were representative of, but not part of the new science curriculum, and could be taught with equal appropriateness by either of the treatments.

The pre-test was administered by class teachers, two days before the experiment began, all testing being commenced at the same time (11 a.m.) with no time restriction. As expected, all children completed the test within thirty-five minutes. The embedded items and those making up the embedding context had been previously scrutinised by the science teachers of the school, to ensure that the level of difficulty, wording, content relevance, and the form of the items used were suitable. Administration by class teachers, it was hoped, would prevent the pupils from associating the experiment with the test. Each teacher received typewritten instructions on the administration of the test, following a briefing session. As far as could be ascertained, there was little between-class variation in administration. Total scores gained by the pupils conformed closely to those expected by the teachers of the science classes. The two senior science teachers who scrutinised the embedded items, did not teach the tested unit. Those who actually taught the unit had no prior knowledge of the embedded item content, but had received information about the remainder of the test.

The Post-tests.

Four post-tests were constructed. The tests of initial learning and of application-transfer were administered to all subjects in the school assembly hall at 10.45 a.m. on the day after the lessons had been completed. The tests of retention and of delayed application-transfer were also administered in the school hall, at the same time of day, fourteen days later. Administration in the school hall ensured that all subjects shared the same conditions, and had the same examiner to give instructions. Because the seating was arranged so that the children had separate desks, and with four adults to organise the issue and collection of test materials, as well as giving assistance with any reading difficulties which might hinder a pupil's response, supervision was deemed satisfactory. There was no sign of copying.

The test of initial learning (Science Test IA), designed to measure knowledge and understanding of the experimental concepts and principles, was adjusted by re-ordering answer choices, using the E.T.S. table of random ordering, and was then presented to the children a fortnight later as a retention test (Science Test IB). The items in the initial learning test were analysed subsequent to the administration. Of twenty items, three were marginal in discriminating power (.18, .18, .19), but the rest were judged to be satisfactory, the mean discrimination index being .47 (range .18 - .71). Ordinarily the three marginal items would have been eliminated, but it was decided to retain them because of their face validity. Re-randomised, as they were in the retention test, only

one of the same twenty items had a discrimination index below .34, and the average discrimination was .49. It was considered that the items were functioning in the intended fashion, and that item validity was at least satisfactory. Of the three marginal items retained from the test of initial learning, only one (Item 4) failed to give a normalised biserial coefficient of .30 or better; it appeared that the diagrams may have contributed to its low discrimination power. In each test, the first six items had difficulty levels of 65% ranging up to 93%, and three of the final six items were at the 30% level or lower. This confirmation of judged difficulty levels, together with the item analysis data, provided some justification of the general procedures adopted in constructing the dependent measures. Apart from the face validity of the tests, some support for the validity of these two tests, and of the five pre-test items, is derived from the increasing discrimination power of the items over time.

The first test of application-transfer (Science Test IIA), like the initial learning and retention tests, was a twenty-item multiple-choice (four alternatives) test. The content of the items ranged widely; six were related to balance and stability in physical education, two to boating, six to everyday objects such as vases, kites and pictures, and the remainder were simple problems. Again a wide range of difficulty levels was evident, four of the first six items giving difficulty indices of .72 up to .81, and two of the last three items falling at .33 and .37, with a mean difficulty level of .54. Five of the twenty items were marginal in dis-

criminating power, falling only just below the cut-off point of .20; the overall average discrimination index being .43. From this data, it was apparent that the test was adequate, but less satisfactory than the tests of initial learning and retention. The heterogeneity of this application-transfer test is shown by a Guttman homogeneity-coefficient of .437 (Thorndike, in Lindquist, 1951, p.581).

The second application-transfer test, (Science Test IIB), differed from the other science content measures, being supplied-response in item-type. The eight items gave a maximum possible score of twenty, as was the case with the other tests. Stringent marking criteria were established for both the verbal and the pictorial responses. It was hoped that the responses would be non-verbatim, reflecting in the subjects' own words the concepts and principles taught. This, indeed, was what an examination of the responses verified. By setting a criterion of 50% or more of the maximum marks for each item to define a correct response, and by using Flanagan's tables with upper and lower 27% groups, an estimation of difficulty levels and discriminative power was made. The mean discrimination index was .59 (range .25 - .77), and the average difficulty level was 42.6% (range 15% - 65%). One further intention was to test a wide range of applications in a heterogeneous test. A Guttman coefficient of homogeneity of .468 suggested that this had been achieved. A random sample of twenty papers, specially reproduced and marked by a staff colleague, gave a correlation of 0.956 between original and check marking. This would have been higher had not the

check marker omitted to record the score for two items.

The control group test (Simple Machines) was not formally analysed as were the other science content tests. This test was included for two reasons - on practical grounds, to equate testing time across all groups, and to give a test of the objectives implicit in the teaching experienced, rather than risk frustration by asking children to transfer what they had not been asked to learn. The same care was taken with the construction of this test as was taken with other tests.

Reliability.

A domain sampling view of reliability carries the clear implication that optimum sampling over a domain, in order to have a reliable measure, should involve a sufficient and representative sample replicated over time. Such a view would suggest tests with a large and skilfully selected range of observations, which are consistent with observations taken over the same domain on another occasion. However, parallel forms are not easily constructed, and may not always be feasible, for practical reasons, in field experiments. Split-half procedures depend also on assumptions of equality in items; equality, that is, in difficulty levels, in opportunity to answer (time), in pupil readiness (prior learning, mood, attitude and so on) to answer each item, and, implicitly, some similarity of content and intention. "Reliability is a matter of the adequacy of the sampling of items as well as the consistency of behaviour of each individual." (Thorndike,

in Lindquist, 1951, p.577). As far as discrimination is concerned, the items are generally very satisfactory.

However, test length, difficulty levels, test heterogeneity and the heterogeneity of the population sampled, will affect the reliability indices resulting from internal consistency. (Garrett, 1958). Internal consistency measures are appropriate when tests are not speeded. As none of the tests used in this study was a timed test, split-half reliabilities using an odd-even split, were calculated. It was realised that the coefficients so gained might be depressed because of the factors referred to above, and because the experimental teaching itself, with a defined population, might decrease within-group variance. The tests covered broad objectives, and so were not narrowly homogeneous, items were deliberately constructed to have a wide range of difficulty, and the tests were limited samplings (maximum of twenty items). Furthermore, it was not really appropriate to have an extensive test development programme, because of the content chosen.

The table shown on the next page gives a corrected (Spearman-Brown formula) internal consistency indices, Guttman homogeneity estimates, the single stability measure possible, and other test data.

TABLE XII: POST-TEST DATA FOR THE EXPERIMENTAL GROUPS

	Test IA	Test IIA	Test IB	Test IIB
Corrected internal consistency coefficients (split-half)	.684	.470	.637	.492
Guttman homogeneity coefficients	.452	.434	.662	.468
Stability coefficients (fourteen-day interval)	.660	-	-	-
Correlation with Otis I.Q. scores	.433	.526	.540	.494
Standard error (raw score units)	1.433	2.082	1.700	2.063
(C-scale units)	1.175	1.434	1.119	1.446
Reliability estimate of double-length test	.812	.639	.778	.659
Range	4 - 18	4 - 17	6 - 20	3 - 18
Mean	12.975	10.496	12.923	11.381
S.D. (raw scores)	2.55	2.86	2.82	2.91
N	150	120	120	120

These data are reported fully because the moderate reliability coefficients must affect the confidence one places in any results from the experiment. A number of points arise from an examination of the data. Had the tests been twice the length, the internal consistency indices would have been more satisfactory. Kerlinger (1966, p.442), in his discussion of reliability, makes this point strongly. "Summarily, more items increase the probability of accurate measurement." The correlations with Otis I.Q. are of interest in that a proportion of variance is indicated, which, if removed by analysis

of covariance, could influence the precision of the whole experimental study. The study is confronted by two sampling problems, insofar as it samples subjects with defined characteristics as well as a knowledge content domain. The low relationships between Otis I.Q. and the cognitive style scores of the experimental group (reported in Table XIV) suggest that to use I.Q. as the covariate may have effects on performance measures without vitiating the subject sampling. Nevertheless, the reliability figures do point to the need for caution when findings are considered.

THE ATTITUDE SCALES.

Two attitudinal scales, a measure of academic achievement motivation and a measure of science interest, were constructed, and the two scales (for administrative purposes) were combined in an instrument called Pupil Opinion Survey. The achievement motivation scale consisted of twenty statements, modification of questions used by Entwistle (1968) and Russell (1969), to which the subjects responded on a five-point Likert-type scale, ranging from Strongly Agree to Strongly Disagree. Ten items designed to measure interest in science were interspersed among the twenty, order and position of items being randomised during construction. Kerlinger (1966, p.487) suggests that the summated rating scale is most useful among the commonly used scales (equal-appearing interval and cumulative), and yields results which are more or less equivalent to the more laboriously constructed ones. A panel

of judges, four Teachers College staff teaching either science or educational psychology, examined the items, agreeing completely on the appropriateness of the weightings proposed for each. In both scales, half of the most favourable responses were negative in orientation, i.e. "strongly disagree" indicated a favourable attitude towards academic achievement, or expressed a positive interest in science. This was necessary to minimise bias due to dispositional tendencies to agree or to disagree with given statements. The weightings recommended by Edwards (1957, pp.149-171) for a five-point scale are 0, 1, 2, 3 and 4, and these were used with this instrument.

The combined measure was tried out with a random sample of Form I children (N = 58) at an Intermediate School similar to the experimental school. The trial administration was conducted on two occasions (the practice-teaching occasions), twenty-nine children responding on each. Student teacher observers recorded comments on the children regarding their eagerness to answer questions, their persistence on lesson tasks, their apparent interest in the science content. It was not possible to make an accurate observation on some children who appeared to co-operate quietly, but whose behaviour was rather unrevealing of attitudes to science and to achievement. However, those children manifesting high interest in science, or a desire to perform well academically, also tended to rate themselves highly on the instrument. The highest scorer on the self-report achievement motivational scale was a girl who displayed great eagerness to ask and to answer questions, to such an extent that other children began to show signs of restlessness. This informal observation suggested that the

instrument was achieving the purposes intended, and need not depend on its obvious face validity alone. The means of high and low groups on individual statements were compared and retained if the high group mean clearly exceeded that of the low group. That this simple approximation method worked is shown by the later, more formal analysis of items. For example, the twenty items of the academic achievement motivational scale gave a mean t -value term of 4.56, only one item (Item 5, $t = 1.72$) failing to reach the value 1.75 which indicates significant differences between high and low groups in average responses, providing there are 25 or more subjects in each of the groups. (Upper and lower groups 40, total $N = 150$).

The trial testings increased confidence in the instrument. No reason was found to alter the instructions, or change the wording of any item; rapport was easily established, and the children appeared to answer frankly, enjoying the non-test situation. It was found necessary to stress care in answering each item, as a very small number of children marked one answer twice, leaving an adjacent answer space unmarked. One source of concern remained. Some children responded only at the extremes of "Strongly Agree" or "Strongly Disagree", not choosing intermediate responses, while other children chose only the intermediate categories and made no extreme responses. It was decided to correct the scores by reducing each item to a three-point range, but, at the same time, keeping a record of the original score and the number of extreme responses made by each subject. As was done with

other tests, frequency distributions were drawn up, and the scores normalised by conversion into C-scale units.

The final test form was administered by class teachers, in all the Form I rooms, at 11.30 a.m. on the same day, twelve days before the experiment, in order to minimise the possibility of pupils associating the instrument and the experimental treatments. The teachers had all been given typed copies of instructions at a briefing meeting. From comments made by the staff of the school, it appeared that the children answered freely, and enjoyed the slightly novel activity. The test was not timed, but most children finished it in twenty minutes.

While an approximation method of gauging item discrimination had been found useful in the preliminary stages, the full t-test method described by Edwards (*ibid.*, pp.152-155) was applied to the final tests. These were considered very satisfactory. The interest in science scale gave a mean t-value term of 6.09 for the ten items.

The reliability of the attitude scales was calculated from test-retest data, with a fortnight interval, to produce stability coefficients of 0.789 for attitudes to academic achievement, and 0.910 for attitudes to science in schools.

In order to have a further basis for checking the validity of the instrument, a Form I class at an Intermediate School other than the experimental school, was given the achievement motivational scale, and the teacher was asked to rate, on a five-point scale, her assessment of her pupils' desire to perform well academically. The obtained correlation (Pearson product moment, $N = 32$) was .259, a low relationship

which requires some additional comment. Twelve of the seventeen boys indicated greater desire to achieve well in school than the teacher would have predicted. This pattern was noted with seven of the fifteen girls. Agreement between the woman teacher's rating and the girls' self-rating occurred with 47% of the cases, but between teacher and boys in only one case. Correlations between pupil self-opinion and teacher ratings are subject to a great many errors, and even high correlations are likely to be suspect. In this case, there are several probable sources of error; one speculated possibility is teacher-misperception of boys, another that the teacher's judgement is related to past achievement, another that the pupils were trying to please the examiner in selecting certain responses. The last possibility, however, is not supported by a scrutiny of individual responses. The Otis I.Q. means for boys (107.6, range 91 - 119) and for girls (104.3, range 93 - 118) did not provide any reason to believe that attainment would differ greatly. As the mean score on the instrument was the same for both boys and girls, the possibility that the teacher's perception of boys differed from her perception of girls seems not improbable.

The final measures were correlated with I.Q., the post-tests, and with each other, the results being shown in the table on the next page.

TABLE XIII: CORRELATION OF ATTITUDE SCALES
WITH I.Q. AND POST-TESTS

	I.Q.	A.M.	I.S.	TEST IA	TEST IIA	TEST IB	TEST IIB
Academic Achievement Motivation (A.M.)	.348	-	.354	.332	.194	.312	.231
Interest in Science (I.S.)	.146	.354	-	.219	.353	.293	.334

(r's above .254 significant at .01; above .195 significant at .05. (d.f. 117)).

It is interesting to note the slightly higher, though non-significantly different, relationships between the two scales and the various objectives measured by the post-tests.

CHAPTER VIII

RESEARCH FINDINGS

The findings reported in this chapter are organised in sections related to the areas covered by the hypotheses. Where statistical data would require very long tabulations, the full tables are given in Appendix D. The data which are relevant to the specific hypotheses, or which, while not predicted, are found to be statistically significant, or approaching statistical significance, at the .05 level chosen for this study, are reported here.

SEX DIFFERENCES AND COGNITIVE STYLE.

Much of the data pertaining to Hypotheses 1.1 and 1.2 have already been reported (pp.113-116), in the discussion on the validity of the cognitive style instruments. In null form, the hypotheses became:

- 1.1 The girls and boys in the population studied will not differ in the number of relational responses made on the cognitive style instrument.
- 1.2 The girls and boys in the population studied will not differ in the number of descriptive responses made on the cognitive style instrument.

The cognitive style scores of the random sample from the experimental school manifested significant differences between the sexes for both descriptive ($t = 2.412$, d.f. 56, $p = < .02$) and relational ($t = 2.432$, d.f. 56, $p = < .02$) modes. The null hypothesis cannot, therefore, be sustained.

The direction of the between-sex differences (boys $>$ girls in descriptive responses, girls $>$ boys in relational responses) accords with the findings in other studies (e.g. Kagan et al., 1963), and with theoretical expectations.

Other statistical data, relevant to cognitive style and sex, emerged from intercorrelations computed in testing for the independence of I.Q. from cognitive style, preparatory to the analyses of covariance. These are reported here because of their relevance, and because of their bearing on the findings.

TABLE XIV: CORRELATION COEFFICIENTS BETWEEN COGNITIVE STYLE AND OTIS I.Q. SCORES

	I.Q.	Descriptive	Categorical	Relational
I.Q.	-	.040	-.041	.148
Descriptive	.097	-	.155	-.739
Categorical	.132	.005	-	-.687
Relational	-.135	-.610	-.709	-

(Girls above diagonal, boys below. N = 60 for each sex)

For both the boys and the girls in the experimental groups, there is a negative relationship between relational and categorical modes, and between descriptive and relational modes. When the near-zero relationship between descriptive and categorical modes is taken into account, it appears that three cognitive style dimensions are being represented. However, when it is remembered that the scores were normalised for each sex on the cognitive style measure, and for I.Q. across the sexes, and also that the sample was a stratified

one, it will be realised that the obtained correlations may result, in part at least, from the assignment procedures. Thus, the statistical pattern may be confirmation of the independence of the groups rather than evidence for the independence of preferred styles. Furthermore, the nature of the instrument itself may be a factor influencing the correlational pattern. While any one descriptive response may be given for one of several appropriate reasons, the choice of one style of response necessarily means a lower score on the other two modes. But to claim that the results are a test artifact for an either-or, two-choice situation (cf. Kagan et al., 1963), does not explain the pattern obtained on triads - a three-choice situation. It does seem possible, then, in spite of the qualifications made, that each of the three styles has its own attributes, and has a complex relationship with the other two. It may well be that the relationship between the styles, as well as the preference for one rather than another, is an important aspect of an individual's cognitive functioning. Kagan has argued in several places that the same styles may have different meanings for the sexes. The relative preferences between the sexes suggested by the research hypotheses do not contradict such an assertion.

In order to test further the hypothesis of sex differences in conceptual style, the raw scores of all the Form I children tested ($N = 302$, 157 boys and 145 girls) were subjected to statistical examination - means, standard deviations and t-values being calculated. The results are summarised in the table which follows.

TABLE XV: SEX DIFFERENCES IN COGNITIVE STYLES
AMONG FORM I PUPILS

Statistic	Descriptive		Categorical		Relational	
	Boys	Girls	Boys	Girls	Boys	Girls
N	157	145	157	145	157	145
Mean	6.97	5.33	6.62	5.76	10.51	12.82
S.D.	3.98	3.31	4.61	4.34	5.80	5.98
d.f.	. 300		300		300	
t	3.890		1.673		3.255	
p	< .001		< .10 (N.S.)		< .001	
Direction	B > G		-		B < G	

(t-tests two-tailed)

The Form I population of the experimental school manifested significant differences between the sexes in the descriptive and relational styles. It is reasonable to reject the null hypothesis, for this data and for this population. Thus, further confirmation of an earlier finding with a random sample (N = 58, p.114) provides evidence of sex differences between cognitive styles.

THE MAJOR HYPOTHESES.

The null form of the hypotheses, as the procedural form for testing statistical significance at the five percent level, in the cases of Hypotheses 2.3 and 2.4, require little change from the research formulation. The six procedural hypotheses are, then:

- 2.1 Pupils receiving Treatment R.E. will not differ, at more than chance levels, in their mean performance on the test of initial learning, from those receiving Treatment R.D.
- 2.2 Pupils receiving Treatment R.D. will demonstrate no more than chance differences in mean scores on the application-transfer tests, from those receiving Treatment R.E.
- 2.3 There will be no significant differences in mean scores on the test of retention between pupils receiving Treatment R.D. and those receiving Treatment R.E.
- 2.4 There will be no significant differences between Teachers, as measured by pupils' mean scores on post-tests of initial learning, retention, and application transfer.
- 2.5 There will be no significant differences between Treatments as measured by the percentages derived from coded observation scores for Categories 4, 5, 8 and 9 of an Amidon-Flanders Interaction Analysis.
- 3.1 For pupils whose preferred cognitive style is descriptive, no more than chance differences in mean scores on post-tests will distinguish those receiving Treatment R.D. from those receiving Treatment R.E.
- 3.2 For pupils whose preferred cognitive style is relational, no more than chance differences in mean scores on post-tests will distinguish those receiving Treatment R.D. from those receiving Treatment R.E.

The analysis of variance used to evaluate the hypotheses is an extension of the examples given by Lewis (1968, pp.130-152), in his account of designs with nesting and crossing. One of the reasons for using a nested design is that it reduces the error term, and so increases the precision of the contrasts made. The justification for the choice of error term lies in the mean squares expectation for the components of the linear model (expanded in Appendix C). The four error terms chosen from the components analysis to test each source of variation

in the design are the mean squares for:

- (a) between pupils (P) within A x B x C x D
(the four major factors),
- (b) pupils x objectives (P x E) within
A x B x C x D,
- (c) pupils by occasions (P x F) within
A x B x C x D, and
- (d) the residual variation (P x E x F) within
A x B x C x D.

Thus, for main effects and interactions not involving objectives or occasions, the error term is P; for main effects and interactions involving objectives and occasions, the error terms are P x E and P x F respectively; while the residual component, P x E x F, becomes the test for interactions involving both objectives and occasions. One caution to be observed with a large number of interactions in a six-factor (extended to seven, with a nested factor) analysis is care in accepting all interactions which appear to be statistically significant. With such a large number, some might be expected to be significant on chance alone.

A table showing a selected summary of hypothesis-related results from the main analysis of variance follows.

**TABLE XVI: SELECTED SUMMARY OF HYPOTHESIS-RELATED
RESULTS FROM THE MAIN ANALYSIS OF VARIANCE**

Source of Variation	Sum of Squares	d.f.	Mean Square	Error Term	F Ratio	p
A. Teachers	1.875	1	1.875	P	< 1.00	
* B. Methods	140.833	1	140.833	P	15.41	(a)
* C. Cognit. Styles	23.829	2	11.914	P	1.30	
* D. Sex	22.533	1	22.533	P	2.47	
E. Objectives	0.133	1	0.133	PxE	< 1.00	
F. Occasions	0.208	1	0.208	PxF	< 1.00	
* B x C	12.280	2	6.140	P	< 1.00	
* B x D	9.076	1	9.076	P	< 1.00	
* B x E	16.876	1	16.876	PxE	6.87	(c)
* C x D	27.155	2	13.578	P	1.49	
D x F	9.634	1	9.634	PxF	4.33	(d)
A x C x E	16.462	2	8.231	PxE	3.35	(d)
* B x C x E	4.588	2	2.294	PxE	< 1.00	
* B x C x F	2.004	2	1.002	PxF	< 1.00	
* B x C x D	16.588	2	8.294	P	< 1.00	
A x B x D x E	7.500	1	7.500	PxE	3.05	(e)
A x B x E x F	7.008	1	7.008	PxExF	4.19	(d)
A x C x E x F	5.954	2	2.977	PxExF	1.78	
C x D x E x F	5.488	2	2.744	PxExF	1.64	
AxBxCxDxExF	9.612	2	4.806	PxExF	2.88	(e)
P within AxBxCxD	877.000	96	9.135	PxExF	5.47	(a)
PxE within AxBxCxD	236.000	96	2.458	PxExF	1.47	(d)
PxF within AxBxCxD	213.400	96	2.223	PxExF	1.33	(e)
Residual (PxExF) within AxBxCxD	160.400	96	1.671			
Summary:	Total sum of squares				1928.792	d.f. 479
	Between-cells sum of squares				441.992	d.f. 95
	Within-cells sum of squares				1486.800	d.f. 384
* = Hypothesis-related results. All other F ratios for interactions, not shown here, fall below 1.00						
	(a) p = .001		(c) p = .025		(e) p = .10	
	(b) p = .01		(d) p = .05		(N.S.)	

Hypotheses 2.1, 2.2 and 2.3.

Hypotheses 2.1 (initial learning), 2.2 (application-transfer) and 2.3 (retention), in their null form, predict no differences between treatments on measures of initial learning, retention (knowledge and understanding) and application-transfer. The main analysis of variance provided two relevant F-ratios whose magnitude, with the degrees of freedom available (1/96), could have been the result of chance factors less than five times in a hundred. The significant results were the first-order interaction (B x E, treatments x objectives), and the treatments main effects ($F = 6.87$, d.f. 1/96, $p < .025$; and $F = 15.41$, d.f. 1/96, $p < .001$ respectively). The B x E (treatments x occasions) interaction, however, was not significant. It can be seen that these results do not appear to sustain null hypotheses 2.1 and 2.2, but, at the same time, do not permit the rejection of null hypothesis 2.3. However, before a conclusion can be reached, further evidence from the analyses of covariance and the additional analyses of variance will need to be examined.

The research hypotheses suggested the achievement of different objectives with different treatments - a B x E interaction. This is shown graphically in Figure 1, using a set of hypothetical means.

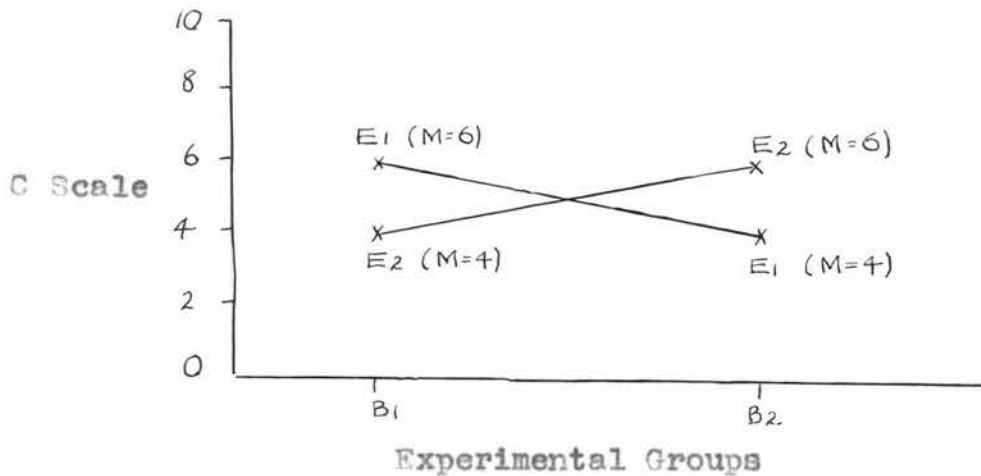


Figure 1: Predicted interaction B x E based on hypothetical data.

The obtained data, however, showed a non-symmetrical interaction pattern (Figure 2), whereas the predicted pattern was symmetrical in form.

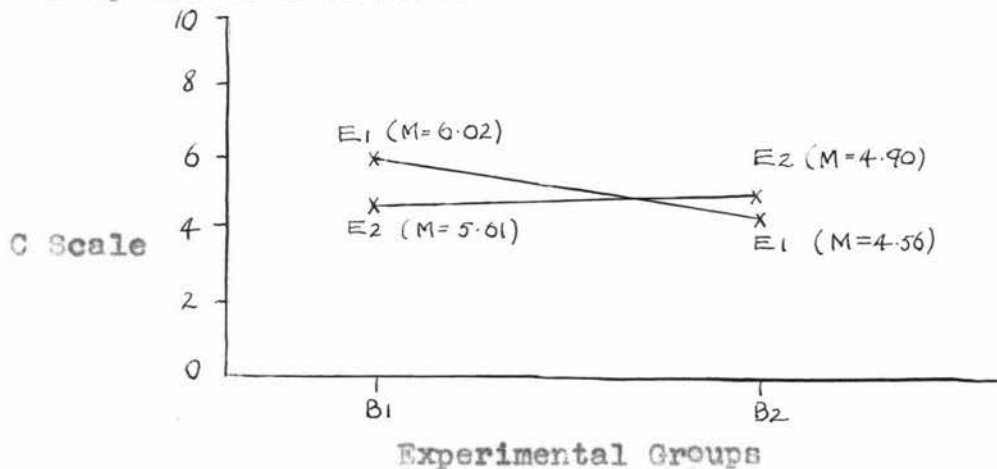


Figure 2: Obtained interaction B x E

The obtained B x E interaction is interpreted as showing the advantage of Treatment R.E. over Treatment R.D. for the knowledge and understanding objectives as indicated by performance on the tests of initial learning and retention (Tests IA and IB). It is apparent that this superiority is not maintained for application-transfer objectives.

The main analysis of variance, however, was essentially the test of Hypotheses 3.1 and 3.2, which were concerned with the interaction of cognitive style and treatments (teaching methods). For a test of the hypotheses particularly concerned with the methods, the four analyses of covariance and the four additional analyses of variance provide further evaluative material (see Table XVII). Each dependent measure was examined, in turn, by these statistics. Significant differences between treatments, for initial learning, were found in both the supplementary analysis of variance ($F = 22.43$, d.f. 1/96, $p = < .001$) and the analysis of covariance ($F = 19.46$, d.f. 1/95, $p = < .001$). Research Hypothesis 2.1 is thus supported, the direction of difference being that predicted (R.E. mean 6.17, R.D. mean 4.52). Further indirect confirmation of this result was produced when a comparison of mean performance of the two treatment groups, R.E. (mean 2.33; S.D. 1.11) and R.D. (mean 1.60; S.D. 1.01), on the five embedded items was made. The two-tailed t-test indicated a significant difference in favour of Treatment R.E. ($t = 3.782$, d.f. 1/118, $p = < .01$).

The data relating to Hypothesis 2.2 (null form) were rather equivocal. Two dependent measures were used, Test IIA and Test IIB. The former is probably the better on which to base comparisons, as it was of the same multiple-choice form as the two tests of knowledge and understanding, whereas the latter was a response-supplied form. Neither test had more than a moderate split-half reliability index. The data from Test IIA gave an F-ratio from the analysis of variance of 6.25 (d.f. 1/96, $p = < .025$) and for the analysis of covariance

of 3.43 (d.f. 1/95, $p = < .10$). Test IIB, with its larger error term, gave a between-treatments main effect F-ratio of 2.02 (d.f. 1/96, $p = < .20$) for the analysis of variance, and less than 1.00 for the analysis of covariance. To reject the null hypothesis categorically is to risk Type I error; to accept it equally categorically is to risk a Type II error. Such evidence as there is suggests a difference, but does not establish one.

Hypothesis 2.3, as a testing procedure, predicted no difference between treatments on the retention measure. The F-ratios for both the supplementary analysis of variance ($F = 12.45$, d.f. 1/96, $p = < .01$) and for the analysis of covariance ($F = 9.56$, d.f. 1/95, $p = < .01$) were both judged to be of sufficient magnitude, with their associate degrees of freedom, to be evaluated as showing significant differences between treatments for retention. The null hypothesis is therefore rejected. The consequence is, of course, to reject the research hypothesis which also predicted no differences. The differences are clearly in favour of Treatment R.E. (R.E. mean 5.70, R.D. mean 4.77) with these subjects, in these conditions, and with the dependent measures used. A comparison between the mean performance on the embedded items also supported the conclusion reached. Treatment R.E. (mean 2.36, S.D. 1.19) produced significantly higher mean scores than did Treatment R.D. (mean 1.55, S.D. 1.20), when two-tailed tests were run ($t = 3.75$, d.f. 1/118, $p = < .01$). Selected F-ratios for the analyses of variance and covariance are shown in Table XVII.

TABLE XVII: SUMMARIES OF MAIN EFFECTS AND
FIRST-ORDER INTERACTIONS FOR ANALYSES OF
COVARIANCE AND ANALYSES OF VARIANCE

Analysis of Covariance				Analysis of Variance		
Source	Adjusted d.f.	M.S.	F	d.f.	M.S.	F
<u>TEST IA: (INITIAL LEARNING).</u>						
A. Teachers	1	1.453		1	0.075	
B. Methods	1	55.318	19.46(a)	1	81.675	22.43(a)
C. Styles	2	0.978		2	2.034	
D. Sex	1	18.007	6.34(d)	1	7.067	1.90
A.B.	1	0.994		1	3.675	1.01
A.C.	2	7.714	2.71(e)	2	9.100	2.50(e)
A.D.	1	2.770		1	4.350	1.19
B.C.	2	6.230	2.19	2	3.100	
B.D.	1	0.000		1	1.817	
C.D.	2	4.271	1.50	2	7.204	1.91
Within	95	2.842		96	3.642	
Total	118			119		
<u>TEST IB: (RETENTION).</u>						
A. Teachers	1	1.599		1	0.300	
B. Methods	1	25.379	9.56(b)	1	48.133	12.45(b)
C. Styles	2	1.016		2	2.158	
D. Sex	1	8.387	3.16(e)	1	0.833	
A.B.	1	0.043		1	0.834	
A.C.	2	3.190	1.20	2	4.575	1.18
A.D.	1	4.186	1.58	1	6.534	1.69
B.C.	2	4.648	1.75	2	2.808	
B.D.	1	1.008		1	0.834	
C.D.	2	1.365		2	5.308	1.37
Within	95	2.654		96	3.867	
Total	118			119		
All F-ratios below 1.00 are omitted. All higher order interactions were non-significant. (See Appendix D).						
(a) p = .001		(c) p = .025		(e) p = .10		
(b) p = .01		(d) p = .05		(N.S.)		

TABLE XVIII: (Continued)

Analysis of Covariance				Analysis of Variance		
Source	Adjusted d.f.	M.S.	F	d.f.	M.S.	F
<u>TEST IIA: (APPLICATION-TRANSFER)</u>						
A. Teachers	1	16.165	6.96(b)	1	4.408	1.18
B. Methods	1	7.964	3.43(e)	1	23.408	6.25(c)
C. Styles	2	2.902	1.25	2	4.758	1.27
D. Sex	1	52.923	22.78(a)	1	27.075	7.23(b)
A.B.	1	2.173		1	0.076	
A.C.	2	3.742	1.61	2	3.058	
A.D.	1	1.916		1	3.752	
B.C.	2	2.122		2	1.658	
B.D.	1	0.026		1	3.675	
C.D.	2	0.741		2	4.375	1.17
Within	95	2.323		96	3.746	
Total	118			119		
<u>TEST IIB: (APPLICATION-TRANSFER; OPEN-ENDED TEST).</u>						
A. Teachers	1	9.851	3.11(e)	1	2.133	
B. Methods	1	2.378		1	8.533	2.02
C. Styles	2	3.004		2	5.034	1.19
D. Sex	1	6.753	2.13	1	0.533	
A.B.	1	8.697		1	16.134	3.81(e)
A.C.	2	0.808		2	1.434	
A.D.	1	2.927		1	4.801	1.13
B.C.	2	2.570		2	2.434	
B.D.	1	0.070		1	3.333	
C.D.	2	1.796		2	1.234	
Within	95	3.168		96	4.233	
Total	118			119		
All F-ratios below 1.00 are omitted. All higher order interactions were non-significant. (See Appendix D).						
(a) p = .001		(c) p = .025		(e) p = .10		
(b) p = .01		(d) p = .05		(N.S.)		

Comparison of Score Patterns.

One feature not commented on so far concerns the relationship between the scores on Test IA (initial learning) and scores on the remaining three measures. While pupils receiving Treatment R.E. gained generally higher mean scores than did pupils receiving Treatment R.D., there was a pattern revealed by the relative mean scores of each group. The children in the R.E. condition had mean scores for Tests IB (retention) and IIA and IIB (application-transfer) which were lower than the mean score for Test IA. In contrast, the children in the R.D. condition, for the same tests, had mean scores which were higher than the initial learning mean score. Figure 3 shows the contrasting pattern.

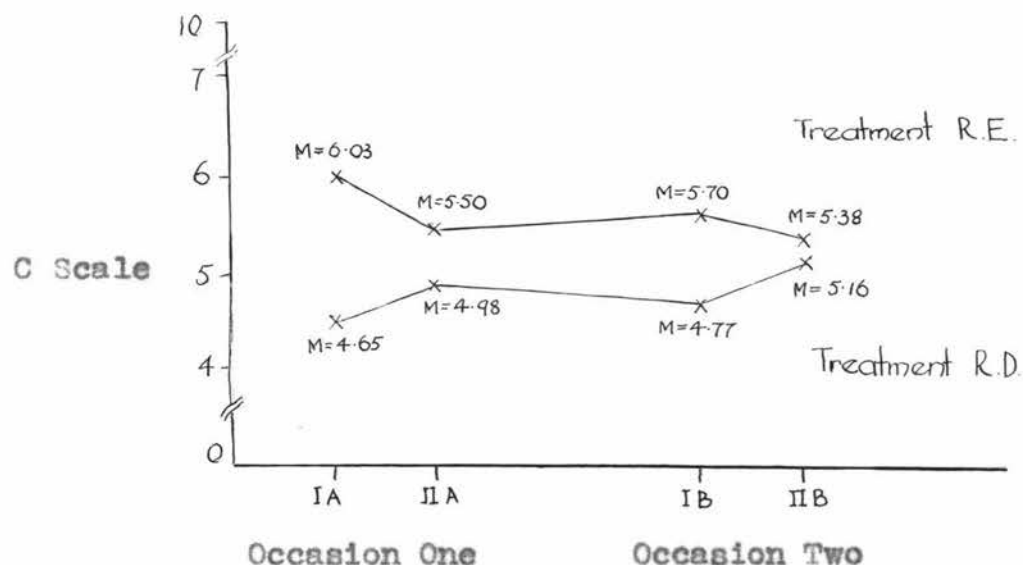


Figure 3: Relative scores on the four post-tests in order of administration across time (adjusted means).

While the results have supported only the hypothesised advantage of Treatment R.E. over Treatment R.D. for initial learning, the pattern of relative scores is in accord with

the predictions. A "post-experimental gain" for Method R.D. has been noted in other studies, e.g. Wittrock (1963a). It has been suggested that tests themselves provide a learning situation, but that does not explain why a different pattern was observed with method R.E. The obliterative subsumption argument proposed by Ausubel (1968) could well provide a satisfactory explanation of the observed pattern.

Hypotheses 3.1 and 3.2.

Null hypotheses 3.1 and 3.2 cannot be rejected by the main analysis of variance evidence, as the F-ratios for second-order interactions B x C x F (methods x styles x occasions), B x C x E (methods x styles x objectives), first-order interaction B x C (methods x styles), and for cognitive style main effects, all fall below 1.00. It cannot, however, be claimed that cognitive style has no effects other than chance influences, or that there is no interaction between preferred modes, different objectives, and teaching situations. Indeed, the significant interaction ($F = 3.55$, d.f. 2/96, $p < .05$) between teacher, cognitive style and objectives (A x C x E) does suggest a dynamic in the situation which is not easily explained. The teachers were similar in many ways, although one was female, the other male. This is one possible factor, although no significant interactions were noted between sex of pupil and teacher (A x D).

Another possibility, in view of the significant methods x objectives (B x E) first-order interaction, is that

some subtle aspect of teacher style interacts with cognitive mode and objectives. An informal and circumstantial clue to such a possibility was observed in the pilot testing of the cognitive style instrument. Of the two classes used, one with a male teacher had a high proportion of relational responses, while the other, with a female teacher, had a high proportion of descriptive responses, the distributions in each class being almost mirror images of one another. As the groups were not random samples, to draw any conclusions is, however, only speculation.

Still another possibility emerged from the analyses of covariance. Main effects for teachers, when I.Q. variance was removed, were indicated for both application-transfer tests ($F = 6.96$, d.f. 1/95, $p = < .01$, and $F = 3.11$, d.f. 1/95, $p = < .10$), but not for the tests of initial learning or retention, for which method was a much more potent factor. The possibility remains, then, that some teacher characteristics, be they teaching style or some other factors, interact with pupil characteristics to facilitate the attainment of some objectives better than others. It has been recognised that some interactions might be spuriously significant, but in this case, the psychological plausibility of teacher effects seems sufficient to be considered actual.

Thus, while the null hypotheses cannot be rejected, within the limits and data of this study it cannot be claimed that cognitive style has no influence. The error term for cognitive style, and for the cognitive style-method interaction in the main analysis is symbolised as P (pupil variance within

A x B x C x D). An examination of the data suggests that individual variability of pupils is itself a major source of influence ($F = 5.47$, d.f. 96/96, $p = .001$), and it is not surprising that cognitive style, as one individual difference variable, contributes insufficient variance to produce an F-ratio which can be evaluated as significant. This is also the case when the variance attributable to I.C. is removed by analysis of covariance.

An unexpected finding: sex differences.

One result not predicted was the significant D x F (sex by occasion) interaction in the main analysis of variance ($F = 4.33$, d.f. 1/96, $p = < .05$).

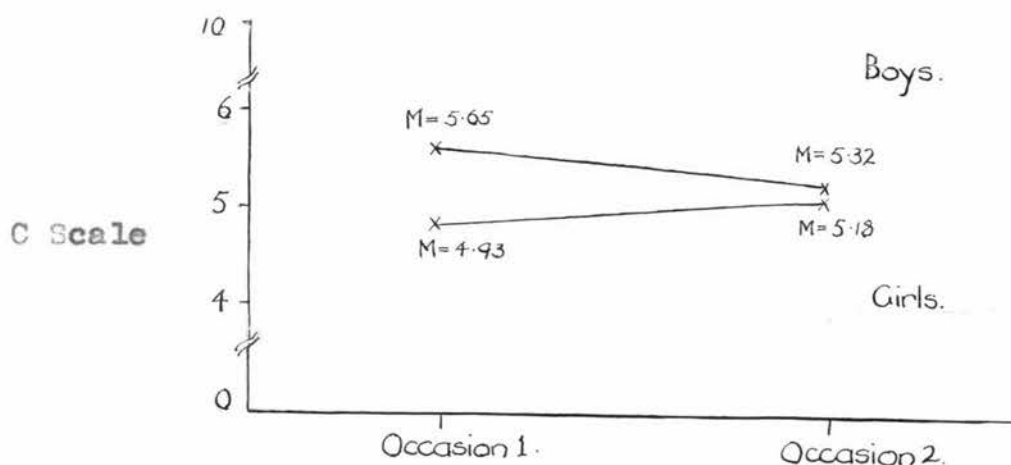


Figure 4: Interaction of sex of pupils and two testing occasions

Sex differences were also evident in the two analyses of covariance based on test results for Science Test IA and Science Test IIA, both tests having been administered on the first post-test occasion. One explanation is that boys tend

to be more interested in science than girls, and so scored at a higher level on the first occasion. This, however, does not explain the improvement in the girls' scores on the second testing occasion. The boys' scores, while remaining at a higher level than the girls', showed a decrease at the second testing. Furthermore, when the means were adjusted for I.Q., the differences between the scores were more marked, but still the same pattern of scores emerged. Figure 5 shows this pattern of adjusted scores, not as a true interaction diagram, but in a way which shows the four tests and the relative performances of the boys and the girls.

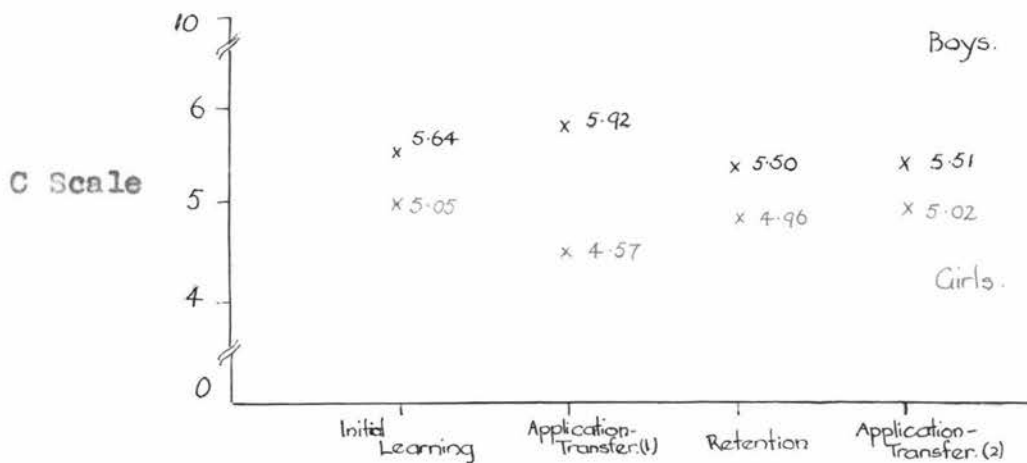


Figure 5: Adjusted mean scores for boys and girls on the four dependent measures.

A feature of the figure was the mirror imaging displayed between the initial learning and the application-transfer results, on Occasion One. It was suggested earlier that Test IIA might be more satisfactory than Test IIB. If this is so, much of the interactional pattern can be understood

in terms of the objectives measured by Test IIA. The analysis of covariance for Test IIA provides a main effect F-ratio for sex which is highly significant ($F = 22.78$, d.f. 1/95, $p = <.001$). Thus, while the main analysis of variance gives a significant D x F interaction, it is suggested that performance on one measure of objectives may account for the obtained interaction. It seems psychologically plausible for boys, who are expected to display more independence than girls (Kagan and Moss, 1962), to display this also in application-transfer tasks, particularly in science. Cofer (in Gagné, 1967, p.138) reports a Sex x Methods interaction, and, in his discussion, recognises the difficulty of making any conclusive interpretation. A further tentative explanation is that regression effects are operating differentially. The initial differences in mean scores between the sexes are such that, if regression to the mean is to occur, only one pattern of regression is possible.

CONTROL GROUP PERFORMANCE.

The control group was, as has been stated, external to the statistical design. The functions of this group, in the investigation, were to provide a check against accidental factors (e.g. television programmes on the science topic), against practice effects with the embedded items, and, should it have been required, against spurious conclusions about attitude changes as measured by the attitude scales. In the

end, the last function was not necessary, as little attitude change occurred. A simple comparison of means of the scores on the five embedded items showed that little practice effect was operating as far as the control group was concerned. The mean scores for the three occasions, one in the pre-test, the other two in post-tests, were 0.70, 0.90 and 0.86. The experimental groups' scores reveal a different pattern. The mean scores of the two treatment groups, the first two means of which were from tests answered in common with the control group, were 0.97, 2.33, 2.36 and 0.93, 1.60, 1.55 respectively. The control group showed a small, non-significant ($t = 0.90$) gain from pre-test to post-test. On the other hand, the four experimental classes all revealed significant gains (t 's range 2.28 to 5.92, p 's .05 to .01, d.f. 1/58). It seems reasonable to conclude that such gains as were observed with the experimental groups arose as a result of the teaching, and not through extraneous influences.

TESTS FOR INDEPENDENCE OF TREATMENTS.

The analysis of variance produced F -ratios which indicate no significant differences between teachers, but did indicate very significant treatment differences. To what extent can the treatments be regarded as independent? The null hypothesis formulated to guide the testing of independence of treatment is given below:

- 2.5 There will be no differences between Treatments R.D. and R.E. in percentages recorded on Categories 4, 5, 8 and 9 of an Amidon-Flanders Interaction-Analysis of teacher-pupil behaviour.

The coded observations for each teacher, in each of the treatments, were combined in Work Matrices (see Appendix E) to give matrices for each teacher, each treatment, for teacher across treatments, and for treatments across teachers. It is the latter matrix which is at present being examined. The categories for interaction analysis are summarised below. Category 10 includes, for this study, pupil practice and manipulation, and problem solving with equipment.

Summary of Categories for Interaction Analysis.

1. Accepts feeling: accepts and clarifies the feeling tone of the students in a nonthreatening manner. Feeling may be positive or negative. Predicting or recalling feelings is included.
2. Praises or encourages: praises or encourages student action or behaviour. Jokes that release tension, but not at the expense of another individual. Nodding head, or saying "um hm?" or "go on" are included.
3. Accepts, clarifies or uses ideas of students. clarifying, building or developing ideas suggested by a student. As teacher brings more of his own ideas into play, shift to Category 5.
4. Asks questions: asking a question about content or procedure with the intent that a student answer.
5. Lecturing: giving facts or opinions about content or procedures; expressing own ideas, asking rhetorical questions.
6. Giving directions: directions, commands, or orders with which a student is expected to comply.
7. Criticising or justifying authority: statements intended to change student behaviour from non-acceptable to acceptable pattern; 'bawling someone out'; stating why the teacher is doing what he is doing; extreme self-reference.

8. Student talk - response: talk by students in response to teacher. Teacher initiates the contact or solicits student statement.
9. Student talk - initiation: talk by students, which they initiate. If "calling on" student is only to indicate who may talk next, observer must decide whether student wanted to talk. If he did, use this category.
10. Practice and problem solving: group work, manipulation of apparatus, activity following teacher directions or questions. Includes pauses, short periods of silence, and periods in which communication cannot be understood by observer.

In total, 2397 observations were coded (1277 for one teacher and 1120 for the other), one thirty-minute tape being lost due to failure of the recorder. The totals, percentages and the significance of differences between percentages (Garrett, 1958, pp.235-238) are tabulated below.

TABLE XVIII: DIFFERENCES BETWEEN TREATMENTS:
PERCENTAGES FOR TEN CATEGORIES OF AN
AMIDON-FLANDERS INTERACTION ANALYSIS

	1	2	3	4	5	6	7	8	9	10
R.E. (totals)	5	48	29	116	651	24	9	37	77	163
R.D. (totals)	18	107	34	287	94	134	8	103	154	299
R.E. (%)	0.45	4.16	2.52	10.06	56.46	2.08	0.78	3.21	6.68	14.14
R.D. (%)	1.45	8.64	2.75	23.18	7.59	10.82	0.65	8.32	12.44	24.15
t	<1.0	4.533	<1.0	8.747	26.416	8.793	<1.0	5.544	4.881	6.335
p	N.S.	.01	N.S.	.01	.01	.01	N.S.	.01	.01	.01

(N for R.E. = 1159; for R.D. = 1238. Total 2397. Because of rounding-off effects with decimals during calculations, percentages do not total 100 exactly.)

The significant differences between treatments in Categories 4, 5, 8 and 9, at the .01 level, are sufficiently different from chance predictions to support a rejection of the null hypothesis. With further evidence arising from the non-significant main effects from teachers in the analysis of variance, it may be argued that the treatments were indeed independent. A feature of the table is the sharpness of the distinctions made; no category falls between the one pole of non-significance and the other of marked significance. Some of the distinctions apparent in the table were not predicted by the hypothesis, but are understandable in terms of the treatments. Treatment R.D., by definition and observation, is characterised by a great many question and answer sequences, and hence provides many opportunities for praise and encouragement. The problem solving activity is shown by the percentage (24.45) in Category 10. At the same time, many of the principles to be derived in Treatment R.D. were explained and demonstrated in Treatment R.E., followed by practice (14.14 % in Category 10). The non-significance of differences in Categories 1 and 3, and in teacher justification (Category 7), suggest, as was hoped, that the teachers were generally similar in warmth and supportiveness in both treatments.

Additional data giving substance to claims for the independence of the treatments were gained from a teaching pattern analysis of the major moves. Following the criteria and rules in the Teaching Pattern Analysis Manual (Amidon and Amidon, 1967), the work matrices were examined and three-step major teaching moves identified. For Treatment R.E., the

matrices for both teachers showed the pattern 5 - 4, 4 - 8, 8 - 2; i.e., teacher exposition, followed by a teacher question, with the pupil response being, in turn, followed by praise or encouragement. For Treatment R.D., the first three steps in the major move covered the same three cells with slightly different emphases. The cells were 4 - 8, 8 - 2, 2 - 4; i.e. teacher question, followed by pupil response, teacher praise, and then a further teacher question. These major patterns are consonant with the criteria defining the treatments.

The Amidon-Flanders is an instrument for which good reliability is claimed, with trained observers. However, there are many subtle interchanges which are not included in the system, and, furthermore, it is based on an observer's frame of reference (cf. Westbury, 1967-68). This study did not make provision for a pupil-report instrument as a cross check. There were some signs that this would have been valuable. For example, the interaction analysis data suggested (non-significantly) that Teacher One was a little warmer, a little more positive in approach to children than was Teacher Two. An algebraic addition of the positive and negative changes of direction in attitudes between pre- and post-testings gave some slight evidence to suggest that the pupils might have a different perception. The evidence from the Pupil Opinion Survey is very tentative because of the stability of the scores on the instrument; regression effects could count, in substantial measure, for directional changes. The overall results from the interaction analysis data, nevertheless, seem

sufficiently clear to confirm the claimed independence of the treatments.

TEACHER EFFECTS.

Reference has already been made to the fact that the F-ratio in the main analysis of variance for teacher main effects was much smaller than 1.00. The non-significance of teachers as a source of variance is regarded as supporting the null hypothesis (2.5), and, thus, the research hypothesis. The work matrices from the interaction-analysis were combined to give category percentages for each teacher across treatments. The t-value for each of the ten categories fell below 1.00 in each case, and failed to reach the criterion of significance. There was, then, agreement between the analysis of variance data and the observations coded in the interaction-analysis system, supporting the null hypothesis conclusion indicated above, and providing some indication of the validity of the teacher observation method used. The lack of statistically significant differences between the teachers is seen as confirming the use of trained student teachers, rather than experienced teachers, in the experiment, since no "method reversion" tendencies were apparent.

However, while there is considerable evidence for teacher similarity in the main analysis of variance, the interaction-analysis data, and six of the eight additional analyses, some discrepancy was observed in two of the analyses of covariance. When I.Q. was covaried, the analyses for the

two application-transfer tests showed increased main effects for teachers (Test IIA, $F = 6.96$, d.f. 1/95, $p = .01$; Test IIB, $F = 3.11$, d.f. 1/95, $p = < .10$). At the same time, treatment main effects decreased, and main effects attributable to the sex of pupils, particularly for Test IIA, increased. Nonetheless, the teacher-by-sex-of-pupil interaction was non-significant. ($F = < 1.00$ in each case). Such results are not easily interpreted. Among possible explanatory factors might be suggested the sex-differences in I.Q. noted by Elley (1969, pp.140-145), while re-standardising the Otis test. It is also possible that some subtle aspect of teacher style facilitates the achievement of some objectives more than others (see p.152). The significant A x C x E interaction in the main analysis of variance is not incompatible with this possibility. A third possibility is that, given similar general ability, boys are more likely to seek to apply and transfer their knowledge than are girls. The boys' adjusted mean on Test IIA was 5.92, the girls' 4.57 (see p.154).

Although some qualification has been made, the evidence overall for consistency between teachers, as well as for the independence of the two treatments, has considerable strength.

CHAPTER IXSUMMARY, DISCUSSION AND GENERAL
CONCLUSIONSSUMMARY OF THE INVESTIGATION.

The present investigation was concerned with the possible interaction between three cognitive styles (descriptive-analytic, categorical-inferential and relational-contextual) and two teaching methods (rule-explained and rule-derived). Three groups of hypotheses were formulated. The first group postulated sex-differences in cognitive style. Boys were expected to make more descriptive responses than girls; girls were expected to make more relational responses than boys. The second group posited differential achievement of objectives for the two teaching treatments. Method R.E., it was proposed, would lead to higher initial learning scores on a test of knowledge and understanding; method R.D. was expected to lead to higher scores on tests of application-transfer; and neither method was anticipated to produce better retention of knowledge and understanding than the other. Two hypotheses were a subset of the second group, predicting, respectively, no differences between teachers as indicated by post-test scores, and significant differences between treatments on Categories 4, 5, 8 and 9 of an Amidon-Flanders Interaction-Analysis. The third group of hypotheses posited interactions between the cognitive styles and the teaching methods. Pupils whose preferred style was descriptive were

predicted to score more highly than others on post-tests, after receiving Treatment R.D., while pupils whose preferred mode was relational were expected to demonstrate more effective learning under Treatment R.E.

The subjects, assigned at random to cells, were equal numbers of boys and girls who had scored at a Stanine level of seven or higher on one category of a specially-constructed, cognitive style instrument. All were Form I pupils from a city Normal Intermediate School, and, apart from the high scores made on the measure of conceptual style, displayed no characteristics which would make them atypical of other children of their age. (Mean age 143.62 months, mean Otis I.Q. 112.55). In all, 120 children (60 boys and 60 girls) were assigned to the four experimental classes, together with 30 to a control group, and a further 30 to a reserve group having the same basic composition (5 pupils x 2 sexes x 3 cognitive styles) as the other classes.

Selected second-year Teachers College students studying science were trained over a period of three months to act as observers and teachers. The experimental teachers (one male, one female, of similar age and ability) were trained to teach both methods. Each taught two classes, (one by each method), giving two fifty-minute lessons to each class. The content material, closely related to the present Forms I to IV science curriculum, was the concept of stability, with its associated principles. Time of day, classrooms, and the order in which the lessons were taught, were all crossed to maintain equivalence between conditions. Tape recordings of

sixty percent of each lesson provided data for the test of the hypothesised treatment differences.

An experimental design was developed which resembled a $2 \times 2 \times 3 \times 2 \times 2 \times 2$ factorial experiment, with teachers, methods, cognitive style, sex, objectives, and occasions being the six factors. Scores from different pupils were obtained for different levels of teacher, method, cognitive style and sex, while scores from the same pupils were also obtained for levels of objectives and occasions. Pupils therefore were regarded as an additional factor nested within teacher, method, cognitive style and sex, and crossed with objectives and occasions. The basic models from which this design was developed are reported in Lewis (1968).

Four dependent measures were constructed to provide data on the achievement of two levels of objectives (knowledge and understanding, and application and transfer) on two occasions (the day following the conclusion of the experimental teaching, and fourteen days later). All post-tests were administered in the school hall, to all subjects at the same time, so equalising the conditions of testing.

Scores from each dependent measure were normalised, using C-scale units. These scores became the dependent measures in the experimental design, each score being classified in seven ways: i.e. as belonging to a particular teacher, method, cognitive style, sex, objective, occasion and pupil.

In addition to the main analysis of variance, an analysis of covariance (with I.Q. covaried) and a supplementary

analysis of variance were conducted with each dependent variable in turn.

SUMMARY OF THE FINDINGS.

1. Hypotheses 1.1 and 1.2, predicting sex differences in preferred cognitive style, were supported by the observations. Boys in the Form I population in the experimental school made significantly more descriptive responses than did the girls ($t = 3.890$, d.f. 300, $p = < .001$). The girls in Form I in the school made significantly more relational responses than did the boys ($t = 3.255$, d.f. 300, $p = < .001$).
2. Hypothesis 2.1 predicted higher initial learning scores for pupils receiving Treatment R.E. than for those receiving Treatment R.D. Significant differences between the means in the direction postulated permitted the acceptance of Hypothesis 2.1
3. Hypothesis 2.2, predicting higher scores on tests of application-transfer for pupils taught by Treatment R.D. as compared with those taught by Treatment R.E., is rejected, as the mean differences were no greater than would be expected on the basis of chance alone.
4. Hypothesis 2.3 predicted no difference in retention scores between the two treatments, but is rejected. The obtained differences were significantly in favour of Treatment R.E.
5. Hypotheses 2.4 and 2.5 were concomitants of the experimental design. The former, proposing no differences

between the teachers as measured by post-test scores, is confirmed by the evidence from the main analysis of variance and six of the eight subsequent analyses. The latter, which posited significant differences between the treatments, is also supported by the data from the interaction analysis.

6. Hypotheses 3.1 and 3.2 were central foci in this study. They predicted interactions between cognitive styles and teaching methods. The first hypothesised, for pupils whose preferred mode was descriptive, better post-test performance under Treatment R.D. than under Treatment R.E. The second predicted better post-test performance following Treatment R.E., in contrast to Treatment R.D., for pupils whose preferred mode was relational. Neither of these hypotheses can be supported by the evidence of the investigation.
7. One finding not predicted was the sex x occasion interaction in the main analysis of variance. An interpretation offered was that, while girls performed fairly consistently on the tasks involving knowledge and understanding, they did not perform as well, relatively speaking, as the boys on the tasks involving application and transfer. However, it was also suggested that regression effects could have occurred, acting differentially, to produce an apparent rather than a real interaction.

DISCUSSION AND GENERAL CONCLUSIONS.

The major research question for which an answer was sought in this study concerned the interaction of cognitive style and two teaching strategies, believed to be commonly employed in New Zealand schools, in the teaching of science. The evidence from this investigation does not indicate that cognitive style was a major factor in the learning of these children, in this experimental situation. It cannot be claimed, on the other hand, that cognitive style will not be an important factor in other situations, with other content, with different age groups. That other individual difference variables were operating is apparent when the highly significant Pupils within - A x B x C x D interaction is considered. The use of I.Q. as a covariate in the analysis of covariance has little effect on the variance contributed by cognitive style, although it did in this way confirm the relative independence of conceptual mode from general ability as defined by the Otis test.

The stability of the defining measure, and the sex-different patterns revealed by the scores, seemed to suggest that the instrument was reliable, and was measuring the same dimensions with some consistency over time. The validation of the instrument on two grounds (the pattern of sex differences, and, for the experimental groups, extreme response sets) rests on the extent to which the observations are in accord with Kagan's earlier findings, and are consonant with those which can be predicted by extrapolation from the hypothetical construct. The argument of coherence does appear to be reasonable

in view of the obtained results. Consequently, it is assumed that individually consistent conceptual styles were identified and manipulated.

If the styles are accepted to be preferred modes, it is not entirely unexpected that cognitive style interactions and main effects fail to be manifested. The strong effects of the teaching treatments, the highly significant individual difference factors, and the complexity as well as the novelty of the situation, provide, along with other dynamics, a nexus of variance from which to partition one variable may have been to anticipate too much. The lack of definition in the findings concerning cognitive style may be a consequence of situational dominance. Wachtel's (1968) findings support such a view. Katz (1968, pp.233-238), investigating the role of irrelevant cues in the formation of concepts by lower class children, concluded that "reflection was not a general response characteristic, but rather one that was appropriately related to the stimulus characteristics of the task." And for this study, the stimulus features were complex, and, apparently, very strong. More precise experiments in better controlled situations may, nevertheless, demonstrate preferred cognitive modes as significant individual difference variables. Even within this investigation, some tentative clues are found in the A x C x E interaction in the main analysis, in the analysis of covariance interactions A x C ($F = 2.71$, d.f. 2/95, $p = .10$), and B x C ($F = 2.19$, d.f. 2/95, $p = < .20$) for initial learning scores, and in the supplementary analysis of variance A x C interaction ($F = 2.50$, d.f. 2/96, $p = .10$), also for

initial learning. When the dependent measures were converted to a common scale to give a measure of attainment in science based on the experimental lessons, a directional (non-significantly different) range of means was observed (descriptive 5.56, categorical 5.24, relational 5.01). Such indications are, however, slight. The limitations of the dependent tests and, possible, the coarseness of the C-scale, may have acted to reduce the accuracy of the observations and the precision of the analyses.

While a great deal of caution must be exercised in considering the cognitive style variable, a little more assurance may be felt regarding some of the findings for the treatment variables. The results in this investigation indicate that the expository-type procedure was superior to the rule-derived-with-guidance procedure for initial learning and retention, as measured by Tests IA and IB. The two tests were designed to measure performance at the knowledge-understanding level. On the other hand, the rule-explained-and-demonstrated procedure showed no significant superiority over the R.D. treatment as far as scores on Tests IIA and IIB (application-transfer) were concerned. While the statistical evidence seems clear, some qualifications are necessary.

The first qualification relates to a difficulty this investigation shares with a number of the research studies reported (Cronbach, 1966, pp.83-84). The difficulty is whether to equate time between treatments, as was done in the experiment here reported, or to train both groups to the same criterion. Had the latter course been adopted, a number of

changes, both in experimental design and results, would have been likely consequences. The discrepancy between the initial learning scores of the two groups was such that it was unlikely that pupils in condition R.D. could have gained similar retention scores to those in condition R.E. A second qualification relates to the post-tests. The within-cells variance of the supplementary analyses of variance reflected the moderate reliability of the instruments. The extent to which I.Q. was a source of error variance was shown by the reduction of the error terms in the analyses of covariance. Thus, while the post-tests were less than optimal, the value of analysis of covariance in increasing the precision of the statistical design is demonstrated in this experiment. The third qualification is again related to an issue Cronbach (*ibid.*) has raised in connection with the logic of experiments on discovery - the need for long-term investigations in the fields of curriculum and teaching methodology. In the study under discussion two teaching periods only, (one hundred minutes of teaching) is a limiting condition. "Educational development comes through continued instruction with intellectually significant subject matter and that is what we should investigate." (*ibid.*, p.90). In spite of this last qualification, the statistical evidence provides clear indications of quite marked effects from such a brief period of instruction.

The pattern of the mean scores between the two treatments has some theoretical and research support. Ausubel, on many occasions (e.g. 1968) has argued that as the subsump-

tion process continues with the integration of new material into cognitive structure, there may be some loss of specific detail. The decrease in the magnitude of the means from initial learning to retention, under Treatment R.E., is, then, consistent with a theory of meaningful reception learning. On the other hand, the means for Treatment R.D. display a reversal of the pattern, increasing in magnitude. A number of research reports (e.g. Wittrock, 1963a) have referred to "post-experimental gain" with "minimally directed" groups. The retention and application transfer scores in this investigation exceeded the initial learning scores for the pupils in condition R.D.

The findings do not provide grounds for drawing conclusions about which were the major method variables. The general sequence (Worthen, 1968) for each treatment was the same, although there were systematic within-phase differences. Each treatment contained a large number of "instantial moves" (Nuthall, 1966). The general set differed for each, as did the amount of manipulation, practice and teacher talk. But, in this study, it is not possible to point to the prepotency of any one within-method factor.

While the findings reveal differences between methods, the interaction between methods and objectives was not only significant statistically, it is also of interest practically. It was argued earlier (p.97), that objectives were independent variables occupying a position of primacy in any teaching enterprise or model. The results do not contradict the contention, and, to an extent, they justify the inclusion of

objectives as an independent variable entering into the main analysis of variance. A further gain, not predicted but resulting from the design, was suggestive of a difference between the sexes in regard to objectives. The girls appeared, relative to their within-scope performance, to be less successful on the extra-scope tasks of application and transfer of Test IIA. The converse was observed for the boys, who were, relative to their within-scope performance on Tests IA and IB, more successful on the tasks of Test IIA. However, this observation needs further testing before any definite claims can be made.

It was not the intention of this study to attempt to establish the advantages of one teaching method over another. The method hypotheses were postulated in the belief that each method had its particular contribution to make in the achievement of different objectives. While only one of the hypotheses was confirmed, the general pattern of obtained scores was consistent with predictions.

To what extent can a degree of generality be claimed for any of the findings of this study? As a pilot study, the investigation had an exploratory orientation rather than one seeking definitive conclusions. Furthermore, a number of restrictions have been suggested in various places in the report. Consequently, the conclusions here proffered as having some generality beyond the limits of the experimental boundaries are few. Insofar as the children in Form I at the experimental school may be regarded as representative of first-year intermediate school children in New Zealand, it

does seem, with this age group, possible to identify one set of cognitive style variables which shows a degree of stability, and manifests different response patterns between the sexes. The pattern found in the study is similar to that found by Kagan (1963,1964) and others (e.g. Sigel, 1967) in American children. At the same time, it seems probable that cognitive style as a preferred mode is, in any specific situation, contingent on the extent to which stimulus properties demand or evoke any particular style of intellectual functioning.

A second general proposition offered as an extension of the findings of the investigation is that the choice of teaching method, or combination of methods, is one major factor deciding the extent to which particular objectives and classes of objectives will be achieved. In this study, as in many others (e.g. Worthen, 1968; Kersh, 1962), the expository procedure was the more effective means of facilitating the attainment of the more immediate knowledge and understanding objectives. The assertion is not unqualified, for the method used and the objectives set were those which were considered appropriate for the particular group of pupils. In addition, many factors in the teaching situation affect the attainment of objectives. For example, teacher style (Heil et al., 1961), learning style (Tallmadge and Shearer, 1969), the interaction of learner characteristics and instructional mode (Ripple et al., 1969), learner abilities (Dunham and Bunderson, 1969), are but a few of the recently reported factors. A factor not discussed is the effect of pupil familiarity with a particular teaching approach or set of

approaches, which may conceivably have influenced the findings of this study.

EDUCATIONAL IMPLICATIONS.

It is unlikely that any simple method of instruction can be found to serve all major cognitive goals. The teaching procedures employed in the experiment reported in this thesis were not equally effective in enabling pupils to achieve two classes of objectives. Treatment R.E. (an expository type of teaching) led to higher scores on tests of knowledge and understanding, but, relative to these scores, scores on tests of application and transfer were lower. Treatment R.D. (a guided-discovery type of teaching) led to a pattern of scores which showed gains on tests of application and transfer as compared with scores on the initial test of knowledge and understanding. A combination of the two methods may have been more effective overall than was either one alone. Johnson and Stratton (1966) compared four single methods of teaching concepts with a mixed method, and found the mixed method best on all criterion tests. The four single methods were about equally effective. Leith and McHugh (1967) found an expository theory passage in a mediating position to be most efficient for a difficult conceptual task. They conclude (p.116) that the place of theory "..... would seem to be after learning of particular concepts rather than before them, especially when the more difficult exemplar is given first." However, there was a significant order of presenta-

tion x ability level interaction which moderates the conclusion. It is suggested that teachers experiment in their own classrooms with varying combinations of expository and inductive teaching.

A further implication is related to teacher education. It was possible to train ten student teachers to observe and teach two quite distinctive methods. All of their education course for one term was taught through the training programme, in the context of a particular research task. The results indicate that the experimental teachers were able to employ both methods without confounding them. Informal conversation with the group of students left the strong impression that they had gained confidence in their teaching ability, in ability to plan, had understood and were able to apply a considerable body of learning theory. They revealed considerable interest in pursuing their individual investigations in science, in cognitive style, in attitudes to science and in regard to the appropriateness of different methods with individual children. It is, then, suggested that involving students in research and survey enterprises may be a beneficial approach in teacher training.

FURTHER RESEARCH.

An exploratory study, by definition, must lead to a number of suggestions concerning further research. The research areas which appear fruitful for investigation as a continuation of questions which emerged from the experiment reported here, are grouped under four headings.

1. Cognitive Style.

There appears to be a need for longitudinal studies of cognitive styles, not necessarily confined to the one construct, and possibly related to both Piaget's theory of intellectual development (particularly the concrete-abstract continuum) and to a personality theory similar to that of Harvey, Hunt and Schroder, one which related perception, cognitive cost and self concept. Such a study would be concerned with possible situational determinants, and especially with the influence of teachers on preferred conceptual modes. Are there marked changes in conceptual style at certain periods in a child's life? To what extent are they related to change of teacher or other factors? If teachers are influential in effecting such changes, are the changes related to the teacher's own personality or own cognitive style? Or are they, alternatively, some function of teaching style? These are among research problems which it would seem profitable to explore. A further aspect of the cognitive style construct studied in the present investigation was a slight clue that the relative balance between the three modes might be as important as the weighting attached to the dominant one. This, together with a study of which stimulus patterns tend to evoke which type of cognitive response style, is another potential study. A specific question for further investigation is the relationship between response set and other cognitive and personality characteristics.

2. Individual Differences.

The variance attributable to pupils within the interacting matrix of pupils, teacher, methods, styles and sex was of considerable strength in the major analysis of variance. Neither Otis I.Q. nor cognitive style accounted for other than a portion of this variance. Sex, however, was a factor interacting with teacher and objectives. It was suggested that boys might transfer and apply learning more readily than girls. Such a proposition requires further testing. When I.Q. was covaried, sex appeared as a factor having significant main effects. Further research on the relationship between sex and the achievement of objectives, between sex and learning variables (other than science content, for example) and between sex and school learning with I.Q. controlled, would appear to be valuable enterprises. It is hardly necessary to recommend further investigation to identify individual differences beyond intelligence and cognitive styles. Cronbach (already cited) has suggested that the potent factors may be non-cognitive. The contention is not denied, but the task of identifying the important variables remains. Until this is accomplished, attempts to develop a comprehensive theory of instruction may be limited.

3. Teaching Strategies.

Insofar as different objectives are achieved to differing extents by different teaching procedures, further investigation on the most appropriate combinations of expository and inductive instruction to achieve multiple cognitive

objectives should be initiated. Further studies in which amount of guidance is varied, or in which procedures are compared, might advisedly be over longer periods of instruction time than the one-hundred-minute-span of this experiment. A research programme dealing with variables of task-presentation (e.g. cues, set, sequence, practice, form of presentation) with task characteristics and with teacher functions could be developed. It is also suggested that research on instruction could be conducted at two levels of analysis simultaneously, the gestalt and the analytic. It was not possible in the present investigation to identify which were the major within-method variables. An additional recommendation is that attempts be made to measure multiple outcomes, cognitive and affective, for a range of teaching procedures, holding content and pupil characteristics constant, or varying certain pupil characteristics systematically. Such a research programme would be exploratory, and rather open-ended, but would, it is believed, contribute to the development of a theory of teaching.

4. Instruments.

A difficulty in measuring multiple outcomes is in locating and developing instruments which are sufficiently sensitive to discriminate subtle changes with greater precision than was achieved in this study. A need for means to measure attitude changes, self-concept modifications and affective outcomes resulting from school learning, is apparent. The task of constructing, testing and refining appropriate scales and instruments is a major one.

APPENDIX A: (INSTRUMENTS)

- A.1 Administration: Triad One. (With tests and answer sheet)..
- A.2 Administration: Triad Two. (With tests and answer sheet).
- A.3 Administration: Science Test - Force, Work and Energy. (With test).
- A.4 Administration: Post-tests. (With the four tests).
- A.5 Criteria for Science Test IIB - Scoring Guide. (With Control Group tests).
- A.6 Administration: Pupil Opinion Survey. (With test).
- A.7 Example of Test Objectives.
- A.8 Item Analysis Data: Cognitive Style measures.
- A.9 Item Analysis Data: Science Content measures.
- A.10 Item Analysis Data: Pupil Opinion Survey.

APPENDIX A.1: ADMINISTRATION - TRIAD ONEPREPARATION.

After ensuring that all pupils had pencils and rubbers ready, the booklets and answer sheets were issued. Children were asked not to open the booklets until instructed to do so, but were instructed to enter name, date etc. on answer sheets.

INTRODUCTORY STATEMENT TO CHILDREN.

"Today I have an exercise for you to do, one which I think you will enjoy very much. It is an exercise, and NOT a test. Have you noticed that when people are grouping things, each one has his own way of doing it? For example, if you were putting some marbles into groups, you might do it in many ways. Each way is right for the person doing it. Today, I want you to do some grouping for me. Look at the cover of the booklet. I shall read it aloud, while you follow what it says."

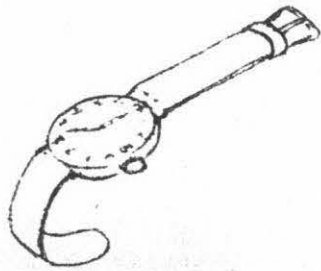
INSTRUCTIONS

When you are told to start you will open the booklet and find some pictures arranged in groups of three. Each group of three is numbered, and each picture has a letter beneath it. The first page is arranged like this:

	Picture	Picture	Picture
1.	a.	b.	c.
	Picture	Picture	Picture
2.	d.	e.	f.

You are to pick from each group of three, two pictures that are alike or go together in some way. On the answer sheet put an X on two letters, to show the pictures you have chosen. Then, after the word "because", write your reason for picking those two pictures. Usually the reason can be stated in only a few words, and you do not need to write a complete sentence, or to worry about your spelling. There are no "right" or "wrong" answers. You may work at your own speed. If you finish before the others, close your booklet, sit quietly and wait for further instructions. Please do not mark your booklet.

Please wait for the signal to begin.



1.

a.



b.



c.

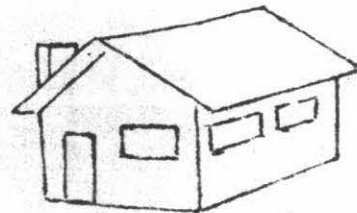


2.

d.



e.



f.

3.



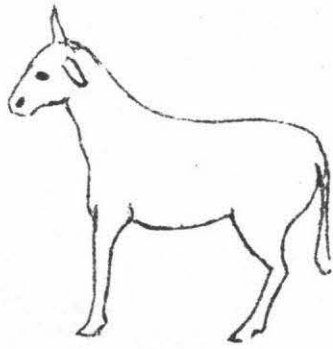
a.



b.



c.

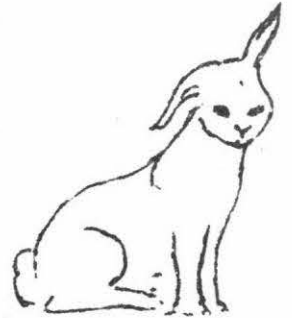


4.

d.



e.



f.



5.

a.



b.

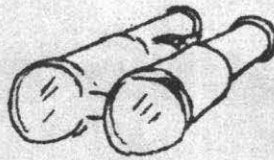


c.

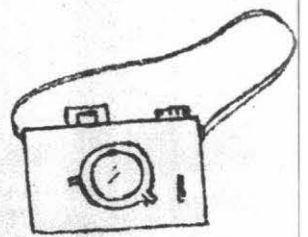


6.

d.



e.



f.



7.

a.



b.

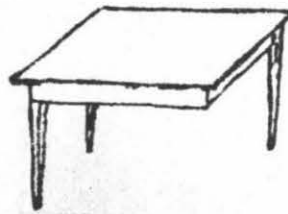


c.



8.

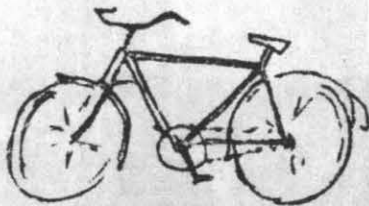
d.



e.

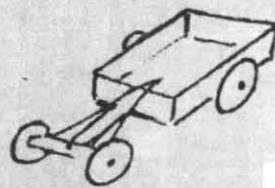


f.



9.

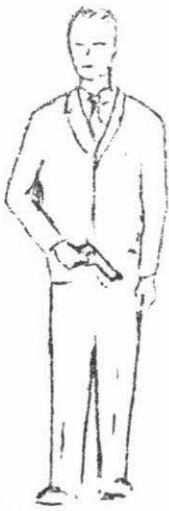
a.



b.



c.

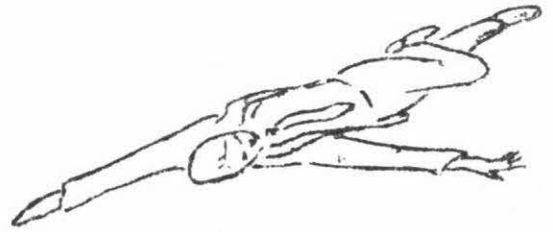


10.

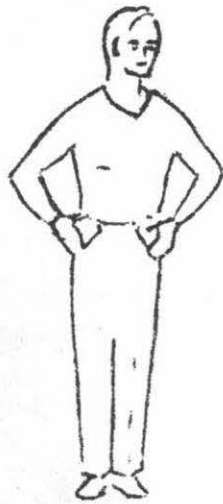
d.



e.



f.



11.

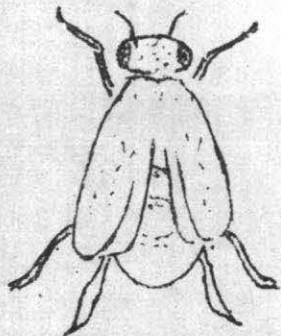
d.



b.

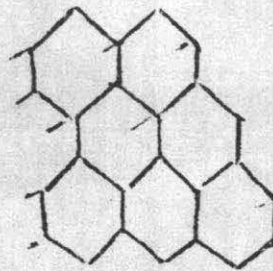


c.



12.

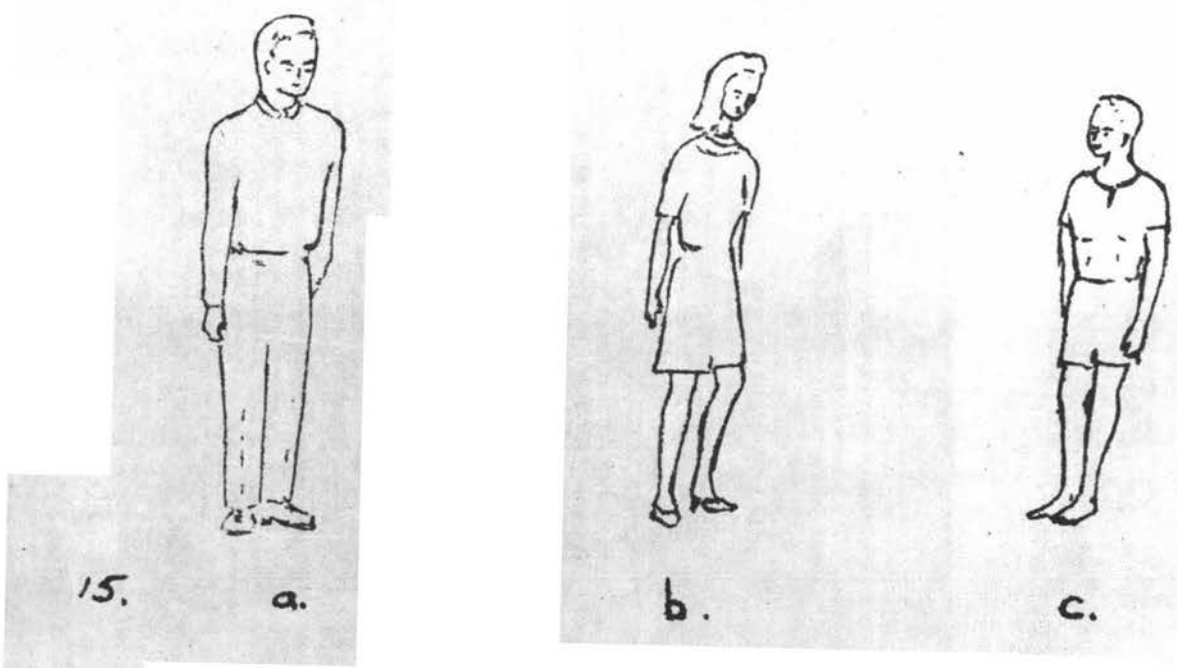
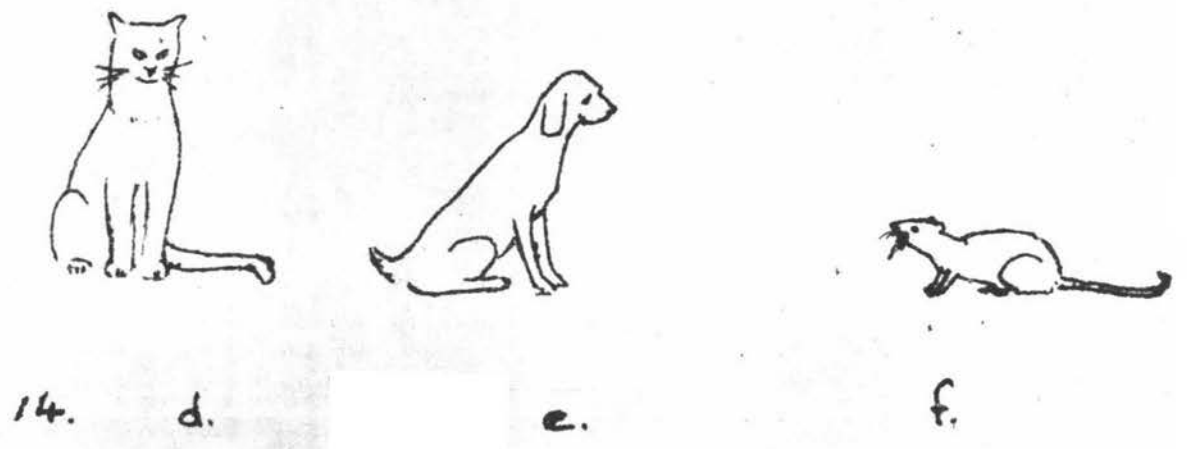
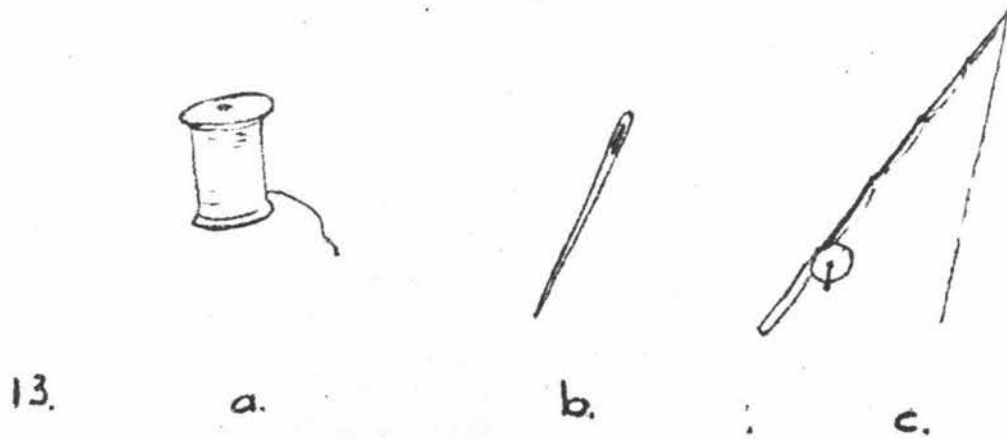
d.



e.



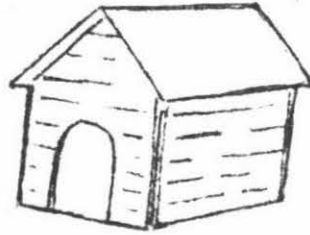
f.





16.

d.



e.



f.



17.

a.



b.



c.



18.

d.



e.



f.



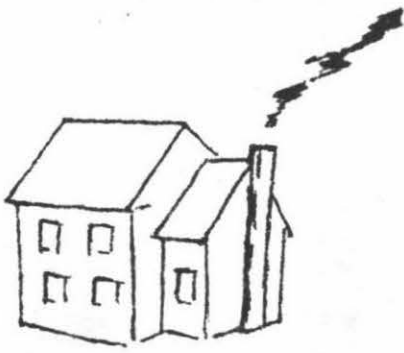
19. a.



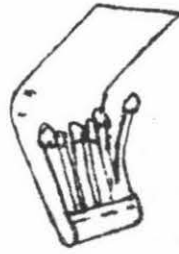
b.



c.



20. d.



e.



f.



21

a.



b.



c.



22.

d



e.



f.



23.

a.



b.



c.



24.

d.



e

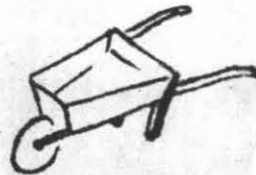


f.



25.

a.



b



c.

ANSWER SHEET

NAME: _____ BOY or GIRL: _____

AGE: _____ years _____ months DATE: _____

Remember: Put an X on the two letters from each group of three to show the pictures that are alike or go together in some way. Then, after the word "because", write your reason for picking those two pictures. You do not need to write complete sentences, or to worry about your spelling. There are no "right" or "wrong" answers.

Number:	<u>Group:</u>	<u>Reason:</u>
1.	a. b. c.	because _____ _____
2.	d. e. f.	because _____ _____
3.	a. b. c.	because _____ _____
4.	d. e. f.	because _____ _____
5.	a. b. c.	because _____ _____
6.	d. e. f.	because _____ _____
7.	a. b. c.	because _____ _____
8.	d. e. f.	because _____ _____

APPENDIX A.2: ADMINISTRATION - TRIAD TWOINTRODUCTORY STATEMENT TO CHILDREN.

"Last time you were kind enough to help with an exercise which had a number of pictures arranged in groups of three. You will remember that you chose pairs of pictures that were alike or would go together in some way. You made some X's on the answer sheet to show your pair out of a group of three, and then you wrote your reason for choosing those two. I think you will remember that. Well, today there are some more exercises which are really very interesting ones - I think you will enjoy doing them. But before we go on to some of the new ones, I should like you to help me with picture grouping. After last time, it was decided to alter the exercise a little. Today I should like you to find three ways of making pairs out of each group of three pictures."

(Demonstration on blackboard).

"On the answer sheet you will find that group number one has room for three lots of two that are alike or will go together in some way.

e.g. 1. a. b. c. because

.....

a. b. c. because

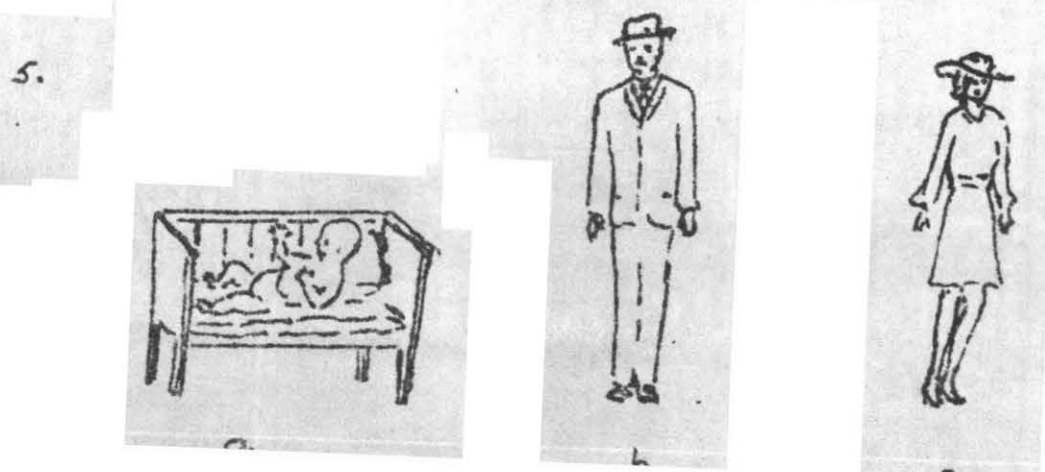
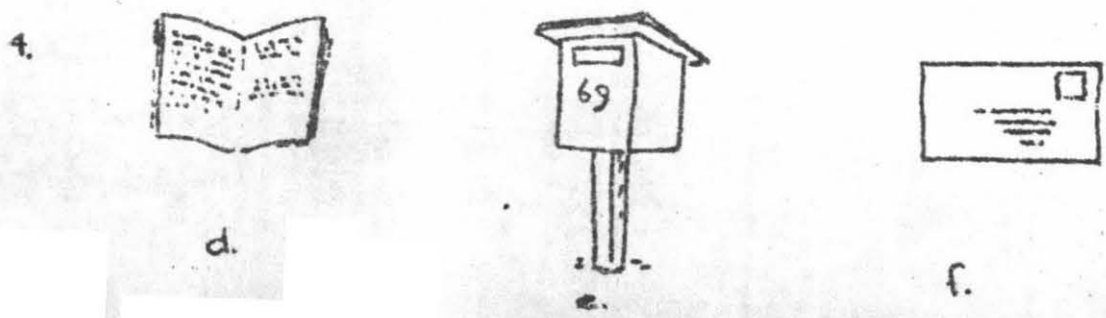
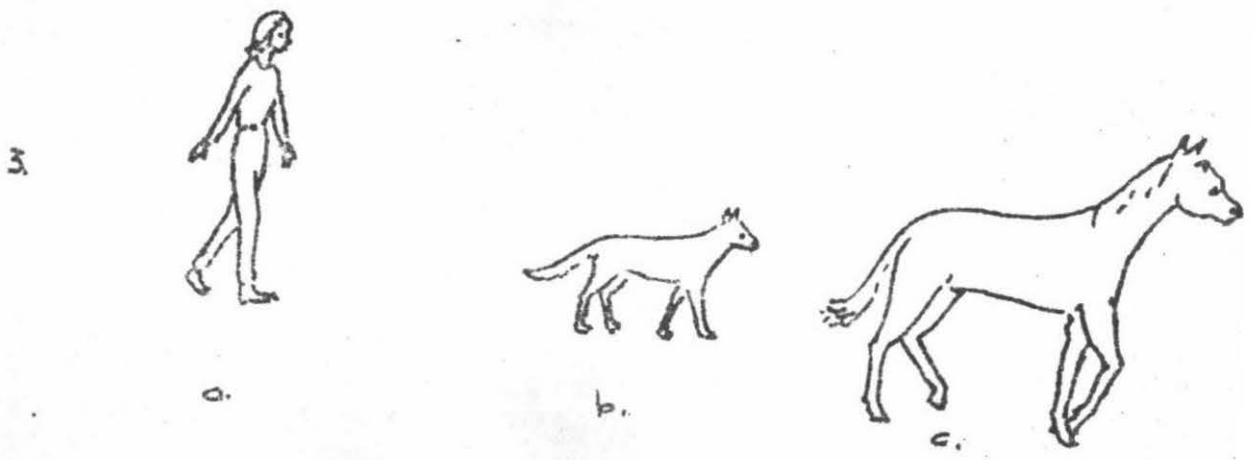
.....

a. b. c. because

.....

You might put a and b together, or a and c, or b and c. However, you might put a and b together twice, but for different reasons. Are there any questions? For a group of three pictures, you are to choose two pictures that are alike or go together in some way. Put a X on your answer sheet over the two letters that go with the two pictures, and write your reason very briefly. You are to try to find three ways of making pairs for each group of pictures. We are interested in your reasons, so you could have the same letters more than once, but for different reasons.

Are you ready? Very well, start now."



ANSWER SHEET

NAME: _____ BOY or GIRL: _____

AGE: _____ years _____ months DATE: _____

Look at each group of three pictures, and try to find three ways of making pairs that are alike or go together in some way. For each pair, put an X on the two letters to show which you have chosen. Then, after the word "because", write your reason for picking those two pictures. You do not need to write complete sentences, or to worry about your spelling. There are no "right" or "wrong" answers.

<u>Number:</u>	<u>Group:</u>	<u>Reason:</u>
1.	a. b. c.	because _____ _____
	a. b. c.	because _____ _____
	a. b. c.	because _____ _____
2.	d. e. f.	because _____ _____
	d. e. f.	because _____ _____
	d. e. f.	because _____ _____

APPENDIX A.3: ADMINISTRATION - SCIENCE TEST:
FORCE, WORK AND ENERGY

This test was administered by class teachers as part of the science programme in the experimental school. The embedded items were numbers 4, 7, 9, 10 and 11. The remainder were designed to test performance on the current science unit. Items 12, 13, 14 and 15 were modifications of items used by Suchman. (See Hedges, W.D. (1966). Testing and Evaluation for the Sciences. California: Wadsworth.)

GENERAL INSTRUCTIONS TO TEACHERS.

Please ensure that all pupils have pencils and rubbers, and ask them to fill in date and name. Observe the usual test conditions. Rapport will be easily established, but as part of the approach, suggest to the children that we are interested to see how well they can handle science questions, some of which might be new to them.

Read out the instructions, trying to ensure that all children understand what is required. Most pupils will finish in fifteen to twenty minutes, but allow time for all of them to finish. If any pupil should have difficulty in reading any question, please assist by reading the question quietly to him. Thank you.

SCIENCE TEST - FORCE, WORK AND ENERGY (A)

Date: _____

Name: _____

Unless the question has special instructions (for example Question Two) you should record your answer by printing, in the space at the right hand side, the letter (A, B, C or D) showing the answer you have chosen to be correct.

1. A push, a pull, and gravity are examples of a particular kind of:
- force.
 - friction.
 - inertia.
 - work.

1. _____

2. Below are four pictures, i, ii, iii and iv. Under the pictures are some phrases to describe what is happening. The phrases have letters beside them. Choose the correct phrases to go with each picture, and put their letters in the right answer space. Each answer space will have at least two letters in it.



i



ii.



iii.



iv.

2. i. _____

ii. _____

iii. _____

iv. _____

- | | |
|-----------------------------|---------------------------|
| A. Work done. | B. No work done. |
| C. Potential energy gained. | D. Kinetic energy gained. |
| E. Gravity acting. | F. Friction acting. |

3. If I had four blocks, each of the same weight, which one would be hardest to push along a wooden floor?
- A glass block.
 - A wooden block.
 - A concrete block.
 - An ice block.

3. _____

4.

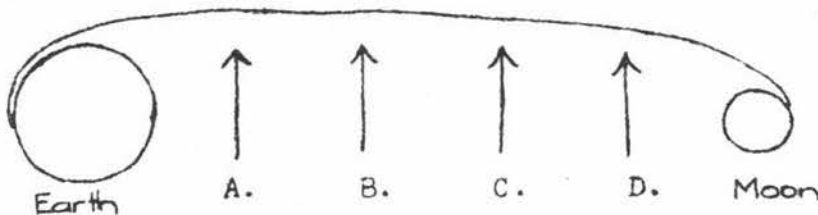


Which is the BEST reason to explain why this racing car is hard to roll over?

- A. The width between the wheels keeps the weight low.
- B. The difference in size between back and front wheels gives more grip.
- C. The large rear tyres increase the amount of friction.
- D. The length of axle gives a wide wheel base.

4. _____

5. Which arrow shows where the moon's gravity would cause a space craft to speed up again in a flight to the moon?



5. _____

6. The pull of gravity makes it harder to lift a heavy weight:

- A. at the equator.
- B. at the south pole.
- C. halfway between equator and pole.
- D. in space above the earth.

6. _____

7. The astronauts on the moon leaned forward in the direction they wished to walk. They did this because:

- A. the solar wind pushed so strongly against them.
- B. they were protecting their eyes from the glare of the strong sunlight.
- C. they had the life-support packs on their backs.
- D. the moon's surface was so soft, it took more energy to walk on it.

7. _____

8. Which statement is NOT correct?

- A. Inertia is the tendency to resist movement.
- B. Inertia acts on all stationary objects.
- C. Inertia is the tendency to continue moving in a set direction.
- D. Inertia is absent on the moon.

8. _____

9. An object like a pyramid is hard to tip over. Which is the BEST reason to explain this?

- A. The object has angled sides.
- B. The object has large flat surfaces.
- C. The object has a low weight centre.
- D. The object has a large mass.

9. _____

10. The Leaning Tower of Pisa, sketched below, does not fall over because:



- A. the earth's rotation acts against its falling.
- B. it has no strong force acting on it.
- C. it has greater weight in the bottom half than in the top half.
- D. the line through which gravity acts falls inside the base.

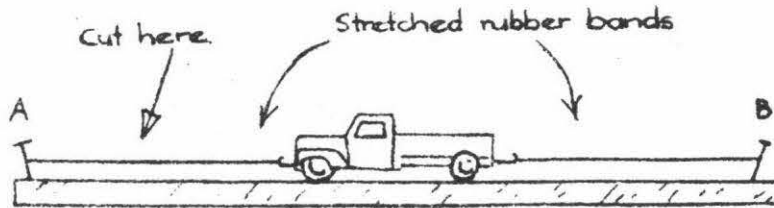
10. _____

11. When a force begins to tilt a stable object, the object tends to:

- A. move away from its original position.
- B. return to its original position.
- C. take up a new position.
- D. move position in line with the force being used.

11. _____

12. Look at the picture and then answer the question.



If the rubber band is cut at the place marked, the toy truck would:

- A. move towards A.
- B. move towards B.
- C. not move.
- D. fly straight up.

12. _____

13. How would you make the truck move in the opposite direction?

- A. Change the rubber bands before starting the experiment.
- B. Drive the nail at A further in.
- C. Take the wheels off the truck.
- D. Cut the other rubber band.

13. _____

14. Which rule explains why the truck moves?

- A. Whenever a force acts on an object, that object must move.
- B. Whenever two forces act on an object, that object does not move.
- C. When an object is moving, the forces acting on it are in balance.
- D. When an object is set in motion, the forces acting on it are not in balance.

14. _____

15. Which form of energy allowed this experiment to work?

- A. Potential energy.
- B. Kinetic energy.
- C. Molecular energy.
- D. Chemical energy.

15. _____

APPENDIX A.4: ADMINISTRATION - POST-TESTSORDER OF ADMINISTRATION.

- Test 1. All children (Science Test - Part IA).
 Test 2. For groups 1, 2, 3 and 4 (who were with Miss - and Mr. -), Science Test - Part IIA.
 For groups 5 and 6, Test - Simple Machines.
 Test 3. All children (Pupil Opinion Survey).

TIMING. It was expected that all children would complete each test. Times anticipated for completion were:

- Test 1. 20 minutes approximately.
 Test 2. 20 minutes approximately.
 Test 3. 15 to 20 minutes.

NON-READERS. If any child had difficulty reading any item, it was read to him without disturbing the other children.

INTRODUCTORY STATEMENT TO CHILDREN.

"On Tuesday and Thursday you had some extra science lessons with four teachers, Miss - and Mr. - in the Art and Science Rooms, and Miss - and Miss - in Rooms 9 and 23. The lessons were all new ones for Form I children. We are really very interested to know what you found out in those lessons, and how well you understood them. The best way for us to find out is to give you some questions to answer. I think you will find the questions very interesting ones. We would like you to answer them as well as you can. There are three lots of questions for you to do."

(Test 1 issued).

"Please put on the top of the paper the name of the teacher you had on Tuesday and Thursday. Then put the date, and your full name, i.e. first and second names, e.g. John James."

(Instructions read through to children to make sure they knew what to do).

"If you have any questions while you are working, just raise your hand and someone will come to help you. Remember to do your best. If you cannot answer a question, go on to the next one and come back to it later. If you can, try to answer all the questions."

Test 2. Procedure as for Test 1.

Test 3 (Pupil Opinion Survey). Procedure as for Test 1.

STATEMENT TO CHILDREN. "You will have met this exercise before. We would like you to do it again for us today. This will help us to know whether we have asked the right questions."

ADMINISTRATION for the second testing occasion, fourteen days later, followed a similar pattern, with only two tests (Science Tests IB and IIB) being given.

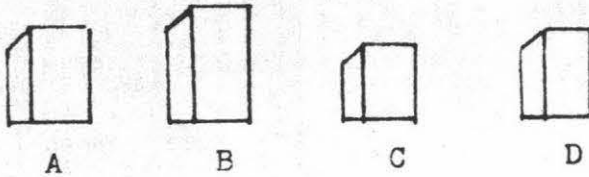
SCIENCE TEST - PART I (A)

Date: _____

Name: _____

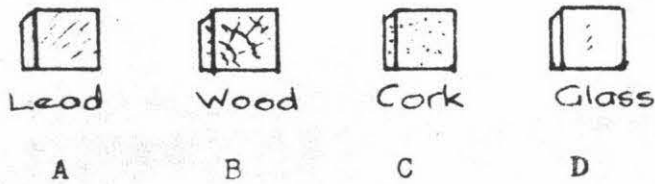
Choose the letter A, B, C or D which shows the answer you decide is correct. Put that letter in the space at the right hand side of the page, beside the number of the question you are answering. Do not spend a long time over any one question. Go on to the end and come back to that question later.

1. Which of the four blocks, each weighing five pounds and having the same base area, is the least stable?



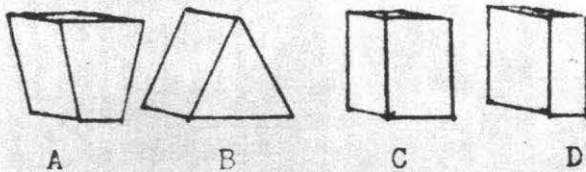
1. _____

2. Which of the four blocks, all of equal size, is the most stable?



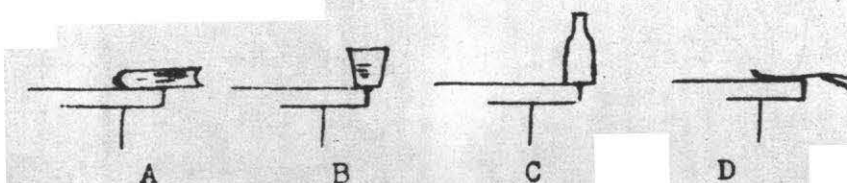
2. _____

3. Look at the scale drawings of the four objects. Each is eight inches high and weighs two pounds. Which is the most stable?



3. _____

4. Look at the pictures below. Which of the book, the glass, the bottle or the spoon is most likely to fall off the table?



4. _____

5. The point within an object through which the pulling force of the earth works is called:
- the centre of mass.
 - the centre of force.
 - the centre of volume.
 - the centre of gravity.

5. _____

6. Which of the following objects is most stable?
- A building brick.
 - An empty petrol tin.
 - A hardboiled egg.
 - An empty tea-cup.

6. _____

7. Which is the BEST reason to explain why this racing car is hard to roll over?



- The width between the wheels keeps the weight low.
- The difference in size between back and front wheels gives more grip.
- The large rear tyres increase the amount of friction.
- The length of axle gives a wide wheel base.

7. _____

8. Which piece of apparatus would give you the most accurate position of the centre of gravity in an odd-shaped piece of cardboard?
- A foot ruler.
 - A weighted string.
 - A pair of compasses.
 - A sharp pin.

8. _____

9. Which statement describes an object which is not very stable?
- The object has a great weight and little height.
 - The object has a large base and little weight.
 - The object has a great height and a large base.
 - The object has a small base and great height.

9. _____

10. An object like a pyramid is hard to tip over. Which is the BEST reason to explain this?
- The object has angled sides.
 - The object has large flat surfaces.
 - The object has a low weight centre.
 - The object has a large mass.

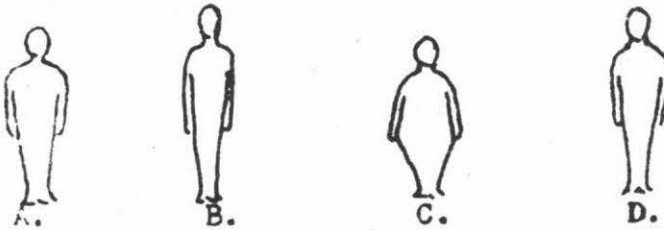
10. _____

11. Irregular objects will balance when:

- A. the weight is evenly spaced around the centre.
- B. the centre of gravity is directly over the base of support.
- C. the point of support is directly above the centre point.
- D. the base is wide enough to hold the weight.

11. _____

12. Who has the lowest centre of gravity?



12. _____

13. The centre of gravity of any object is:

- A. the point in the exact centre of the object.
- B. the point on which the object will turn.
- C. the point through which gravity takes its effect.
- D. the point where all forces act equally on the object.

13. _____

14. Which mark would be nearest the centre of gravity of this piece of cardboard?



14. _____

15. When a force begins to tilt a stable object, the object tends to:

- A. take up a new position.
- B. move away from its original position.
- C. return to its original position.
- D. move position in line with the force being used.

15. _____

16. The astronauts on the moon leaned forward in the direction they wished to walk. They did this because:
- they had the life support packs on their backs.
 - the moon's surface was so soft, it took more energy to walk on it.
 - they were protecting their eyes from the glare of the strong sunlight.
 - the solar wind pushed so strongly against them.

16. _____

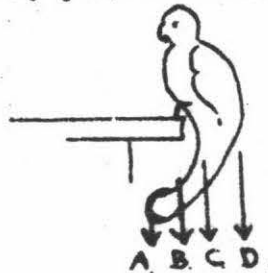
17. The Leaning Tower of Pisa, sketched below, does not fall over because:



- the earth's rotation acts against its falling.
- it has no strong force acting on it.
- the line through which gravity acts falls inside the base.
- it has greater weight in the bottom half than in the top half.

17. _____

18. Which arrow shows the line on which gravity is acting on the toy parrot balanced on the table edge?



18. _____

19. Weight may be described as:
- the mass of an object.
 - the overall density of an object.
 - the effect of gravity on an object.
 - the total volume of an object.

19. _____

20. To increase the stability of an object you would increase:

- the height of the object.
- the weight of the object.
- the volume of the object.
- the length of the object.

20. _____

SCIENCE TEST - PART I (B)

Date: _____

Name: _____

Choose the letter A, B, C or D which shows the answer you decide is correct. Put that letter in the space at the right hand side of the page, beside the number of the question you are answering. Do not spend a long time over any one question. Go on to the end and come back to that question later.

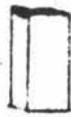
1. Which of the four blocks, each weighing four pounds and having the same base area, is the least stable?



A



B



C



D

1. _____

2. Which of the four blocks, all of equal size, is the most stable?



Wood.

A



Glass

B



Cork

C

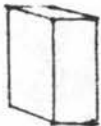


Lead.

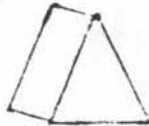
D

2. _____

3. Look at the scale drawings of the four objects. Each is ten inches high and weighs three pounds. Which is the most stable?



A



B



C



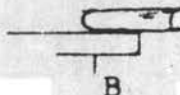
D

3. _____

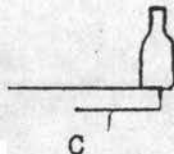
4. Look at the pictures below. Which of the book, the glass, the bottle or the spoon is most likely to fall off the table?



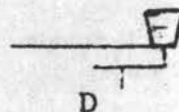
A



B



C



D

4. _____

11. Irregular objects will balance when:

- A. the centre of gravity is directly over the base of support.
- B. the weight is evenly spaced around the centre.
- C. the point of support is directly above the centre point.
- D. the base is wide enough to hold the weight. 11. _____

12. Who has the lowest centre of gravity?

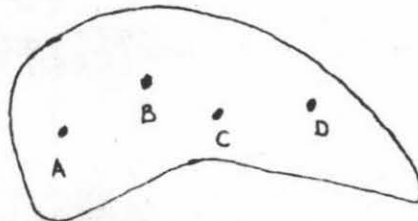


12. _____

13. The centre of gravity of any object is:

- A. the point where all forces act equally on the object.
- B. the point through which gravity takes its effect.
- C. the point in the exact centre of the object.
- D. the point on which the object will turn. 13. _____

14. Which mark would be nearest the centre of gravity of this piece of cardboard?



14. _____

15. When a force begins to tilt a stable object, the object tends to:

- A. return to its original position.
- B. move away from its original position.
- C. move position in line with the force being used.
- D. take up a new position. 15. _____

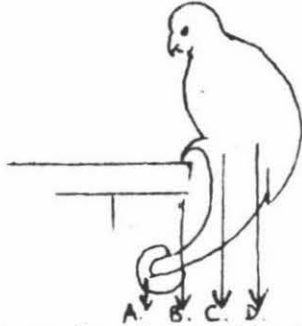
16. The astronauts on the moon leaned forward in the direction they wished to walk. They did this because:

- A. they had life support packs on their backs.
- B. the moon's surface was so soft, it took more energy to walk on it.
- C. they were protecting their eyes from the glare of the strong sunlight.
- D. the solar wind pushed so strongly against them. 16. _____

17. The Leaning Tower of Pisa, sketched below, does not fall over because:



- A. the line through which gravity acts falls inside the base.
 B. the earth's rotation acts against its falling.
 C. it has greater weight in the bottom half than in the top half.
 D. it has no strong force acting on it. 17. _____
18. Which arrow shows the line on which gravity is acting on the toy parrot balanced on the table edge?



19. Weight may be described as:
 A. the total volume of an object.
 B. the effect of gravity on an object.
 C. the overall density of an object.
 D. the mass of an object. 19. _____
20. To increase the stability of an object you would increase:
 A. the length of the object.
 B. the weight of the object.
 C. the height of the object.
 D. the volume of the object. 20. _____

5. The point within an object through which the pulling force of the earth works is called:
- A. the centre of mass.
 - B. the centre of gravity.
 - C. the centre of volume.
 - D. the centre of force.

5. _____

6. Which of the following objects is most stable?
- A. An empty petrol tin.
 - B. A hardboiled egg.
 - C. An empty tea-cup.
 - D. A building brick.

6. _____

7. Which is the BEST reason to explain why this racing car is hard to roll over?



- A. The width between the wheels keeps the weight low.
- B. The large rear tyres increase the amount of friction.
- C. The difference in size between back and front wheels gives more grip.
- D. The length of axle gives a wide wheel base.

7. _____

8. Which piece of apparatus would give you the most accurate position of the centre of gravity in an odd-shaped piece of cardboard?
- A. A weighted string.
 - B. A pair of compasses.
 - C. A foot ruler.
 - D. A sharp pin.

8. _____

9. Which statement describes an object which is not very stable?
- A. The object has a great height and a large base.
 - B. The object has a small base and great height.
 - C. The object has a great weight and little height.
 - D. The object has a large base and little weight.

9. _____

10. An object like a pyramid is hard to tip over. Which is the BEST reason to explain this?
- A. The object has a large mass.
 - B. The object has angled sides.
 - C. The object has large flat surfaces.
 - D. The object has a low weight centre.

10. _____

SCIENCE TEST - PART II (A)

Date: _____

Name: _____

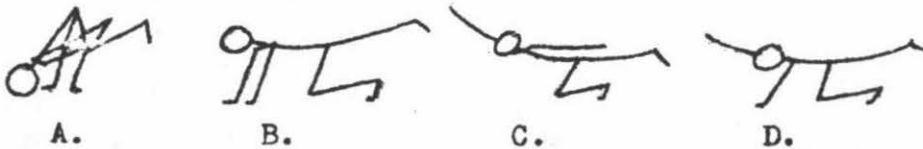
1. A runner needs to alter his centre of balance when he sprints around a sharp curve. He does this by:
- leaning out on the bend.
 - leaning into the bend.
 - keeping straight up and down.
 - keeping his weight forward.

1. _____

2. When a gymnast is walking along a narrow beam she can improve her stability most easily by:
- spreading her arms.
 - bending her knees.
 - pointing her toes out.
 - keeping her head forward.

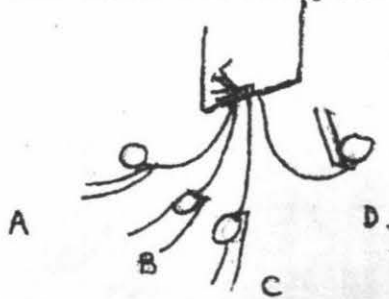
2. _____

3. Which is the least stable position - one which would be difficult to hold for long?



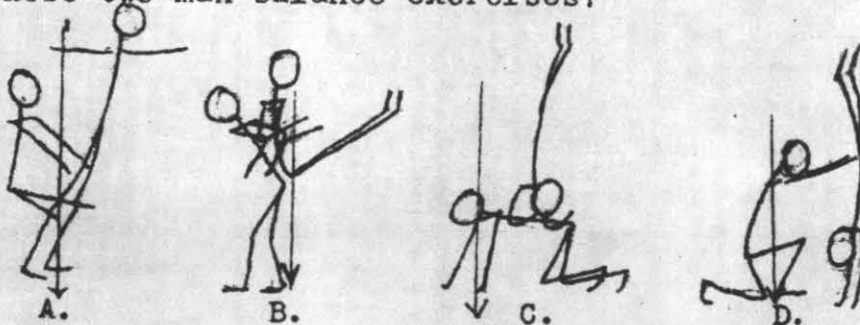
3. _____

4. Which is the most stable position?



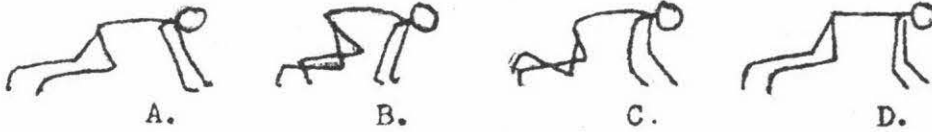
4. _____

5. Which sketch shows the correct line of gravity in these two-man balance exercises?



5. _____

6. Which runner has set his centre of balance to get the best start?



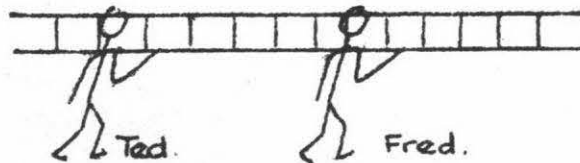
6. _____

7. My model yacht capsizes too easily. How can I BEST make it more stable?

- A. Reduce the amount of sail.
 B. Reduce its overall length.
 C. Increase the weight of the keel.
 D. Increase the width of the hull.

7. _____

8. Two boys, Ted and Fred, carry a ladder on their shoulders as shown in the sketch below.

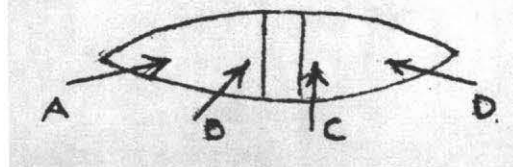


Who carries the heavier weight?

- A. Ted, because he is further from the balance point.
 B. Fred, because Fred is nearer the turning point.
 C. Fred, because he is nearer the centre of gravity.
 D. Neither, each carries an equal share of the weight.

8. _____

9. A canoe is not very stable. If the centre of gravity is positioned near the seat, which arrow shows the safest way to step from the shore into the canoe?



9. _____

10. I have some tall, heavy flowers to put in one of four clay vases. Which vase will I use if I wish to keep my flower arrangement stable?



A.



B.



C.



D.

10. _____

11. Which of these pictures shows a correct balance of a 3 pound weight and a 4 pound weight?



A.



B.



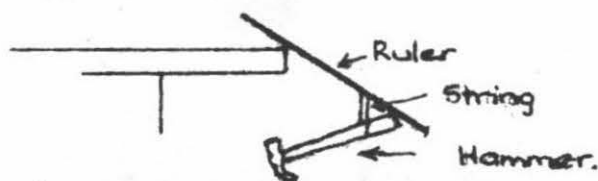
C.



D.

11. _____

12. Look at the picture of the group of objects balanced on the table edge.

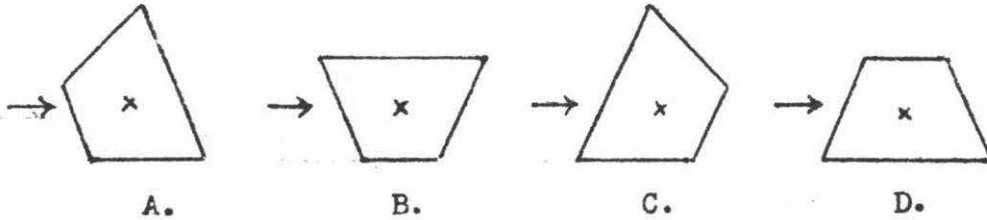


Which is the BEST reason to explain why these objects balance as they do?

- The centre of gravity 's in the hammer near the string.
- The centre of gravity is in line with the edge of the table.
- The low centre of gravity increases friction on the ruler.
- The objects are arranged to have a low centre of gravity.

12. _____

13. Below are drawings of the same piece of wood in four different positions. If a tipping force is pushing each one in the direction shown by the arrows, which one is in the least stable position? Note: The centre of gravity is marked.

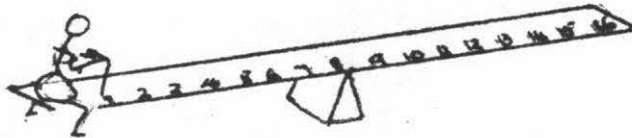


13. _____

14. Which is the BEST reason to explain how a picture hangs perfectly straight.
- The hook is in the exact middle of the length of string.
 - There is an equal area of picture each side of the centre line.
 - The support point is right over the centre of gravity.
 - The weight of the picture is lower than the point of support.

14. _____

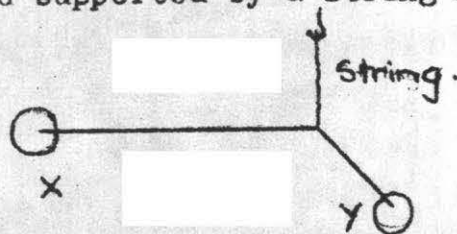
15. A girl weighing 50 lbs. is sitting on the end of the see-saw. Where would a boy weighing 100 lbs. have to sit so that the see-saw balances?



- A. At 16? B. At 14? C. At 12? D. At 10?

15. _____

16. The picture below shows two objects joined by a light wire and supported by a string from a hook.



Why do the weights, X and Y, stay in this position?

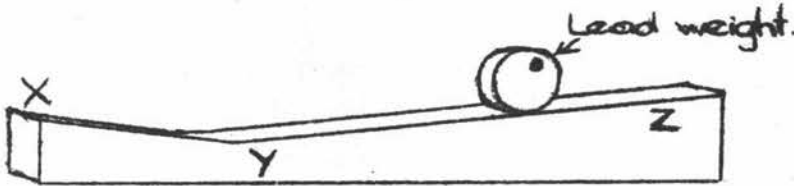
- X weighs the same as Y.
- X is heavier than Y.
- Y is heavier than X.
- Y is below X.

16. _____

17. Which is the most sensible reason for adding a longer tail to a kite?

The tail will make the kite more stable by:

- A. increasing the weight of the kite.
 - B. increasing the mass of the kite.
 - C. reducing air resistance on the kite.
 - D. lowering the centre of gravity of the kite. 17. _____
18. Look at the picture of the solid wooden wheel. Notice where a piece of lead, half the weight of the wheel, has been fitted into it. If the wheel can move in the gently sloping path, which way will it move?



- A. The wheel will first roll downhill to Y.
 - B. The wheel will first roll uphill to Z.
 - C. The wheel will immediately roll all the way to X.
 - D. The wheel will remain still. 18. _____
19. Which is the most difficult to balance on its end?
- A. A new pencil.
 - B. A drinking straw.
 - C. A wooden clothes peg.
 - D. A cigarette. 19. _____
20. Which would be safest to use when travelling downhill along a winding road?
- A. A tricycle.
 - B. A bicycle.
 - C. A go-kart.
 - D. A pram. 20. _____

SCIENCE TEST - PART II (B)

Date: _____

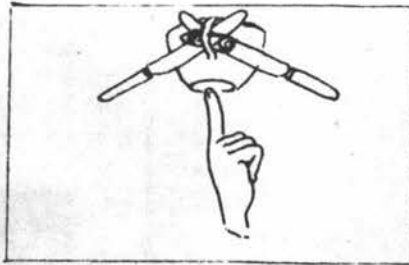
Name: _____

1. What is meant by "top-heavy"?

2. Why is it easier to ride a three-wheeled cycle than a two-wheeled cycle?

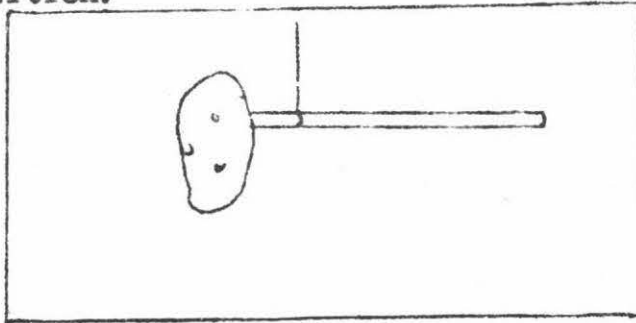
3. Why should you never stand upright when changing places in a rowing boat?

4. Look at the drawing:

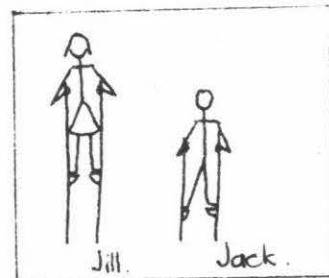


Why will the cup, with knives crossed and held in place in the handle by a roll of paper, balance on a finger-tip?

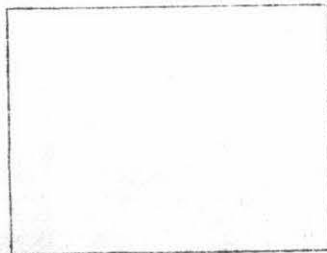
5. A potato with a meat skewer in it is suspended from a string. The potato and skewer alone will not hang in the way shown in the picture. If I have two dinner forks, I can make the potato and skewer hang so that the skewer is level. Draw the two forks in on the picture, so that the objects hang in a balanced position.



6. Two children, Jack and Jill, are on stilts, as shown in the picture. Why is Jill easier to push over than Jack is?



7. You have been asked to design a cup which will be very stable and hard to knock over. Draw your cup in the box below.



8. Some objects are very stable indeed. What three things will be true of such objects?

(a) _____

(b) _____

(c) _____

APPENDIX A.5: CRITERIA FOR SCIENCE TEST - PART II: B

Item	Guide to Scoring	Score Given	Maximum
1.	Simple statement, limited to e.g. "Is heavy" Descriptive statement, e.g. "Heavy in top half" "Unstable" Description with reason, e.g. "Has a high centre of gravity and is unstable"	0 1 2	2
2.	Simple statement, e.g. "You don't need to balance it" Descriptive, implies reason, e.g. "Has a wide base" Statement with reason, e.g. "The wide base makes the weight centre lower"	1 2 3	3
3.	Simple statement of consequence, e.g. "Will tip over" Description, reason implied, e.g. "It will become less stable" Reason given, e.g. "It will alter the centre of gravity, and be less stable"	1 2 3	3
4.	Simple statement, e.g. "It is balanced" Description, reason implied, e.g. "The weight is increased evenly" Reason in principles given, e.g. "The centre of gravity is lowered" "Balanced on line of support"	1 2 3	3
5.	Give <u>zero</u> if no possibility of balance Give <u>one</u> if a balance appears possible, but not clearly accurate enough to be sure Give <u>two</u> if a clear fulcrum pattern is shown, or if counterbalanced with forks in potato	0 1 2	2
6.	Give <u>one</u> if descriptive only, e.g. "Jill is taller" Give <u>two</u> if reason is supplied, e.g. "Jack has a lower weight centre"	1 2	2
7.	Give <u>one</u> if base increased but cup would be difficult to drink from Give <u>two</u> if base wide, or sides straight, or additional weight added, and still easy to drink from	1 2	2
8.	Give <u>one</u> point for each clear principle, with maximum of <u>three</u> , e.g. heavy weight, large base area, low centre of gravity, line of gravity within base zone, low and squat.		3

SCIENCE TEST - SIMPLE MACHINES

Date: _____

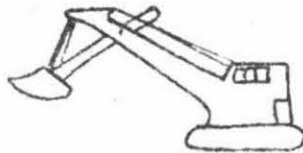
Name: _____

Unless the question has special instructions (for example Question Four) you should record your answer by printing, in the space at the right hand side, the letter (A, B, C or D) showing the answer you have chosen to be correct.

1. A screw is an example of which of the following kinds of simple machine?
- The lever.
 - The wheel and axle.
 - The inclined plane.
 - The wedge.

1. _____

2.



This earth digger is using the principle of:

- the lever and pulley.
- the wedge and wheel.
- the wheel and axle.
- the pulley and inclined plane.

2. _____

3. Which pattern is found in a first class lever?
- Fulcrum, load, effort.
 - Load, effort, fulcrum.
 - Effort, fulcrum, load.
 - None of these.

3. _____

4. In the answer space beside each of the following, put the number which tells what kind of a lever the object is:

- = first class lever
- = second class lever
- = third class lever

4.

- a wheelbarrow
- sugar tongs
- a spade
- scissors
- nutcrackers
- a see-saw
- tweezers

i. _____

ii. _____

iii. _____

iv. _____

v. _____

vi. _____

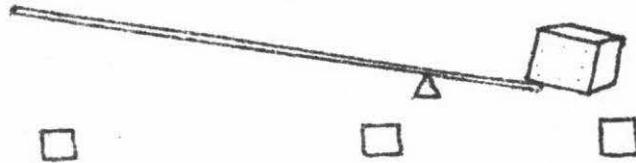
vii. _____

5. Below is a list of simple machines. Alongside is a list of everyday objects. In the answer space at the right, put the correct letter to show what kind of simple machine each object is. You will need to use some of the letters more than once.

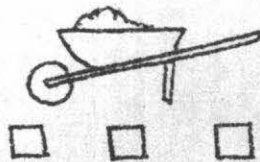
- | | | | | |
|--------------------|-------|---------------------|-------|-------|
| A. Lever. | i. | a ship's gangplank. | i. | _____ |
| B. Wheel and axle. | ii. | | ii. | _____ |
| C. Inclined plane. | iii. | a pair of pliers. | iii. | _____ |
| D. Wedge. | iv. | a screwdriver. | iv. | _____ |
| E. Pulley. | v. | a lift (elevator). | v. | _____ |
| | vi. | a garden fork. | vi. | _____ |
| | vii. | a flight of stairs. | vii. | _____ |
| | viii. | a door knob. | viii. | _____ |

6. Print the letters F (fulcrum), E (effort), and L (load or resistance) in their correct places in the boxes.

(1)



(11)

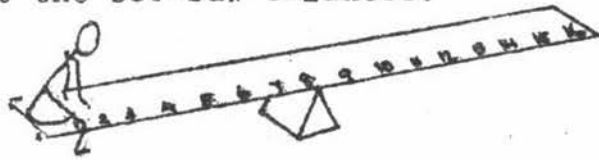


7. I wish to lift a weight of 2,000 pounds up to a height of six feet. Which of the following simple machines would I use if I had only manpower to help me?

- A. Inclined plane.
- B. Wedge.
- C. Wheel and axle.
- D. Lever.

7. _____

8. A girl weighing 50 pounds is sitting on the end of the see-saw. Where would a boy weighing 100 pounds have to sit so that the see-saw balances?



- A. At 16? B. At 14? C. At 12? D. At 10?

8. _____

9. Which arrangement of lever and fulcrum would you use to lift a very heavy load?



A.



B.



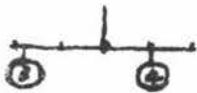
C.



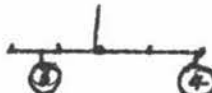
D.

9. _____

10. Which of these pictures shows a correct balance of a three pound weight and a four pound weight?



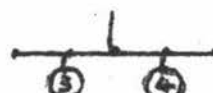
A.



B.



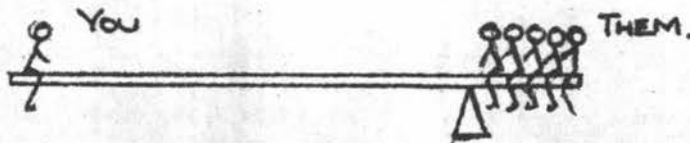
C.



D.

10. _____

11. Look at the drawing and then choose the reason which BEST explains why this balance works.



- A. YOU are five times as heavy as any one of THEM.
 B. The plank is five times heavier than the usual see-saw plank.
 C. The five of THEM take up five times as much space as YOU do.
 D. The long arm of the plank is five times longer than the short arm.

11. _____

12. Draw a simple machine, name the parts, and say how it works.

SCIENCE TEST - FORCE, WORK AND ENERGY (B)

Date: _____

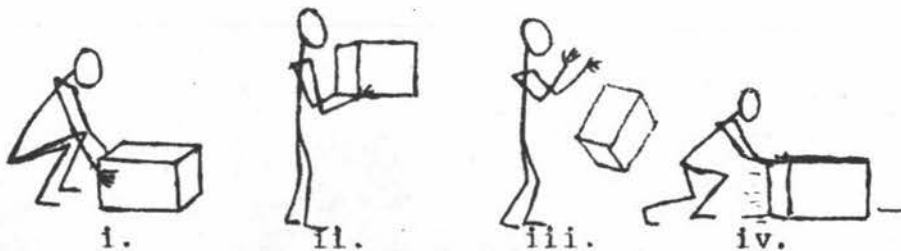
Name: _____

Unless the question has special instructions (for example Question Two) you should record your answer by printing, in the space at the right hand side, the letter (A, B, C or D) showing the answer you have chosen to be correct.

1. A push, a pull, and gravity are examples of a particular kind of:
- work.
 - inertia.
 - friction.
 - force.

1. _____

2. Below are four pictures, i, ii, iii and iv. Under the pictures are some phrases to describe what is happening. The phrases have letters beside them. Choose the correct phrases to go with each picture, and put their letters in the right answer space. Each answer space will have at least two letters in it.



2. i _____

ii _____

iii _____

iv _____

- No work done.
- Work done.
- Kinetic energy gained.
- Potential energy gained.
- Friction acting.
- Gravity acting.

3. If I had four blocks, each of the same weight, which one would be hardest to push along a wooden floor?
- A concrete block.
 - A glass block.
 - A wooden block.
 - An ice block.

3. _____

4.

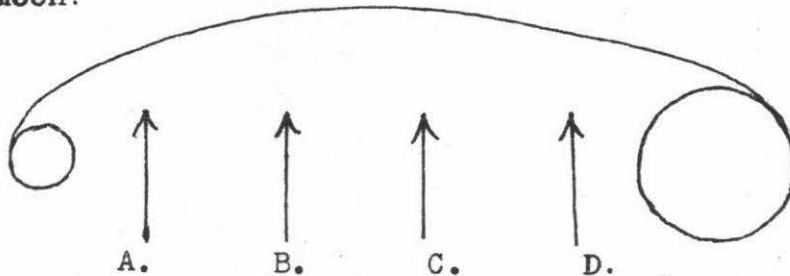


Which is the BEST reason to explain why this racing car is hard to roll over?

- A. The difference in size between back and front wheels gives more grip.
- B. The length of axle gives a wide wheel base.
- C. The width between the wheels keeps the weight low.
- D. The large rear tyres increase the amount of friction.

4. _____

5. Which arrow shows where the moon's gravity would cause a space craft to speed up again in a flight to the moon?



5. _____

6. The pull of gravity makes it harder to lift a heavy weight:

- A. at the south pole.
- B. in space above the earth.
- C. halfway between equator and pole.
- D. at the equator.

6. _____

7. The astronauts on the moon leaned forward in the direction they wished to walk. They did this because:

- A. the moon's surface was so soft, it took more energy to walk on it.
- B. they were protecting their eyes from the glare of the strong sunlight.
- C. the solar wind pushed so strongly against them.
- D. they had the life-support packs on their backs.

7. _____

8. Which statement is NOT correct?

- A. Inertia is the tendency to resist movement.
- B. Inertia is the tendency to continue moving in a set direction.
- C. Inertia is absent on the moon.
- D. Inertia acts on all stationary objects.

8. _____

9. An object like a pyramid is hard to tip over. Which is the BEST reason to explain this?

- A. The object has a low weight centre.
- B. The object has large flat surfaces.
- C. The object has a large mass.
- D. The object has angled sides.

9. _____

10. The Leaning Tower of Pisa, sketched below, does not fall over because:



- A. it has greater weight in the bottom half than in the top half.
- B. it has no strong force acting on it.
- C. the line through which gravity acts falls inside the base.
- D. the earth's rotation acts against its falling.

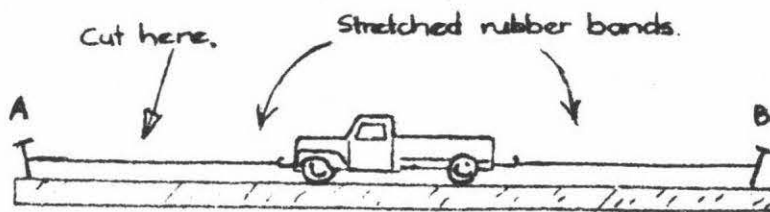
10. _____

11. When a force begins to tilt a stable object, the object tends to:

- A. take up a new position.
- B. move away from its original position.
- C. move position in line with the force being used.
- D. return to its original position.

11. _____

12. Look at the picture and then answer the question.



If the rubber band is cut at the place marked, the toy truck would:

- A. not move.
- B. move towards B.
- C. move towards A.
- D. fly straight up.

12. _____

13. How would you make the truck move in the opposite direction?

- A. Change the rubber bands before starting the experiment.
- B. Take the wheels off the truck.
- C. Drive the nail at A further in.
- D. Cut the other rubber band.

13. _____

14. Which rule explains why the truck moves?

- A. Whenever two forces act on an object, that object does not move.
- B. When an object is moving, the forces acting on it are in balance.
- C. Whenever a force acts on an object, that object must move.
- D. When an object is set in motion, the forces acting on it are not in balance.

14. _____

15. Which form of energy allowed this experiment to work?

- A. Molecular energy.
- B. Chemical energy.
- C. Potential energy.
- D. Kinetic energy.

15. _____

APPENDIX A.6: ADMINISTRATION - PUPIL OPINION SURVEY

The portion of the Survey relevant to this study is Part One, which includes scales of

- (a) attitudes to academic achievement, and
- (b) interest in science.

The items for (b) are numbers 2, 7, 16, 17, 18, 19, 21, 26, 28 and 29. The remaining twenty items comprise (a).

The test was administered by class teachers.

GENERAL INSTRUCTIONS TO TEACHERS.

Please ensure that children have pencils and rubbers. Issue the booklets face up, asking the children to write their names, whether they are a boy or a girl (because of possible name/sex confusion, e.g. Leslie), date, and their age. Actual years and months can be filled in later from class registers.

INTRODUCTORY STATEMENT TO CHILDREN.

"It is not often that we ask girls and boys to tell us what they feel about a number of things which are part of being at school. Today, we are interested in finding out what pupils think. The booklet just issued has a number of statements about school life; some you might agree with, and some you might disagree with. The important thing is to give your own opinion, and not one you might think will please me. In fact, I (your class teacher) will not be reading your answers.

Now let us read through the front part, under 'General Directions', to see how we shall answer."

NOTE TO TEACHERS.

- Please read the front page aloud to children, stressing
- (a) the non-test nature of the Survey, and
 - (b) the need for care in placing ticks in the correct box beside the statement given.

Work through the practice example, to ensure that pupils understand what to do, before telling them to begin.

The Survey is expected to take fifteen to twenty minutes for children to complete. If any pupil has difficulty reading any statement, please read it quietly to him. Thank you.

PUPIL OPINION SURVEYGeneral Directions:

We are interested in how girls and boys feel about a number of things, and we would like you to help us find out. The answers are needed for part of a survey of pupils' opinions.

1. This is not a test, so there is no right answer. Answer the questions as honestly as you can. Then the answers we get will be truly the opinions of girls and boys. Your answers are not going to be shown to other people.
2. Do not spend time puzzling over any of the questions, but give the first answer that comes easily to you. Some questions are similar to other ones, but no two are exactly alike. Your answers may differ in these cases. Your answer is what is true of you MOST TIMES.
3. While answering the survey, you will find special instructions headed "What to do." Please follow these instructions as they apply to you, when you come to them.
4. The questions and sentences you will read are in this booklet. We would like you to put your answers in the answer space, alongside the same number as the question.
5. Put your answer by making a tick in the box which shows your opinion. Answer each question only once.
6. If you have any questions while you are doing the survey, raise your hand and someone will come to help you.

Name: _____

Boy/Girl: _____

Age: _____ years _____ months

Date: _____

Practice Example:

Code: Strongly agree Agree Neither agree nor disagree Disagree Strongly disagree

A

B

C

D

E

X. I like icecream.

A	B	C	D	E

Y. I like porridge.

A	B	C	D	E

Part I: What To Do.

Below are a number of sentences. To answer, put a tick in the box beside the letter which best tells your opinion. This is the code we use:

Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly Disagree
A	B	C	D	E
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Begin with sentence number one. Be sure you put your opinion of sentence 1 in one of the boxes beside 1 in your answer space.

	A	B	C	D	E
1. I would rather work with my hands than do ordinary schoolwork.					
2. I enjoy trying to find reasons to explain why things happen as they do.					
3. I am more interested in games and sport than I am in schoolwork.					
4. If I had to choose between taking part in a competition or being the judge, I would choose to be the judge.					
5. I stick to a project, or piece of work, until it is finished even though it is dull and boring.					
6. It worries me when others get better marks than I do.					
7. I enjoy lessons when we are shown experiments and told why they work.					
8. I prefer to sit at the back of the classroom					
9. My friend stopped running hard when he saw that he was going to lose. I would have done the same if I had been running in the race.					
10. I like being asked questions in class.					

- | | A | B | C | D | E |
|---|---|---|---|---|---|
| 11. My friends think I don't take my schoolwork seriously. | | | | | |
| 12. I work hard most of the time. | | | | | |
| 13. When someone is being praised I find myself wishing it were me. | | | | | |
| 14. I should like to leave school as soon as I am fifteen. | | | | | |
| 15. If I get lower marks than usual in a test I feel disappointed. | | | | | |
| 16. If someone told me that our bad weather is caused by atomic bombs and rocket blasts I would want to know how he could prove it. | | | | | |
| 17. Collecting things like rocks and plants is a dull hobby. | | | | | |
| 18. I would rather find out things for myself than be told. | | | | | |
| 19. Science is something scientists do and does not affect my life very much. | | | | | |
| 20. I often compare my work with the work of others. | | | | | |
| 21. Science is one of the most interesting subjects at school. | | | | | |
| 22. My mind often wanders off the subject at school. | | | | | |
| 23. I usually leave my homework until the last minute. | | | | | |
| 24. I enjoy trying to find the answers to difficult problems. | | | | | |
| 25. I enjoy being a class leader. | | | | | |
| 26. I like knowing how things move and why they move | | | | | |

- | | A | B | C | D | E |
|---|---|---|---|---|---|
| 27. If I do badly in one school subject I find I do not try as hard next time we have that subject. | | | | | |
| 28. Art does more for man than science does. | | | | | |
| 29. I like reading books about space exploration. | | | | | |
| 30. I tell my parents about my successes at school. | | | | | |

Part II: What To Do.

Please look at the five subjects listed below:

ART, MATHEMATICS, SCIENCE, SOCIAL STUDIES, STORY-WRITING.

In the answer spaces, at the right write them in the order in which you like them, starting at the top with the one you like most and finishing at the bottom with the one you like least of the five.

APPENDIX A.7: EXAMPLE OF TEST OBJECTIVESKNOWLEDGE AND UNDERSTANDING LEVEL

	Item No.	No. of items
<u>Discrimination of unstable objects:</u>		3
Relative to height	1.	
Relative to base and height	9.	
Relative to centre of gravity and support	4.	
<u>Identification of stable objects:</u>		3
Relative to weight	2.	
Relative to base area	3.	
Relative to height, weight and base	6.	
<u>Concept of gravity, defining:</u>	19.	1
<u>Centre of gravity, concept:</u>		8
Attributes of centre of gravity	5.	
Locating centre of gravity	8.12.14.	
Relative to centre of gravity and base	7.10.	
Defining centre of gravity	13.	
Relative to balance and support	11.	
<u>Line of gravity, concept:</u>		2
Relative to base	17.	
Relative to support point	18.	
<u>Understanding ways of increasing stability:</u>		2
By adjusting centre of gravity	16.	
By altering one dimension	20.	
<u>Understanding relationship between a force and stability:</u>	15.	1

APPENDIX A.8: ITEM ANALYSIS DATACOGNITIVE STYLE MEASURES.

TRIAD ONE					TRIAD TWO				
Item	≠ (D)	≠≠ (C)	≠≠≠ (R)	Mean	Item	≠ (D)	≠≠ (C)	≠≠≠ (R)	Mean
K. 1	.72	.26	.78	.58	1.a.	.66	.73	.84	.72
2	.59	.50	.53	.54	b.	.64	.49	.37	.50
3	.63	.50	.80	.64	c.	.71	.51	.12	.45
K. 4	.65	.56	.37	.57	2.a.	.53	.70	.30	.51
K. 5	.68	.42	.71	.60	b.	.71	.88	.63	.72
6	.51	.25	.63	.49	c.	.68	.77	.58	.64
7	.68	.42	.37	.49	3.a.	.63	.53	.59	.58
K. 8	.80	.42	.68	.63	b.	.80	.62	.79	.72
9	.70	.25	.73	.56	c.	.50	.31	.68	.49
K.10	.58	.30	.40	.42	4.a.	.77	.53	.71	.67
K.11	.48	.50	.43	.47	b.	.22	.70	.60	.51
12	.37	.23	.77	.44	c.	.55	.51	.51	.52
13	.77	.58	.84	.73	5.a.	.66	.71	.81	.73
14	.75	.66	.74	.71	b.	.84	.50	.98	.77
K.15	.60	.45	.43	.49	c.	.76	.38	.68	.61
K.16	.30	.36	.63	.43	a = First choice analysis b = Second choice analysis c = Third choice analysis				
17	.51	.60	.68	.59					
18	.63	.71	.51	.62					
K.19	.84	.55	.88	.76	≠ = Descriptive mode ≠≠ = Categorical-inferential mode ≠≠≠ = Relational-contextual mode				
K.20	.63	not chosen	.78	.37					
21	.48	.30	.70	.49					
22	.58	.60	.75	.64					
23	.66	.53	.79	.66					
24	.58	.30	.52	.47					
25	.70	.25	.73	.56					

K. = From a published Kagan item.
 Remaining items newly constructed.

APPENDIX A.9: ITEM ANALYSIS DATASCIENCE CONTENT MEASURES.

	Embedded Items		Initial Learning		Retention		First Application		Second Application-Transfer	
	Diff.	Disc.	Diff.	Disc.	Diff.	Disc.	Diff.	Disc.	Diff.	Disc.
1	.27	.17	.80	.61	.88	.51	.53	.56	.15	.51
2	.24	.27	.74	.56	.87	.55	.81	.18	.48	.67
3	.16	.21	.74	.66	.89	.48	.72	.45	.47	.71
4	.23	.36	.66	.18	.65	.13	.75	.66	.57	.68
5	.08	.19	.92	.43	.93	.37	.57	.31	.30	.38
6			.92	.19	.90	.43	.74	.56	.37	.77
7			.35	.57 \times	.38	.56 \times	.51	.41	.42	.25
8			.60	.71	.75	.41	.58	.57	.65	.75
9			.71	.38	.82	.58	.33	.17		
10			.38	.47 \times	.46	.41 \times	.75	.66		
11			.47	.67	.53	.63	.40	.47		
12			.79	.47	.85	.51	.20	.18		
13			.34	.32	.43	.42	.45	.13		
14			.76	.53	.78	.47	.53	.16		
15			.18	.58 \times	.30	.73 \times	.50	.67		
16			.22	.27 \times	.35	.48 \times	.52	.60		
17			.56	.57 \times	.53	.71 \times	.66	.45		
18			.20	.18	.30	.34	.33	.51		
19			.58	.42	.57	.60	.58	.49		
20			.66	.61	.65	.48	.37	.51		

Embedded items marked \times in Initial Learning and Retention tests

APPENDIX A.10: ITEM ANALYSIS DATAPUPIL OPINION SURVEY.

Item Number	Academic Achievement Motivation	Item Number	Interest in Science
1	8.75	2	5.78
3	2.98	7	9.19
4	1.72	16	3.06
5	2.96	17	8.04
6	2.92	18	4.24
8	1.97	19	7.70
9	4.77	21	7.96
10	9.72	26	4.90
11	6.99	28	2.46
12	5.65	29	7.65
13	5.25		
14	4.39		
15	2.89		
20	2.54		
22	5.02		
23	4.84		
24	5.58		
25	7.96		
27	9.20		
30	5.02		

Figures reported are t-values.
$$t = \frac{\bar{X}_H - \bar{X}_L}{\sqrt{\frac{S_H^2}{n_H} + \frac{S_L^2}{n_L}}}$$

Edwards (1957, pp.149-171) suggests 1.75 as a significant discriminating value if there are 25 or more subjects in each of the upper and lower groups.

APPENDIX B: ADDITIONAL DETAILS OF TRAINING PROGRAMME

- B.1 The Training Programme.
- B.2 A Teaching Paradigm.
- B.3 Summary of Lesson Principles.
- B.4 Summarised Lesson Guides.
- B.5 Simplified Task Structures.

APPENDIX B.1: THE TRAINING PROGRAMME

This appendix gives further details of the programme of training for the teachers and observers in this study. Subsections of the appendix include:


- (a) task structures of the experimental lessons and their links with the Form I science curriculum unit taught just prior to the experimental work;
- (b) summaries of the major concepts and principles of stability (the content of the research lessons), and
- (c) the content of the lesson guides developed during the training programme.

A general description of the training programme is found in the main text (pp.93-95).

The initial approach to the study of learning, through lecture and discussion, was a global S-O-R model owing much to Frandsen (1961), Mowrer (1951), Lindgren (in Bower and Hollister, 1967), and to C.G.N.Hill (personal communication). A number of basic distinctions were made - between psychological and educational views of learning (cf. Komisar, 1966), between performance and learning, and between stimulus and response as foci of attention rather than definitive entities. The perspective provided allowed the students to consider different theoretical orientations as reflecting different levels of analysis, different emphases (i.e. on stimulus or response or learner variables), different frames of reference (observer or participant), and different kinds of tasks and outcomes. The major views surveyed were conditioning theory (classical and operant - Skinner), imitation theory (Bandura and Walters, 1963), cognitive theory (Ausubel, 1968; Suchman, 1961; Bruner, 1966). Particular attention was given to meaningful reception learning, discovery learning and reinforcement theory.

A teaching paradigm was built up with the group of students, following the study of learning, some task analysis activities (see attached task structures), the study of the conceptual basis of the science lessons (see summary of principles attached), and consideration of abilities, attitudes and prior experiences, especially in science, of Form I children. The general model follows.

APPENDIX B.2: A TEACHING PARADIGM

(a)	(b)	(c)	(d)
<u>Objectives</u>	<u>Predispositions</u>	<u>Process Conditions</u>	<u>Learning and Consequences</u>
for particular children taught - Cognitive - Affective	- Attention - Motivation - Prior learning (cognitive content and structure) - Abilities, attitudes and modes and developmental features	- Task variables e.g. nature of task sequence form (e.g. iconic) analysis - Teacher functions e.g. stimulus presenting motivating attention gaining guiding thinking (e.g. Gagne and Aschner) - Conditions and interactions e.g. practice reinforcement etc. - Strategies and tactics e.g. expository guided discovery	- As process and product - Feedback  - Transfer - As affecting predispositions for further learning
General facilitating or inhibiting factors e.g. teacher and pupil personality factors school and community factors - sociological, environmental, financial			

APPENDIX B.3: SUMMARY OF LESSON PRINCIPLESSTABILITY

A body may be placed in several positions and still balance. However, some positions are less secure than others, less stable.

1. Stability is directly proportional to the area of the base of the body.
2. Stability is directly proportional to the weight of a body.
3. Stability is inversely proportional to the height of a body.
4. Stability in a given direction is directly proportional to the horizontal distance between the vertical line through the centre of gravity and a pivoting edge of the body.
5. Depending on the distribution of mass of an object, the lower in space the concentration of mass (weight for these children), the further the object must be tipped before it loses stability.
6. An object is stable when its centre of gravity is directly over its supporting base. The nearer to the centre of that base the line of gravity falls, the less likely it is that the object will become unstable.
7. If, when an object is displaced, the line of gravity of the object falls outside the base of support, then the equilibrium is unstable and the object will seek a new base until stable equilibrium is reached.
8. The equilibrium of a body is said to be stable if, on being slightly displaced, the body tends to return to its original position. Stability depends primarily on the location of the centre of gravity in relation to the supporting base, or point of support.

APPENDIX B.4: SUMMARISED LESSON GUIDES.I - RULES EXPLAINED AND DEMONSTRATED (R.E.)LESSON ONE.Concepts and Principles.

Centre of gravity, balance point, location of centre of gravity; attributes of stable objects, weight, height, base area.

General Set.

To know and understand.

Sequence.

- (a) Introduction. Facilitate relatability, anchorage, by revising concept of gravity taught in preceding unit. Gravity reviewed as force that pulls; weight as the measure of the amount of force; direction of force towards earth centre. Thus, can use plumb-line to indicate vertical direction of pull. Gravity acts, pulling on all objects. The closer to earth's centre, the greater the pull or weight and force. Exemplify using globe.
- (b) Centre of gravity. Discuss feel of heavy objects, e.g. shot. Weight does not seem to be on the surface; it pulls through the object; seems to be, as it were, inside the object. This part where weight seems to be concentrated, where the downward pull of gravity seems to be working, is called the centre of gravity.
- (c) Centre of gravity, as centre of balance. Demonstrate with ruler, by placing index fingers at ends, and moving fingers towards centre. Demonstrate with rulers using different commencing positions for forefingers. Note that the fingers always meet at 6" mark, in line with centre of gravity. Children practise with rulers, using different starting points. Explain that centre of gravity of an object is constant. Ruler balances at this point.
- (d) Locating centre of gravity. Demonstrate with cardboard (regular and irregular shapes) and plumb-line to show intersecting lines meet on surface above centre of gravity. Children practise locating centre of gravity with cardboard shapes. Demonstrate and get children to estimate position of centre of gravity in various objects, e.g. gourd, book, ball, vase, stone. Establish notion of constant centre of gravity irrespective of spatial orientation of objects.

- (e) Attributes of stable objects. Demonstrate with objects having same base areas, similar weight, but differing height. Establish idea of low centre of gravity, i.e. low squat objects are more stable than tall objects. Demonstrate with objects how larger base area contributes to stability, how objects with same dimensions but differing weight are stable in proportion to the amount of weight. Children experiment with equipment (i.e. practise) to facilitate understanding.
- (f) Review principles, attributes and concepts explained and demonstrated. Try to establish set of anticipation for next lesson.

LESSON TWO.

Concepts and Principles.

Balance point of objects, support point in line with centre of gravity, both above and below object, and with irregular as well as regular 3-D objects; increasing stability of objects at rest and in balance; line of gravity within base, lowering centre of gravity by widening base, increasing weight below line of support.

General Set.

To know and understand.

Sequence.

- (a) Balance point of objects. Demonstrate with needle, string and cardboard circles, squares and irregular shapes. how these will hang evenly and flat if suspended directly above centre of gravity. Demonstrate support from below, through line of gravity. Apply to solid objects, and estimate balance or support point. Establish principle that when balancing any object, the point of support is directly above or underneath the centre of gravity. Children practise with objects.
- (b) Line of gravity and base zone. Demonstrate with large wooden blocks, pieces of cardboard, how objects when slightly displaced tend to return to original position. Use plumb-line attached to block to show how, when the line of gravity falls outside the base, the block is no longer stable, and topples. Exemplify with reference to many objects in environment, e.g. tractors on hillsides, Leaning Tower of Pisa. Children practise with own rulers, to find point at which objects become unstable.

- (c) Increasing stability by lowering centre of gravity.
Demonstrate and explain how objects can be made stable by increasing the base area (e.g. ruler with plasticine, two rulers as a step-ladder, etc.) by increasing the weight (e.g. add weight to ruler, thicker wood, steel etc.). Exemplify principles at work in constructing tall buildings (wide bases, lower weight centre with deep foundations, steel girders etc.). Provide a demonstration with ruler balanced on edge, on a needle. weight added at ends of ruler so that it then rotates and maintains equilibrium.
- (d) Revise principles with reference to demonstrations - centre of gravity, line of gravity, support, attributes of stable objects. Relate to pupils' own experiences.

II - RULES DERIVED WITH GUIDANCE

LESSON ONE.

Concepts and Principles. As for Lesson One, R.E.

General Set. To find out, to understand.

Sequence.

- (a) Centre of gravity. Ask children to feel and observe objects (book, shot, ball, chunky stone, blocks of wood etc.). Guiding questions. Where is it heavy? Any particular place? Outside, or on the surface? Inside? Does the weight seem to pull through the object and your hand in any particular place? Where? Why might it seem to be there? What could we call this point?
Additional situational examples if needed. Man carrying a ladder, a mattress, a full shopping bag.
- (b) Centre of gravity as centre of balance. Ask children to place forefingers at 1" and 11" marks on ruler, and push fingers together. Ask them to predict where the fingers will meet. Try 1" and 8" marks, and repeat. Guiding questions. At what point do you think fingers will meet? What happened? Why did they meet there? Where does the weight seem to be? What name could we give to this point? Can we make any general conclusion?

- (c) Locating centre of gravity. Issue equipment - stands, plumb-lines, cardboard shapes (regular and irregular). Ask children to try and find the balance point, centre of gravity, of the shapes. Use regular shapes first, e.g. cardboard circles.
Guiding questions. Can you find the balance point, the centre of gravity? Is it the same as the middle point? Use regular and irregular shapes). Does the centre of gravity alter position if we turn the shape? What can we conclude?
 Ask children to make predictions about more objects (e.g. stones, candle, bottles, blocks of wood) and to estimate where centre of gravity would be. Try turning objects to different positions to consider the constancy of the centre of gravity.
- (d) Attributes of stable objects. Ask children to compare objects (e.g. two matchboxes, one full, one empty; two essence bottles, one empty and the other containing sand; two candles, one tall and one short; two Cuisinaire rods, one long and the other short; two wooden blocks with different base areas; objects with bases of different shapes). Ask children to estimate position of centre of gravity.
Guiding questions. Which is the harder to tip over? Which is more stable? Why? What makes one object more stable than another? Are objects more stable in some positions than in others? Why?
 Draw conclusions from children, and summarise them on the blackboard.
- (e) Revise. What have we found out today? To conclude, ask children to draw a stable vase, or an unstable building. Try to establish set of anticipation for the next lesson.

LESSON TWO.

Concepts and Principles. As for Lesson 2, R.E.

General Set. To find out, to understand.

Sequence.

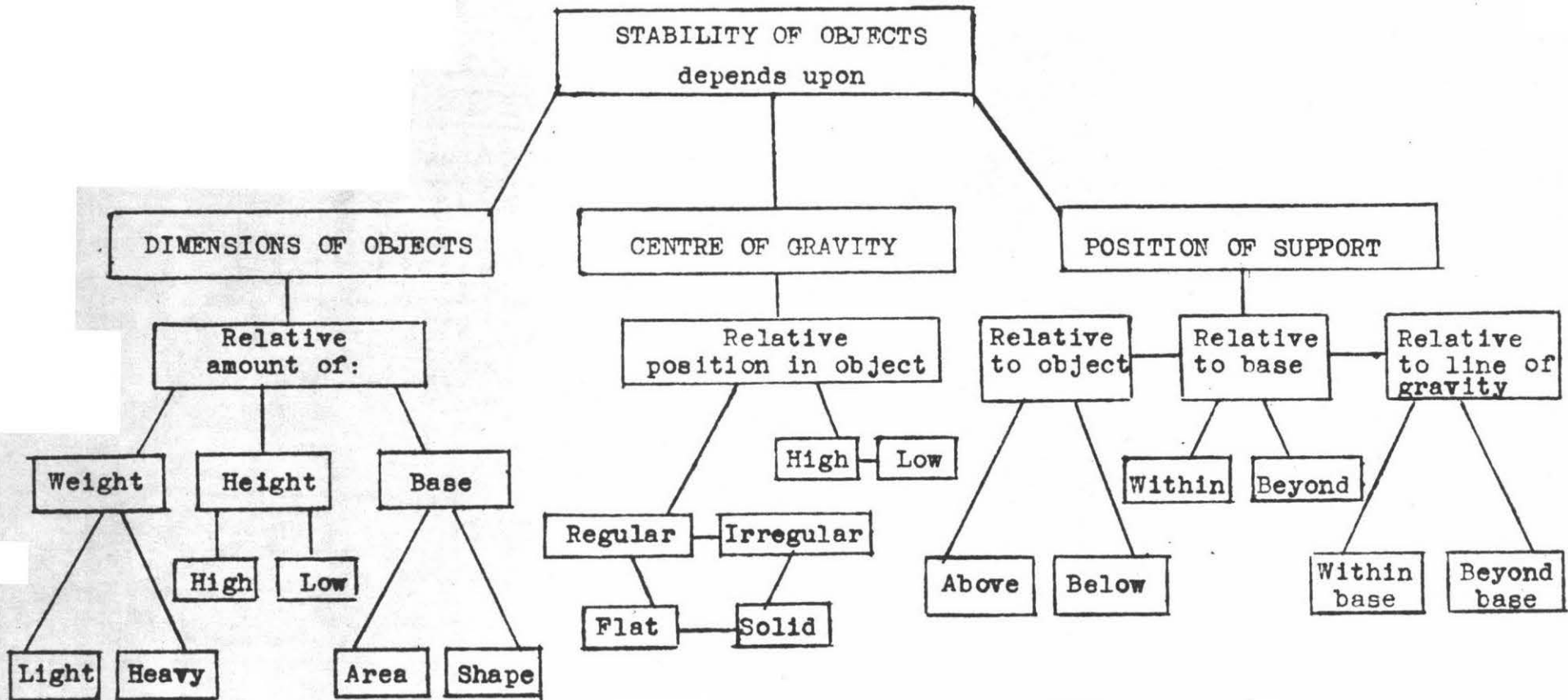
- (a) Balance point of objects. Issue equipment (cardboard shapes, string, needles, washers as weights). Ask children to suspend shapes so that they hang evenly and level.
Guiding questions. Where will I attach the string make the shape hang evenly? What do you notice about

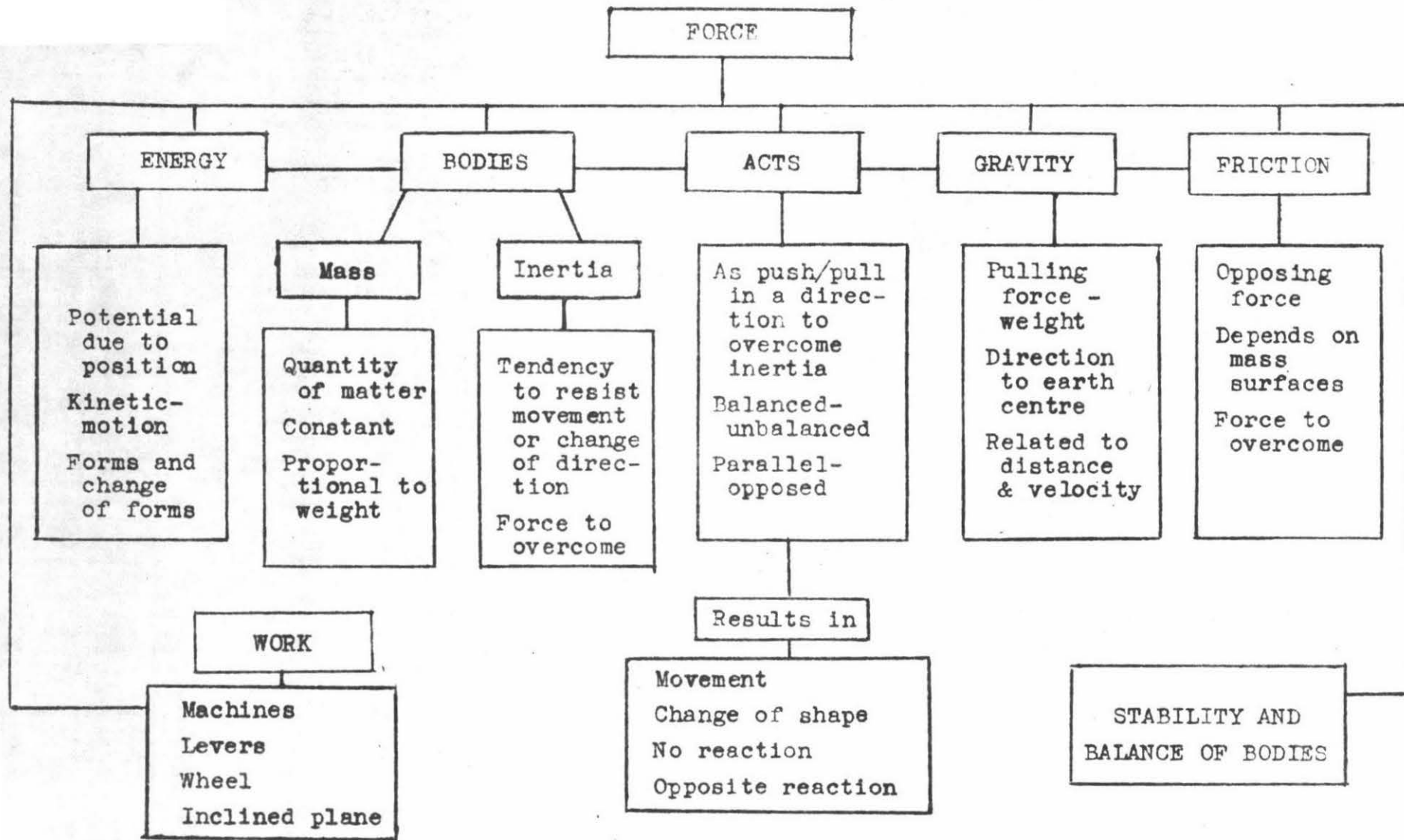
the point from which the card is hanging? Where is the support point in relation to the balance point (the centre of gravity)?

Repeat, with support point below the object. Try, through questions to establish relationship between line of gravity and support point. Through questions and answers, extend to other environmental objects (e.g. Christmas decorations).

- (b) Line of gravity and base zone. Refer to equipment - blocks of wood, cardboard rectangles, plumb-lines. Ask children to displace objects to differing angles, noting position of plumb-line.
Guiding questions. What does the object do when it is tilted only slightly? More? Considerably? When does the object become unstable? What happens to the line of gravity? What can we say about stable objects?
- (c) Increasing stability by lowering centre of gravity. Issue equipment (bottle containing cork with a needle inserted point upwards, washers, fine wire, rulers for those who forgot their own). Ask children to try to balance the ruler on its edge, on needle point.
Guiding questions. What happens when we try to balance the ruler on its edge on the point of the needle? How could we make it balance? What do the washers do? What happens to the centre of gravity of the ruler? Where is the line of gravity? In what other ways can we make objects more stable?
- (d) Review principles derived, by class discussion. Questions and answers with teacher, hints, but no 'telling'. Ask for environmental examples based on children's own experience.

SIMPLIFIED TASK STRUCTURE (EXPERIMENTAL LESSONS)





(Control group lessons)

(Experimental lessons)

APPENDIX C: THE GENERAL MODELS OF THE ANALYSES
OF VARIANCE AND OF COVARIANCE

- C.1 Analysis of Variance: General model for a seven-way experiment, one factor being nested and doubly-crossed.
- C.2 Sample entries from a components analysis for a seven-way experiment, one factor being nested and doubly-crossed.
- C.3 Analysis of Covariance: General model for a four-factor covariance design.

APPENDIX C. 1: ANALYSIS OF VARIANCE: GENERAL MODEL

The model for a seven-way experiment containing a double crossing of a nested factor may be written as:

$$\begin{aligned}
 x_{ijklrst} = & M + A_i + B_j + C_k + D_l + E_r + F_s \\
 & + (AB)_{ij} + (AC)_{ik} + \dots + (EF)_{rs} \\
 & \quad \text{Sum of fifteen terms} \\
 & + (ABC)_{ijk} + (ABD)_{ijl} + \dots + (DEF)_{lrs} \\
 & \quad \text{Sum of twenty terms} \\
 & + (ABCD)_{ijkl} + (ABCE)_{ijkr} + \dots + (CDEF)_{klrs} \\
 & \quad \text{Sum of fifteen terms} \\
 & + (ABCDE)_{ijklr} + (ABCDF)_{ijkls} + \dots + (BCDEF)_{jklrs} \\
 & \quad \text{Sum of six terms} \\
 & + (ABCDEF)_{ijklrs} + P_{ijklt} \\
 & + (PE)_{ijklrt} + (PF)_{ijklst} + (PEF)_{ijklrst}
 \end{aligned}$$

where

$x_{ijklrst}$ is the score of person (t) in level (i) of factor A, level (j) of factor B, level (k) of factor C, level (l) of factor D, level (r) of factor E and level (s) in factor F;

M is a component to all the scores;

A_i is a component common to all scores in level (i) of factor A;

B_j is a component to all scores in level (j) of factor B;

• • • • •

$(AB)_{ij}$ is a component resulting from the interaction of level (i) of factor A and level (j) of factor B;

• • • • •

Appendix C.1: (Continued)

(ABC)
ijk is a component resulting from the interaction of level (i) of factor A, level (j) of factor B and level (k) of factor C;

• • • • •
P_{ijklt} is a component^{common} to all the scores of pupil (t) in level (i) of factor A, level (j) of factor B, level (k) of factor C and level (l) of factor D;

• • • • •
(PE)
ijklrt is a component resulting from the interaction between pupil (t) (in level (i) of factor A, level (j) of factor B, level (k) of factor C and level (l) of factor D), and level (r) of factor E;

and

(PEF)
ijklrst (the residual term) is a component resulting from the interaction between pupil (t) (again in level (i) of factor A, level (j) of factor B, level (k) of factor C and level (l) of factor D) and level (r) of factor E and level (s) of factor F.

Generally, (i) runs from 1 to a, (j) from 1 to b, (k) from 1 to c, (l) from 1 to D, (r) from 1 to c, (s) from 1 to f, and with an equal number of scores (n) in each cell, (t) runs from 1 to n. The sixty-eight contributions to the score $x_{ijklrst}$ are all independent of each other, and the $A_s, B_s, C_s, \dots, (PEF)_s$ are regarded as being drawn from normally distributed populations with means of zero and variances of $\sigma^2_A, \sigma^2_B, \sigma^2_C, \dots, \sigma^2$ respectively.

The model, described above, was developed from a study of Winer (1962) and Lewis (1968). The notation employed follows that used by Lewis.

APPENDIX C 2: SAMPLE ENTRIES FROM A COMPONENTS ANALYSIS FOR A SEVEN-WAY
EXPERIMENT, ONE FACTOR BEING NESTED AND DOUBLY CROSSED

Source of Variation	Degrees of Freedom	Mean Square	Expectation	Appropriate Error Term
A	a-1	$\sigma^2_{PEF.ABCD}$	$+ ef\sigma^2_{P.ABCD} + pbcdef\sigma^2_A$	P
B	b-1	$\sigma^2_{PEF.ABCD}$	$+ ef\sigma^2_{P.ABCD} + pacdef\sigma^2_B$	P
C	c-1	$\sigma^2_{PEF.ABCD}$	$+ ef\sigma^2_{P.ABCD} + pabdef\sigma^2_C$	P
D	d-1	$\sigma^2_{PEF.ABCD}$	$+ ef\sigma^2_{P.ABCD} + pabcef\sigma^2_D$	P
E	e-1	$\sigma^2_{PEF.ABCD}$	$+ f\sigma^2_{PE.ABCD} + pabcdf\sigma^2_E$	PE
F	f-1	$\sigma^2_{PEF.ABCD}$	$+ e\sigma^2_{PF.ABCD} + pabcde\sigma^2_F$	PF
AB	(a-1)(b-1)	$\sigma^2_{PEF.ABCD}$	$+ ef\sigma^2_{P.ABCD} + pcdef\sigma^2_{AB}$	P
AC	(a-1)(c-1)	$\sigma^2_{PEF.ABCD}$	$+ ef\sigma^2_{P.ABCD} + pbdef\sigma^2_{AC}$	P
.....				
DE	(d-1)(e-1)	$\sigma^2_{PEF.ABCD}$	$+ f\sigma^2_{PE.ABCD} + pabcf\sigma^2_{DE}$	PE
DF	(d-1)(f-1)	$\sigma^2_{PEF.ABCD}$	$+ e\sigma^2_{PF.ABCD} + pabce\sigma^2_{DF}$	PF
EF	(e-1)(f-1)	$\sigma^2_{PEF.ABCD}$	$+ ef\sigma^2_{P.ABCD} + pabcd\sigma^2_{EF}$	PEF
.....				
ACE	(a-1)(c-1)(e-1)	$\sigma^2_{PEF.ABCD}$	$+ f\sigma^2_{PE.ABCD} + pbd\sigma^2_{ACE}$	PE
.....				

Appendix C. 2: (Continued)

Source of Variation	Degrees of Freedom	Mean Square	Expectation	Appropriate Error Term
ABEF	(a-1)(b-1) ... (f-1) product of 4 terms	$\sigma^2_{PEF.ABCD}$	+ $ef\sigma^2_{P.ABCD}$ + $pcd\sigma^2_{ABEF}$	PEF
ABCDEF	(a-1)(b-1) ... (f-1) product of 6 terms	$\sigma^2_{PEF.ABCD}$	+ $ef\sigma^2_{P.ABCD}$ + $p\sigma^2_{ABCDEF}$	PEF
P within ABCD	abcd (p-1)	$\sigma^2_{PEF.ABCD}$	+ $ef\sigma^2_{P.ABCD}$	PEF
PE within ABCD	abcd (p-1)(e-1)	$\sigma^2_{PEF.ABCD}$	+ $f\sigma^2_{PE.ABCD}$	PEF
PF within ABCD	abcd (p-1)(f-1)	$\sigma^2_{PEF.ABCD}$	+ $e\sigma^2_{PF.ABCD}$	PEF
PEF within ABCD	abcd (p-1)(e-1)(f-1)	$\sigma^2_{PEF.ABCD}$		

1. These are written for a levels of A, b levels of B, c levels of C, d levels of D, e levels of E, f levels of F and p levels of P within each of the abcd cross classifications of A,B,C,D.
2. A,B,C,D,E,F are taken to be fixed effects, and P a random effect. The choice of notation follows a recommendation in Lewis (1968) for mixed models.
3. An appropriate error term to evaluate a source of variation is given by the mean square, estimating all but the last component in the mean square expectation for the source of variation (Lewis, 1968). This last component in the mean square expectation is the component involving the specific effect being tested.

APPENDIX C. 3: ANALYSIS OF COVARIANCE: GENERAL MODEL

Assuming an A x B x C x D factorial experiment having (n) observations in each cell, the model for a four factor covariance design may be written as:

$$\begin{aligned}
 Y_{ijklt} = & M + A_i + B_j + C_k + D_l + (AB)_{ij} + (AC)_{ik} + (AD)_{il} \\
 & + (BC)_{jk} + (BD)_{jl} + (CD)_{kl} + (ABC)_{ijk} + (ABD)_{ijl} \\
 & + (ACD)_{ikl} + (BCD)_{ijl} + (ABCD)_{ijkl} + R x_{ijklt} \\
 & + e_{ijklt}
 \end{aligned}$$

where

Y_{ijklt} is the score of person (t) in level (i) of factor A, level (j) of factor B, level (k) of factor C, and level (l) of factor D on the dependent measure;

M is a component common to all the scores from the dependent measure;

A_i is a component common to all the dependent measure scores in level (i) of factor A;

B_j is a component common to all the dependent measure scores in level(j) of factor B;

$(AB)_{ij}$ is a component resulting from the interaction of level (i) of factor A and level (j) of factor B;

.

$(ABCD)_{ijkl}$ is a component resulting from the interaction of level (i) of factor A, level (j) of factor B, level (k) of factor C and level (l) of factor D;

R is the regression of the covariate on the dependent measure, and is common to all the groups;

X_{ijklt} is the score on the covariate of person (t) in level (i) of factor A, level (j) of factor B, level (k) of factor C and level (l) of factor D, this score being expressed as a deviation from the overall mean score of the covariate;

and

e_{ijklt} is a component specific to person (t) in level (i) of factor A, level (j) of factor B, level (k) of factor C and level (l) of factor D.

Appendix C. 3: (Continued)

Generally, (i) runs from 1 to a, (j) from 1 to b, (k) from 1 to c, (l) from 1 to d, and with an equal number of scores (n) in each cell, t runs from 1 to n. The eighteen contributions to the score Y_{ijklt} are all independent of each other, and e_{ijklt} is such that for any given (ijkl) entry it can be regarded as drawn from a normally distributed population with means of zero and a variance of σ^2 .

The model described above, was developed from a study of Winer (1962) and Lewis (1968). The notation employed follows that used by Lewis.

APPENDIX D: TABLES OF RESULTS FROM THE
VARIOUS ANALYSES

- D.1 The main analysis of variance model: partitioning pattern and degrees of freedom.
- D.2 Table A. The main analysis of variance.
- D.3 Table B. Analysis of covariance, Test IA.
- D.4 Table C. Analysis of covariance, Test IIA.
- D.5 Table D. Analysis of covariance, Test IB.
- D.6 Table E. Analysis of covariance, Test IIB.
- D.7 Table F. Analysis of variance, Test IA.
- D.8 Table G. Analysis of variance, Test IIA.
- D.9 Table H. Analysis of variance, Test IB.
- D.10 Table I. Analysis of variance, Test IIB.
- D.11 Table J. Means and adjusted means for selected factors.
- D.12 Table K. Treatment means and other data.

THE MAIN ANALYSIS OF VARIANCE MODEL
(Partitioning pattern and degrees of freedom)

TOTAL VARIATION

BETWEEN CELLS

Main Effects:

d.f.
Terms.

A.	B.	C.	D.	E.	F.
1	1	2	1	1	1
6					

Interactions:

First Order

d.f.
Terms.

AxB.....	ExF
20	15

Second Order

d.f.
Terms.

AxBxC.....	DxExF
30	20

Third Order

d.f.
Terms.

AxBxCxD.....	Cx Dx Ex F
25	15

Fourth Order

d.f.
Terms.

AxBxCxDxE..	BxCxDxExF
11	6

Fifth Order

d.f.
Terms.

AxBxCxDxExF
2
1

WITHIN CELLS

P - Between pupils
within AxBxCxD
d.f. 96
Terms 1

PxE - Pupils x objectives
within AxBxCxD
d.f. 96
Terms 1

PxF - Pupils x occasions
within AxBxCxD
d.f. 96
Terms 1

PxExF - Residual variation
within AxBxCxD
d.f. 96
Terms 1

Total d.f. 479

TABLE A. MAIN ANALYSIS OF VARIANCE

Source	Sum of Squares	d.f.	Mean Square	Error Term	F - Ratio	P - Value	
A	1.875	1	1.875	P	9.135	1.00	
B	140.833	1	140.833	P	9.135	15.41	.001
C	23.829	2	11.914	P	9.135	1.30	
D	22.533	1	22.533	P	9.135	2.47	.20
E	0.133	1	0.133	PxE	2.458	1.00	
F	0.208	1	0.208	PxF	2.223	1.00	
AxB	10.800	1	10.800	P	9.135	1.18	
AxC	12.838	2	6.419	P	9.135	1.00	
AxD	19.200	1	19.200	P	9.135	2.10	
AxE	4.800	1	4.800	PxE	2.458	1.95	
AxF	0.209	1	0.209	PxF	2.223	1.00	
BxC	12.280	2	6.140	P	9.135	1.00	
BxD	9.076	1	9.076	P	9.135	1.00	
BxE	16.876	1	16.876	PxE	2.458	6.87	.025
BxF	4.034	1	4.034	PxF	2.223	1.81	
CxD	27.155	2	13.578	P	9.135	1.49	
CxE	2.255	2	1.122	PxE	2.458	1.00	
CxF	1.755	2	0.878	PxF	2.223	1.00	
DxE	1.409	1	1.409	PxE	2.458	1.00	

TABLE A. MAIN ANALYSIS OF VARIANCE (Continued)

Source	Sum of Squares	d. f.	Mean Square	Error Term	F - Ratio	P - Value
DxF	9.634	1	9.634	PxF 2.223	4.33	.05
ExF	0.534	1	0.534	PxExF 1.671	1.00	
AxBxC	1.288	2	0.644	P 9.135	1.00	
AxBxD	1.408	1	1.408	P 9.135	1.00	
AxCxD	2.112	2	1.056	P 9.135	1.00	
BxCxD	16.588	2	8.294	P 9.135	1.00	
AxBxE	0.208	1	0.208	PxE 2.458	1.00	
AxCxE	16.462	2	8.231	PxE 2.458	3.35	.05
BxCxE	4.588	2	2.294	PxE 2.458	1.00	
AxDxE	0.075	1	0.075	PxE 2.458	1.00	
BxDxE	0.533	1	0.533	PxE 2.458	1.00	
CxDxE	2.279	2	1.140	PxE 2.458	1.00	
ABF	2.700	1	2.700	PxF 2.223	1.21	
ACF	1.079	2	0.540	PxF 2.223	1.00	
BCF	2.004	2	1.002	PxF 2.223	1.00	
ADF	0.133	1	0.133	PxF 2.223	1.00	
BDF	0.075	1	0.075	PxF 2.223	1.00	
CDF	1.379	2	0.690	PxF 2.223	1.00	

TABLE A. MAIN ANALYSIS OF VARIANCE (Continued)

Source	Sum of Squares	d.f.	Mean Square	Error Term	F - Ratio	P - Value
AEF	0.033	1	0.033	PxExF 1.671	1.00	
BEF	0.008	1	0.008	PxExF 1.671	1.00	
CEF	0.129	2	0.064	PxExF 1.671	1.00	
DEF	1.875	1	1.875	PxExF 1.671	1.12	
ABCD	1.979	2	0.990	P 9.135	1.00	
ABCE	2.329	2	1.164	PxE 2.458	1.00	
ABDE	7.500	1	7.500	PxE 2.458	3.05	.10
ACDE	3.238	2	1.619	PxE 2.458	1.00	
BCDE	6.926	2	3.463	PxE 2.458	1.41	
ABCF	1.462	2	0.731	PxF 2.223	1.00	
ABDF	0.075	1	0.075	PxF 2.223	1.00	
ACDF	0.954	2	0.477	PxF 2.223	1.00	
BCDF	4.512	2	2.256	PxF 2.223	1.01	
ABEF	7.008	1	7.008	PxExF 1.671	4.19	.05
ACEF	5.954	2	2.977	PxExF 1.671	1.78	.20
BCEF	1.129	2	0.564	PxExF 1.671	1.00	
ADEF	0.008	1	0.008	PxExF 1.671	1.00	
BDEF	0.033	1	0.033	PxExF 1.671	1.00	
CDEF	5.488	2	2.744	PxExF 1.671	1.64	

TABLE A. MAIN ANALYSIS OF VARIANCE (Continued)

Source	Sum of Squares	d.f.	Mean Square	Error Term	F - Ratio	P - Value
ABCDE	0.388	2	0.194	PxE 2.458	1.00	
ABCDF	0.388	2	0.194	PxF 2.223	1.00	
ABCEF	1.954	2	0.977	PxExF 1.671	1.00	
ABDEF	2.700	1	2.700	PxExF 1.671	1.62	
ACDEF	1.029	2	0.514	PxExF 1.671	1.00	
BCDEF	0.104	2	0.052	PxExF 1.671	1.00	
ABCDEF	9.612	2	4.806	PxExF 1.671	2.88	.10
Pupils (P)						
within ABCD	877.000	96	9.135	PxExF 1.671	5.47	.001
PxE(within ABCD)	236.000	96	2.458	PxExF 1.671	1.47	.05
PxF(within ABCD)	213.400	96	2.223	PxExF 1.671	1.33	.10
Residual(PEF) within ABCD	160.400	96	1.671			
Total	441.992	479				
<u>Summary</u>						
1. Total sum of squares	1928.792	479				
2. between cells s.s.	441.992	95				
3. Within cells s.s.	1486.800	384				

TABLE B. ANALYSIS OF COVARIANCE. TEST IA. (OBJECTIVE: Understanding
OCCASION: One)

Source	Sum of Squares	d.f.	Mean Square	F - Ratio	P - Value
A. Teachers	1.453	1	1.453	1.00	
B. Methods	55.318	1	55.318	19.46	.001
C. Cognitive Style	1.956	2	0.978	1.00	
D. Sex	18.007	1	18.007	6.34	.025
A.B.	0.994	1	0.994	1.00	
A.C.	15.430	2	7.715	2.71	.10
A.D.	2.770	1	2.770	1.00	
B.C.	12.461	2	6.230	2.19	.20
B.D.	0.000	1	0.000	1.00	
C.D.	8.543	2	4.271	1.50	
A.B.C.	2.077	2	1.038	1.00	
A.B.D.	1.229	1	1.229	1.00	
A.C.D.	0.872	2	0.436	1.00	
B.C.D.	3.327	2	1.664	1.00	
A.B.C.D.	2.236	2	1.118	1.00	
Error	270.030	95	2.842		
Total		118			

TABLE C. ANALYSIS OF COVARIANCE. TEST IIA. (OBJECTIVE: Application
OCCASION: One)

Source	Sum of Squares	d.f.	Mean Square	F - Ratio	P - Value
A. Teachers	16.165	1	16.165	6.96	.01
B. Methods	7.964	1	7.964	3.43	.10
C. Cognitive Style	5.805	2	2.902	1.25	
D. Sex	52.923	1	52.923	22.78	.001
A.B.	2.173	1	2.173	1.00	
A.C.	7.483	2	3.742	1.61	
A.D.	1.916	1	1.916	1.00	
B.C.	4.244	2	2.122	1.00	
B.D.	0.026	1	0.026	1.00	
C.D.	1.482	2	0.741	1.00	
A.B.C.	2.067	2	1.034	1.00	
A.B.D.	0.048	1	1.048	1.00	
A.C.D.	1.880	2	0.940	1.00	
B.C.D.	3.872	2	1.936	1.00	
A.B.C.D.	4.136	2	2.068	1.00	
Error	220.647	95	2.323		
Total		118			

TABLE D. ANALYSIS OF COVARIANCE. TEST IB. (OBJECTIVE: Understanding
OCCASION: Two)

Source	Sum of Squares	d.f.	Mean Square	F - Ratio	P - Value
A. Teacher	1.599	1	1.599	1.00	
B. Methods	25.379	1	25.379	9.56	.01
C. Cognitive Style	2.032	2	1.016	1.00	
D. Sex	8.387	1	8.387	3.16	.10
A.B.	0.043	1	0.043	1.00	
A.C.	6.379	2	3.190	1.20	
A.D.	4.186	1	4.186	1.58	
B.C.	9.296	2	4.648	1.75	.20
B.D.	1.008	1	1.008	1.00	
C.D.	2.730	2	1.365	1.00	
A.B.C.	1.511	2	0.756	1.00	
A.B.D.	7.094	1	7.094	2.67	.20
A.C.D.	0.655	2	0.328	1.00	
B.C.D.	1.827	2	0.914	1.00	
A.B.C.D.	6.691	2	3.346	1.26	
Error	252.162	95	2.654		
Total		118			

TABLE E. ANALYSIS OF COVARIANCE. TEST IIB. (OBJECTIVE: Application
OCCASION: Two)

Source	Sum of Squares	d.f.	Mean Square	F - Ratio	P - Value
A. Teachers	9.851	1	9.851	3.110	.10
B. Methods	2.378	1	2.378	1.00	
C. Cognitive Style	6.009	2	3.004	1.00	
D. Sex	6.753	1	6.753	2.132	.20
A.B.	8.697	1	8.697	2.75	.20
A.C.	1.617	2	0.808	1.00	
A.D.	2.927	1	2.927	1.00	
B.C.	5.139	2	2.570	1.00	
B.D.	0.070	1	0.070	1.00	
C.D.	3.591	2	1.796	1.00	
A.B.C.	1.415	2	0.708	1.00	
A.B.D.	3.032	1	3.032	1.00	
A.C.D.	7.464	2	3.732	1.18	
B.C.D.	11.340	2	5.670	1.80	
A.B.C.D.	3.198	2	1.599	1.00	
Error	300.980	95	3.168		
Total		118			

TABLE F. ANALYSIS OF VARIANCE. TEST IA. (OBJECTIVE: Understanding
OCCASION: One)

Source	Sum of Squares	d.f.	Mean Square	F - Ratio	P - Value
A. Teachers	0.075	1	0.075	1.00	
B. Methods	81.675	1	81.675	22.43	.001
C. Cognitive Style	4.067	2	2.034	1.00	
D. Sex	7.067	1	7.067	1.90	.20
A.B.	3.675	1	3.675	1.01	
A.C.	18.200	2	9.100	2.50	.10
A.D.	4.350	1	4.350	1.19	
B.C.	6.200	2	3.100	1.00	
B.D.	1.817	1	1.817	1.00	
C.D.	14.408	2	7.204	1.91	.20
A.B.C.	2.600	2	1.300	1.00	
A.B.D.	0.066	1	0.066	1.00	
A.C.D.	2.525	2	1.262	1.00	
B.C.D.	1.858	2	0.929	1.00	
A.B.C.D.	4.809	2	2.404	1.00	
Error	349.600	96	3.642		
Total		119			

TABLE G. ANALYSIS OF VARIANCE. TEST IIA (OBJECTIVE: Application
OCCASION: One)

Source	Sum of Squares	d.f.	Mean Square	F - Ratio	P - Value
A. Teachers	4.403	1	4.408	1.18	
B. Methods	23.408	1	23.408	6.25	.025
C. Cognitive Style	9.517	2	4.758	1.27	
D. Sex	27.075	1	27.075	7.23	.01
A.B.	0.076	1	0.076	1.00	
A.C.	6.117	2	3.058	1.00	
A.D.	3.725	1	3.725	1.00	
B.C.	3.317	2	1.658	1.00	
B.D.	3.675	1	3.675	1.00	
C.D.	8.750	2	4.375	1.17	
A.B.C.	2.449	2	1.224	1.00	
A.B.D.	0.958	1	0.958	1.00	
A.C.D.	0.100	2	0.050	1.00	
B.C.D.	6.450	2	3.225	1.00	
A.B.C.D.	0.367	2	0.184	1.00	
Error	359.608	96	3.746		
Total		119			

TABLE H. ANALYSIS OF VARIANCE. TEST IB. (OBJECTIVE: Understanding
OCCASION: Two)

Source	Sum of Squares	d.f.	Mean Square	F - Ratio	P - Value
A. Teachers	0.300	1	0.300	1.00	
B. Methods	48.133	1	48.133	12.45	.01
C. Cognitive Style	4.317	2	2.158	1.00	
D. Sex	0.833	1	0.833	1.00	
A.B.	0.834	1	0.834	1.00	
A.C.	9.150	2	4.575	1.18	
A.D.	6.534	1	6.534	1.69	.20
B.C.	5.617	2	2.808	1.00	
B.D.	0.834	1	0.834	1.00	
C.D.	10.617	2	5.308	1.37	
A.B.C.	1.116	2	0.558	1.00	
A.B.D.	2.132	1	2.132	1.00	
A.C.D.	2.116	2	1.058	1.00	
B.C.D.	3.516	2	1.758	1.00	
A.B.C.D.	2.218	2	1.109	1.00	
Error	371.200	96	3.867		
Total		119			

TABLE I. ANALYSIS OF VARIANCE. TEST IIB. (OBJECTIVE: Application
OCCASION: Two)

Source	Sum of Squares	d.f.	Mean Square	F - Ratio	P - Value
A. Teachers	2.133	1	2.133	1.00	
B. Methods	8.533	1	8.533	2.02	.20
C. Cognitive Style	10.067	2	5.034	1.19	
D. Sex	0.533	1	0.533	1.00	
A.B.	16.134	1	16.134	3.81	.10
A.C.	2.867	2	1.434	1.00	
A.D.	4.801	1	4.801	1.13	
B.C.	4.867	2	2.434	1.00	
B.D.	3.333	1	3.333	1.00	
C.D.	2.467	2	1.234	1.00	
A.B.C.	1.000	2	0.500	1.00	
A.B.D.	8.533	1	8.533	2.02	.20
A.C.D.	2.599	2	1.300	1.00	
B.C.D.	16.467	2	8.234	1.95	
A.B.C.D.	4.733	2	2.366	1.00	
Error	406.400	96	4.233		
Total		119			

TABLE J. MEANS AND ADJUSTED MEANS FOR SELECTED FACTORS

Design Source	Dependent Variable	Factor	Unadjusted Means			Adjusted Means		
			R.E.	R.D.	Total	R.E.	R.D.	Total
Analysis of Covariance	Test IA	Treatments	6.17	4.52	5.34	6.03	4.65	5.34
	Test IIA	Treatments	5.68	4.80	5.24	5.50	4.98	5.24
	Test IB	Treatments	5.87	4.60	5.23	5.70	4.77	5.23
	Test IIB	Treatments	5.53	5.00	5.27	5.38	5.16	5.27
			Boys	Girls	Total	Boys	Girls	Total
Analysis of Covariance	Test IA	Sex	5.58	5.10	5.34	5.64	5.05	5.34
	Test IIA	Sex	5.72	4.77	5.24	5.92	4.57	5.24
	Test IB	Sex	5.32	5.15	5.23	5.50	4.96	5.23
	Test IIB	Sex	5.33	5.20	5.27	5.51	5.02	5.27
			Level I	Level II	Level III			
Main Analysis of Variance	Achievement in Science as a whole	Teachers	5.33		5.21		-	
		Treatments	5.81		4.73		-	
		Styles	5.56		5.24		5.01	
		Sex	5.49		5.05		-	
		Objectives	5.29		5.25		-	
		Occasions	5.29		5.25		-	

APPENDIX D.12TABLE K. TREATMENT MEANS AND OTHER DATA

Statistics	Test IA	Test IIA	Test IB	Test IIB
R.E. Mean	6.166	5.683	5.866	5.533
R.D. Mean	4.516	4.800	4.600	5.000
Difference	1.650	.883	1.266	.533
S.E. _D	.596	.589	.587	.605
Total Mean	5.341	5.242	5.233	5.266
Total S.D.	2.056	1.966	1.986	2.040
Raw Score Total Mean	12.94	10.50	12.881	11.38
N	120	120	120	120

APPENDIX E: WORK MATRICES FOR TEACHERS AND
TREATMENTS

The Matrices show the percentages in each category for each teacher within each treatment, together with the major teaching moves for each teacher in each treatment.

WORK MATRIX

TEACHER ONE: TREATMENT R.F.

	1	2	3	4	5	6	7	8	9	10	
1					1						
2	1		8	9	8		2		5	1	
3		1	4	5	2	1	1		1	2	
4				8	5	10		19	19	7	
5				20	287	2			5	20	
6				4	4	5			1	5	
7				1	1						
8		14	1	1	2				1		
9		18	4	4	2		3		35	4	
10	1			9	22	2	3			57	Matrix Total
TOTAL	2	33	17	61	334	20	9	19	67	96	658
%	0.03	5.06	2.61	9.36	53.44	3.06	1.38	2.91	10.28	14.72	

WORK MATRIX

TEACHER TWO: TREATMENT R.E.

	1	2	3	4	5	6	7	8	9	10	
1	2				1						
2			8	6		1					
3			3	3	4			2			
4		1		15	13			14	4	9	
5				17	273	2				25	
6				4	1					1	
7											
8		9	1		1			2		2	
9		5							5	3	
10	1			10	24	1				30	Matrix Total
TOTAL	3	15	12	55	317	4	0	18	10	67	501
%	0.60	2.99	2.39	10.98	63.27	0.80	0.0	3.59	2.00	13.37	

WORK MATRIX

TEACHER ONE: TREATMENT R.D.

	1	2	3	4	5	6	7	8	9	10	
1	2			7						1	
2	2	2	6	29	5	4		7	3	4	
3	1	1	7	4	2	1		1	1	1	
4	1	1	1	78	4	2		38	17	15	
5				8	26			2		3	
6				5		37		3	1	11	
7										1	
8	1	41	4	8			2	24		1	
9	1	15	2	5					53		
10	2	2		16	1	12		3	2	78	Matrix Total
TOTAL	10	63	20	160	38	56	2	78	77	115	619
%	1.62	10.18	3.23	25.85	6.14	9.05	0.31	12.60	12.44	18.58	

WORK MATRIX

TEACHER TWO: TREATMENT R.D.

	1	2	3	4	5	6	7	8	9	10		
1	2			2	1			1				
2	1	1	9	22	2	5			4	3		
3			2	6	2	2			1	1		
4	1			51	7	7		20	18	14		
5				12	35	4			1	3		
6	1	1		9	2	46	1		2	23		
7				3		1	1			2		
8		18		3			2	2		2		
9	1	22	3	1		3	2		49	4		
10	2	2		18	7	10		2	2	132		Matrix Total
TOTAL	8	44	14	127	56	78	6	25	77	184		619
%	1.29	7.11	2.26	20.52	9.05	12.60	0.97	4.04	12.44	29.73		

BIBLIOGRAPHY

- Allman, J. and Rokeach, M. (1967). A note on the use of paper-pencil items to probe cognitive and affective processes. Educational and Psychological Measurement. 27: 127-133.
- Amidon, E. and Amidon, Peggy. (1967). Teaching Pattern Analysis Manual. Minneapolis, Minnesota: Association for Productive Teaching, Inc.
- Amidon, E. and Flanders, N.A. (1967). The Role of the Teacher in the Classroom. Minneapolis, Minnesota: Association for Productive Teaching.
- Anderson, R.C. (1959). Learning in discussions: resumé of the authoritarian-democratic studies. Harvard Educational Review. 29: 201-215.
- Anderson, R.C., Faust, G.W. and Roderick, Marianne C. (1968). Overprompting in programmed instruction. Journal of Educational Psychology. 59: 88-93.
- Ausubel, D.P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. Journal of Educational Psychology. 51: 267-272.
- Ausubel, D.P. (1961). Learning by discovery: rationale and mystique. Bulletin of the National Association of Secondary-School Principals. 45: 18-58.
- Ausubel, D.P. (1962). A subsumption theory of meaningful verbal learning and retention. Journal of General Psychology. 66: 312-324.
- Ausubel, D.P. (1963a). The Psychology of Meaningful Verbal Learning. New York: Grune and Stratton.
- Ausubel, D.P. (1963b). Cognitive structure and the facilitation of meaningful verbal learning. Journal of Teacher Education. 14: 217-222.
- Ausubel, D.P. (1964a). Adults versus children in second-language learning: psychological considerations. Modern Languages Journal. 48: 420-424.

- Ausubel, D.P. (1964b). Some psychological and educational limitations of learning by discovery. The Arithmetic Teacher. 11: 290-302.
- Ausubel, D.P. (1965). Perception versus cognition in meaningful reception learning. Journal of General Psychology. 73: 185-187.
- Ausubel, D.P. (1967a). A cognitive-structure theory of school learning. Pp. 207-260 in Siegel, L. (ed.), Instruction: Some Contemporary Viewpoints. San Francisco: Chandler.
- Ausubel, D.P. (1967b). Learning Theory and Classroom Practice. Bulletin No. 1/1967: Ontario Institute for Studies in Education.
- Ausubel, D.P. (1968). Educational Psychology: A Cognitive Viewpoint. New York: Holt, Rinehart and Winston.
- Ausubel, D.P. and Ausubel, Pearl. (1966). Cognitive development in adolescence. Review of Educational Research. 36: 403-413.
- Ausubel, D.P. and Blake, E., Jr. (1958). Proactive inhibition in the forgetting of meaningful school material. Journal of Educational Psychology. 52: 145-149.
- Ausubel, D.P. and Fitzgerald, D. (1962). Organizer, general background, and antecedent learning variables in sequential verbal learning. Journal of Educational Psychology. 53: 243-249.
- Ausubel, D.P. and Fitzgerald, D. (1961). The role of discriminability in meaningful verbal learning and retention. Journal of Educational Psychology. 52: 266-274.
- Ausubel, D.P., Robbins, L.C. and Blake, E. (1957). Retroactive inhibition and facilitation in the learning of school materials. Journal of Educational Psychology. 48: 334-343.

- Ausubel, D.P., Stager, Mary and Gaité, A.J.H. (1968). Retroactive facilitation in meaningful verbal Learning. Journal of Educational Psychology. 59: 250-255.
- Ausubel, D.P., Stager, Mary and Gaité, A.J.H. (1969). Proactive effects in meaningful verbal learning and retention. Journal of Educational Psychology. 60: 59-64.
- Ausubel, D.P. and Youssef, M. (1963). Role of discriminability in meaningful parallel learning. Journal of Educational Psychology. 54: 331-336.
- Baddeley, A.D. (1963). A Zeigarnik-like effect in the recall of anagrams. Quarterly Journal of Experimental Psychology. 15: 63-64.
- Baldwin, A.L. (1967). Theories of Child Development. New York: Wiley.
- Bandura, A. and Walters, R.H. (1963). Social Learning and Personality Development. New York: Holt, Rinehart and Winston.
- Berlyne, D.E. (1965). Structure and Direction in Thinking. New York: Wiley.
- Bibergall, Jo Anne. (1966). Learning by discovery: its relation to science teaching. Educational Review. 18: 222-231.
- Biddle, B.J. and Adams, R.S. (1967). Teacher behaviour in the classroom context. Pp. 99-141 in Siegel, L. (ed.), Instruction: Some Contemporary Viewpoints. San Francisco: Chandler.
- Bing, Elizabeth. (1963). Effect of childrearing practices on development of differential cognitive abilities. Child Development. 34: 631-648.
- Blum, A. and Adcock, Carolyn. (1968). Attention and early learning: a selected review. Boston Journal of Education. 150: 28-40.

- Blum, A.H. and Broverman, D.M. (1967). Children's cognitive style and response modification. Journal of Genetic Psychology. 110: 95-103.
- Bruner, J.S. (1961). The act of discovery. Harvard Educational Review. 31: 21-32.
- Bruner, J.S. (1966). Toward a Theory of Instruction. New York: Norton.
- Bruner, J.S., Goodnow, Jacqueline and Austin, G.A. (1956). A study of Thinking. New York: Wiley.
- Bruner, J.S. and Olver, Rose R. (1963). Development of equivalence transformations in children. In Wright, J.C. and Kagan, J. (eds.), Basic cognitive processes in children. Monograph of Society for Research in Child Development. 28 (Whole No.86). 125-141.
- Bruner, J.S. and Tajfel, H. (1961). Cognitive risk and environmental change. Journal of Abnormal Psychology. 62: 231-241.
- Bugelski, B.R. (1964). The Psychology of Learning Applied to Teaching. Indianapolis: Bobbs-Merrill.
- Campbell, D.T. and Stanley, J.C. (1963). Experimental and quasi-experimental designs for research on teaching. Pp.171-246 in Gage, N.L. (ed.), Handbook of Research on Teaching. Chicago: Rand McNally.
- Carroll, J.B. (1963). A model of school learning. Teachers College Record. 64: 723-733.
- Carroll, J.B. (1964). Words, meanings and concepts. Harvard Educational Review. 34: 178-202.
- Carroll, J.B. (1965). School learning over the long haul. Pp. 249-269 in Krumboltz, J.D. (ed.), Learning and the Educational Process. Chicago: Rand McNally.
- Carroll, J.B. (1965). School learning over the long haul. Pp.249-269 in Krumboltz, J.D. (ed.), Learning and the Educational Process. Chicago: Rand McNally.

- Carin, A. and Sund, R.B. (1964). Teaching Science through Discovery. Columbus, Ohio: Merrill.
- Cattell, R.B. (ed.), (1966). Handbook of Multivariate Experimental Psychology. Chicago: Rand McNally.
- Cochran, W.G. and Cox, G.M. (1957). Experimental Design. New York: Wiley.
- Collis, K.F. (1969). The influence of a teaching theory on mathematics achievement and attitude. Australian Journal of Education. 13: 12-31.
- Connaughton, I.M. and Skurnik, L.S. (1969). The comparative effectiveness of several short-cut item analysis procedures. British Journal of Educational Psychology. 3: 225-232.
- Corman, B.R. (1957). The effect of varying amounts and kinds of information as guidance in problem solving. Psychological Monograph. 71. No.2. (Whole No.431).
- Cox, D.R. (1958). Planning of Experiments. New York: Wiley.
- Craig, R.C. (1956). Directed versus independent discovery of established relations. Journal of Educational Psychology. 47: 223-234.
- Cronbach, L.J. (1965). Issues current in educational psychology. Pp. 109-126 in Morrisett, L.N. and Vinsonhaler, J.F. Mathematical learning. Monograph of the Society for Research in Child Development. 30. No.1. (Serial No.99).
- Cronbach, L.J. (1966). The logic of experiments on discovery. Pp. 77-92 in Shulman, L.S. and Keislar, E.R. (eds.), Learning by Discovery: A Critical Appraisal. Chicago: Rand McNally.
- Cronbach, L.J. (1967). How can instruction be adapted to individual differences? Pp. 23-44 in Gagné, R.M. (ed.), Learning and Individual Differences. Columbus, Ohio: Merrill.

- Cronbach, L.J. (1968). A parsimonious reinterpretation of the Wallach-Kogan data. American Educational Research Journal. 5: 491-511.
- Cropley, A.J. (1969). Creativity, intelligence and intellectual style. Australian Journal of Education. 13: 3-7.
- Cunningham, J.D. (1967). Interaction analysis: a useful technique for research and science supervision. Science Education. 51: 27-32.
- Dearden, R.F. (1967). Instruction and learning by discovery. Pp. 135-155 in Peters, R.S. (ed.), The Concept of Education. London: Routledge and Kegan Paul.
- De Cecco, J.P. (1968). The Psychology of Learning and Instruction: Educational Psychology. Englewood Cliffs: Prentice Hall.
- Dickstein, Lois S. (1968). Field independence in concept attainment. Perceptual and Motor Skills. 27: 635-642.
- Duncan, C.P. (1956). Recent research on human problem solving. Psychological Bulletin. 56: 397-429.
- Dunham, J.L. and Bunderson, C.V. (1969). Effect of decision-rule instruction upon the relationship of cognitive abilities to performance in multiple-category concept problems. Journal of Educational Psychology. 60: 121-125.
- Durost, W. (1968). The characteristics, use and computation of stanines. Pp. 132-133 in Gronlund, N.E. (ed.), Readings in Measurement and Evaluation. New York: MacMillan.
- Edwards, A.L. (1957). Techniques of Attitude Scale Construction. New York: Appleton, Century, Crofts.

- Edwards, A.L. (1968). Experimental Design in Psychological Research. New York: Holt, Rinehart and Winston.
- Edwards, J.C. (1969). Effect of instruction and concomitant variables on multiple-categorization ability. Journal of Educational Psychology. 60: 138-143.
- Elley, W.B. (1969). Changes in mental ability in New Zealand school children, 1936-1969. New Zealand Journal of Educational Studies. 4: 140-150.
- Entwistle, N.J. (1968). Academic motivation and school attainment. British Journal of Educational Psychology. 38: 181-188.
- Ervin, Susan M. (1960). Transfer effects of learning a verbal generalization. Child Development. 31: 537-554.
- Festinger, L. (1957). A Theory of Cognitive Dissonance. Evanston, Illinois: Row, Peterson.
- Fisher, S.T., Weiss, D.J. and Dawis, R.V. (1968). A comparison of Likert and pair comparisons techniques in multivariate attitude scaling. Educational and Psychological Measurement. 28: 81-94.
- Forgus, R.H. and Schwartz, R.J. (1957). Efficient retention and transfer as affected by learning method. Journal of Psychology. 43: 135-139.
- Frandsen, A.N. (1961). Educational Psychology: The Principles of Learning in Teaching. New York: McGraw-Hill.
- Frase, L.T. (1969). Paragraph organization of written materials: the influence of conceptual clustering upon the level and organization of recall. Journal of Educational Psychology. 60: 394-401.

- Gage, N.L. (ed.), (1963). Handbook of Research on Teaching. Chicago: Rand McNally.
- Gage, N.L. (1964). Theories of teaching. Pp. 268-285 in Hilgard, E.R. (ed.), Theories of Learning and Instruction. Part I of the 63rd Yrbk of the National Society for the Study of Education. Chicago: University of Chicago Press.
- Gage, N.L. (1964). Towards a cognitive theory of teaching. Teachers College Record. 65: 408-412.
- Gage, N.L. (1966). Research on cognitive aspects of teaching. Pp. 29-44 of the Report of the Seminar on Teaching, "The Way Teaching Is." Association for Supervision and Curriculum Development and the Center for the Study of Instruction. Washington, D.C.: National Education Association.
- Gage, N.L. (1968). An analytical approach to research on instructional methods. Journal of Experimental Education. 37: 119-125.
- Gage, N.L. and Unruh, W.R. (1967). Theoretical formulations for research on teaching. Review of Educational Research. 37: 358-370.
- Gagné, R.M. (1962a). Military training and principles of learning. American Psychologist. 17: 83-91.
- Gagné, R.M. (1962b). The acquisition of knowledge. Psychological Review. 69: 355-365.
- Gagné, R.M. (1965a). The Conditions of Learning. New York: Holt, Rinehart and Winston.
- Gagné, R.M. (1965b). Educational objectives and human performance. Pp.1-24 in Krumboltz, J.D. (ed.), Learning and the Educational Process. Chicago: Rand McNally.
- Gagné, R.M. (1966a). Varieties of learning and the concept of discovery. Pp.135-150 in Shulman, L.S. and Keislar, E.R. (eds.), Learning by Discovery: A Critical Appraisal. Chicago: Rand McNally.

- Gagné, R.M. (1966b). The learning of principles. Pp.81-96 in Klausmeier, H.J. and Harris, C.W. (eds.), Analyses of Concept Learning. New York: Academic Press.
- Gagné, R.M. (ed.), (1967). Learning and Individual Differences. Columbus, Ohio: Charles E. Merrill.
- Gagné, R.M. (1968). Contributions of learning to human development. Psychological Review. 75: 177-191.
- Gagné, R.M. (1969a). Review of Ausubel's Educational Psychology, 1968. American Educational Research Journal. 6: 287-290.
- Gagné, R.M. (1969b). Context, isolation and interference effects on the retention of fact. Journal of Educational Psychology. 60: 408-414.
- Gagné, R.M. and Brown, L.T. (1961). Some factors in the programming of conceptual learning. Journal of Experimental Psychology. 62: 313-321.
- Gagné, R.M. and Dick, W. (1962). Learning measures in a self-instructional program in solving equations. Psychological Reports. 10: 131-146.
- Gagné, R.M. and Smith, E.C., Jr. (1962). A study of the effects of verbalization on problem solving. Journal of Experimental Psychology. 63: 12-18.
- Gagné, R.M. and Wiegand, Virginia K. (1968). Learning and retention of concrete rules. Journal of Educational Psychology. 59: 355-361.
- Gallagher, J.J. and Jenne, W.C. (1967). The relationship of cognitive style to classroom expressiveness and associated variables. Journal of Educational Research. 60: 273-279.
- Gardner, R.W. (1964). The development of cognitive structures. In Scheerer, C. (ed.), Cognition: Theory, Research, Promise. New York: Harper and Row.

- Gardner, R.W., Holzman, P.S., Klein, G.S., Linton, H.B. and Spence, D.P. (1959). Cognitive control: a study of individual consistencies in cognitive behaviour. Psychological Issues. 1. No.4.
- Gardner, R.W. and Long, R.I. (1962). Cognitive controls of attention and inhibition: a study of individual consistencies. British Journal of Psychology. 53: 381-388.
- Garrett, D. (1969). Some Personality Theories and Personality Variables in Adult Learning. Massey University: Unpublished Monograph.
- Garrett, H.E. (1958). Statistics in Psychology and Education. New York: McKay.
- Glaser, R. (1966). Variables in discovery learning. Pp. 13-26 in Shulman, L.S. and Keislar, E.R. (eds.), Learning by Discovery: A Critical Appraisal. Chicago: Rand McNally.
- Glick, O. (1968). The educational process in the classroom. School Review. 76: 339-351.
- Gordon, I.J. (1966). Studying the Child in School. New York: Wiley.
- Gordon, L.V. (1967). Validity of scoring methods for bipolar scales. Educational and Psychological Measurement. 27: 1099-1106.
- Guilford, J.P. (1950). Fundamental Statistics in Psychology and Education. New York: McGraw-Hill.
- Guthrie, J.T. (1967). Expository versus a discovery method. Journal of Educational Psychology. 58: 45-49.
- Harvey, O.J., Hunt, D.E. and Schroder, H.M. (1961). Conceptual Systems and Personality Organization. New York: Wiley.
- Haslerud, G.M. and Meyers, Shirley. (1958). The transfer value of given and individually derived principles. Journal of Educational Psychology. 49: 293-298.

- Heil, L.M. and Powell, Marion. (1961). Effect of Certain Kinds of Teachers on Children's Feelings. Office of Testing and Research, Brooklyn College: Mimeo'd.
- Heil, L.M., Powell, Marion and Feifer, I. (1960). Characteristics of Teacher Behaviour Related to the Achievement of Children in Several Elementary Grades. Washington: U.S. Department of Health, Education and Welfare, Office of Education, Cooperative Research Branch.
- Hendrix, Gertrude. (1947). A new clue to transfer of training. Elementary School Journal. 48: 197-208.
- Hendrix, Gertrude. (1961). Learning by discovery. Mathematics Teacher. 43: 334-339.
- Hess, R.D. and Shipman, Virginia. (1965). Early experience and the socialization of cognitive modes in children. Child Development. 36: 869-886.
- Hilgard, E.R. (1964). A perspective on the relationship between learning theory and educational practices. Pp. 402-415 in Hilgard, E.R. (ed.), Theories of Learning and Instruction. Part I of the 63rd Yrbk of the National Society for the Study of Education. Chicago: University of Chicago Press.
- Hilgard, E.R. and Bower, G.H. (1966). Theories of Learning. 3rd ed. New York: Appleton, Century, Crofts.
- Hilgard, E.R., Edgren, R.D. and Irvine, R.P. (1954). Errors in transfer following learning with understanding: further studies with Katona's card-trick experiments. Journal of Experimental Psychology. 47: 457-464.
- Hilgard, E.R., Irvine, R.P. and Whipple, J.E. (1953). Rote memorization, understanding, and transfer: an extension of Katona's card-trick experiments. Journal of Experimental Psychology. 46: 288-292.

- Hill, W.F. (1963). Learnings: A Survey of Psychological Interpretations. San Francisco: Chandler.
- Hill, W.F. (1963). What can the psychology of learning offer education? A review of some recent opinions from psychologists. Journal of Teacher Education. 14: 443-448.
- Hirschfield, P.P. (1965). Response set in impulsive children. Journal of Genetic Psychology. 107: 117-126.
- Holt, J. (1964). How Children Fail. New York: Pitman.
- Holtzman, W.H. (1966). Intelligence, cognitive style and personality: a developmental approach. Pp.1-32 in Brim, O.G. (Jnr.), Crutchfield, R.S. and Holtzman, W.H. Intelligence Perspectives 1965: The Terman-Otis Memorial Lectures. New York: Harcourt, Brace and World.
- Hunt, J.McV. (1960). Experience and the development of motivation: some reinterpretations. Child Development. 31: 489-504.
- Hunt, J.McV. (1961). Intelligence and Experience. New York: Ronald.
- Jackson, P.W. (1966). The way teaching is. Pp.7-27 of The Way Teaching Is. Report of the Seminar on Teaching. Association for Supervision and Curriculum Development and the Center for the Study of Instruction. Washington, D.C.: National Education Association.
- Jacobson, L.I., Dickinson, T.C., Fleishman, Joyce, M. and Haraguthi, Rosemary Stone. (1969). Relationship of intelligence and mediating processes to concept learning. Journal of Educational Psychology. 60: 109-112.
- Johnson, D.M. and Stratton, R.P. (1966). Evaluation of five methods of teaching concepts. Journal of Educational Psychology. 57: 48-53.

- Johnson, P.E. (1969). On the communication of concepts in science. Journal of Educational Psychology. 60: 32-40.
- Kagan, J. (1965a). Reflection-impulsivity and reading ability in primary grade children. Child Development. 36: 609-628.
- Kagan, J. (1965b). Impulsive and reflective children: the significance of conceptual tempo. Pp. 133-161 in Krumboltz, J.D. (ed.), Learning and the Educational Process. Chicago: Rand McNally.
- Kagan, J. (1965c). Personality and the learning process. Daedalus. 94: 553-563.
- Kagan, J. (1966a). Learning, attention and the issue of discovery. Pp. 151-161 in Shulman, L.S. and Keislar, E.R. (eds.), Learning by Discovery: A Critical Appraisal. Chicago: Rand McNally.
- Kagan, J. (1966b). Reflection-impulsivity: the generality and dynamics of conceptual tempo. Journal of Abnormal Psychology. 71: 17-24.
- Kagan, J. and Garn, S.M. (1963). A constitutional correlate of early intellectual functioning. Journal of Genetic Psychology. 102: 83-89.
- Kagan, J. and Lewis, M. (1965). Studies of attention in the human infant. Merrill-Palmer Quarterly. 11: 95-127.
- Kagan, J. and Moss, H.A. (1962). Birth to Maturity. New York: Wiley.
- Kagan, J., Moss, A.H. and Sigel, I.E. (1963). Psychological significance of styles of conceptualization. Pp. 73-124 in Wright, J.C. and Kagan, J. (eds.), Basic cognitive processes in children. Monograph of the Society for Research in Child Development. 28. No.2. (Serial No.86).

- Kagan, J., Pearson, L. and Welch, Lois. (1966a).
Conceptual impulsivity and inductive reasoning.
Child Development. 37: 582-594.
- Kagan, J., Pearson, L. and Welch Lois. (1966b).
Modifiability of an impulsive tempo. Journal of
Educational Psychology. 57: 359-365.
- Kagan, J. and Rosman, Bernice L. (1964). Cardiac and
respiratory correlates of attention and an
analytic attitude. Journal of Experimental Child
Psychology. 1: 50-63.
- Kagan, J., Rosman, Bernice L., Day, Deborah, Albert, J. and
Phillips, W. (1964). Information processing in the
child: significance of analytic and reflective
attitudes. Psychological Monograph. 78, No.1.
(Whole No.578).
- Kaplan, A. (1964). The Conduct of Inquiry. San Francisco:
Chandler.
- Karp, S.A. (1963). Field dependence and overcoming
embeddedness. Journal of Consulting Psychology.
27: 294-302.
- Katona, G. (1940). Organizing and Memorizing. New York:
Columbia University Press.
- Katz, P.A. (1968). Role of irrelevant cues in the
formation of concepts by lower-class children.
Journal of Educational Psychology. 59: 233-238.
- Kendler, H.H. (1966). Reflections on the conference.
Pp. 171-176 in Shulman, L.S. and Keislar, E.R.
(eds.), Learning by Discovery: A Critical Appraisal.
Chicago: Rand McNally.
- Kerlinger, F.N. (1966). Foundations of Behavioral
Research. New York: Holt, Rinehart and Winston.
- Kersh, B.Y. (1958). The adequacy of "meaning" as an
explanation for the superiority of learning by
independent discovery. Journal of Educational
Psychology. 49: 282-292.

- Kersh, B.Y. (1962). The motivating effect of learning by directed discovery. Journal of Educational Psychology. 53: 65-71.
- Kersh, B.Y. (1964). Learning by discovery: what is learned? The Arithmetic Teacher. 11: 226-231.
- Kersh, B.Y. (1965). Learning by discovery: instructional strategies. The Arithmetic Teacher. 12: 414-417.
- Kersh, B.Y. and Wittrock, M.C. (1962). Learning by discovery: an interpretation of recent research. Journal of Teacher Education. 13: 461-468.
- Klausmeier, H.J. and Wiersma, W. (1965). The effects of I.Q. level and sex on divergent thinking of seventh grade pupils of low, average and high I.Q. Journal of Educational Research. 58: 300-302.
- Klein, G.S., Barr, Harriet Linton and Wolitzky, D.L. (1967). Personality. Annual Review of Psychology. 18: 508-518.
- Kirk, J. and Amidon, E. (1967). When student teachers study interaction. Elementary School Journal. 68: 97-104.
- Kittell, J.E. (1957). An experimental study of the effect of external direction during learning on transfer and retention of principles. Journal of Educational Psychology. 48: 391-405.
- Kohlberg, L. (1968). Early education: a cognitive - developmental view. Child Development. 39: 1013-1062.
- Komisar, B.P. (1966). The non-science of learning. The School Review. 74: 249-264.
- Kornreich, L.B. (1969). Discovery versus programmed instruction in teaching a strategy for solving concept-identification problems. Journal of Educational Psychology. 60: 384-388.

- Kounin, J.S. and Gump, P.V. (1961). The influence of punitive and nonpunitive teachers upon children's concepts of school misconduct. Journal of Educational Psychology. 52: 44-49.
- Kuslan, L.I. and Stone, A.H. (1968). Teaching Children Science: An Inquiry Approach. California: Wadsworth.
- Lawrence, P.J. (1967). The mind of the child. Australian Journal of Education. 11: 97-109.
- Lee, L.C., Kagan, J. and Rabson, Alice. (1963). Influence of a preference for analytic categorization upon concept acquisition. Child Development. 34: 433-442.
- Leith, G.O.M. and McHugh, G.A.R. (1967). The place of theory in learning consecutive conceptual tasks. Educational Review. 19: 110-117.
- Lewis, D.G. (1968). Experimental Design in Education. London: University of London Press.
- Lindgren, H.C. (1967). Theories of human learning revisited. Pp.177-192 in Bower, E.M. and Hollister, W.G. (eds.) Behavioral Science Frontiers in Education. New York: Wiley.
- Lindquist, E.F. (ed.) (1951). Educational Measurement. Washington, D.C.: American Council on Education.
- Lindquist, E.F. (1953). Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin.
- McCall, R.B. and Kagan, J. (1967a). Stimulus-schema discrepancy and attention in the infant. Journal of Experimental Child Psychology. 5: 381-390.
- McCall, R.B. and Kagan, J. (1967b). Attention in the infant: effects of complexity, contour, perimeter and familiarity. Child Development. 38: 939-952.

- Maccoby Eleanor E. (ed.), (1967). The Development of Sex Differences. London: Tavistock.
- McDonald, F.J. (1968). A model of mathemagenic behaviors as intervening variables in classroom learning. Paper presented at the American Research Association Convention, Feb., 8-10, 1968. Chicago, Illinois.
- McGuigan, F.J. (1960). Experimental Psychology: A Methodological Approach. Englewood Cliffs: Prentice Hall.
- McKeachie, W.J. (1961). Motivation, teaching methods, and college learning. Pp.111-142 in Jones, M.R. (ed.), Nebraska Symposium on Motivation. Lincoln, Nebraska: University of Nebraska Press.
- Masling, J. and Stern, G. (1969). Effect of the observer in the classroom. Journal of Educational Psychology. 60: 351-354.
- Melton, A.W. (ed.) (1964). Categories of Human Learning. New York: Academic Press.
- Messick, S. and Fritzky, F. (1963). Dimensions of analytic attitude in cognition and personality. Journal of Personality. 31: 346-370.
- Meux, M.O. (1967). Studies of learning in the school setting. Review of Educational Research. 37: 539-562.
- Moroney, M.J. (1956). Facts from Figures. London: Penguin.
- Mussen, P.H. (1965). Discussion. Pp. 126-130 in Morrisett, L.N. and Vinsonhaler, J. (eds.), Mathematical Learning. Monograph of the Society for Research in Child Development. 30. No.1. (Serial No. 99).
- Newton, J.M. and Hickey, A.E. (1965). Sequence effects in programed learning of a verbal concept. Journal of Educational Psychology. 56: 140-147.

- Nuffield Foundation. (1967). Physics: Teachers Guide I.
London: Longmans - Penguin.
- Nuthall, G.A. (1966). An Experimental Comparison of Instructional Strategies in the Teaching of Concepts.
Unpublished doctor's thesis. Urbana: University of Illinois.
- Nuthall, G.A. (1968). Alternative strategies for teaching concepts. American Educational Research Journal.
4: 561-584.
- Oliver, D.W. and Shaver, J.P. (1966). Teaching Public Issues in the High School. Boston: Houghton-Mifflin.
- Olson, D.R. (1965). Note on Haslerud and Meyer's transfer of principles experiment. Journal of Educational Psychology. 56: 107-108.
- Osler, Sonia F. and Fivel, Myrna W. (1961). Concept attainment: I. The role of age and intelligence in concept attainment by induction. Journal of Experimental Psychology. 62: 1-8.
- Osler, Sonia F. and Trautman, Grace. (1961). Concept attainment: II. Effect of stimulus complexity upon concept attainment at two levels of intelligence. Journal of Experimental Psychology. 62: 9-13.
- Paradowski, W. (1967). Effect of curiosity on incidental learning. Journal of Educational Psychology. 58: 50-55.
- Ray, W.E. (1961). Pupil discovery vs. direct instruction. Journal of Experimental Education. 29: 271-280.
- Reynolds, J.H. (1966). Cognitive transfer in verbal learning. Journal of Educational Psychology. 57: 382-388.

- Reynolds, J.H. (1968). Cognitive transfer in verbal learning: II. Transfer effects after pre-familiarization with integrated versus partially integrated verbal-perceptual structures. Journal of Educational Psychology. 59: 133-138.
- Ripple, R.E., Millman, J. and Glock, M.O. (1969). Learner characteristics and instructional mode: a search for disordinal interactions. Journal of Educational Psychology. 60: 113-120.
- Rothkopf, E.Z. (1965). Some theoretical and experimental approaches to problems in written instruction. Pp. 193-221 in Krumboltz, J.D. (ed.), Learning and the Educational Process. Chicago: Rand McNally.
- Roughead, W.G. and Scandura, J.M. (1968). "What is learned" in mathematical discovery. Journal of Educational Psychology. 59: 283-289.
- Rowell, J.A., Simon, J. and Wiseman, R. (1969). Verbal reception, guided discovery and the learning of schemata. British Journal of Educational Psychology. 39: 233-244.
- Rudin, S.A. (1961). Measuring the teacher's effectiveness as a lecturer. Journal of Genetic Psychology. 98: 147-154.
- Russell, I.L. (1969). Motivation for school achievement: measurement and validation. Journal of Educational Research. 62: 263-266.
- Satterly, D. (1968). Perceptual, representational and conceptual characteristics of primary school children. British Journal of Educational Psychology. 38: 78-82.
- Sax, G. (1968). Empirical Foundations of Educational Research. Englewood Cliffs: Prentice Hall.

- Scandura, J.M. (1964). An analysis of exposition and discovery modes of problem solving instruction. Journal of Experimental Education. 33: 149-159.
- Scandura, J.M. (1966). Algorithm learning and problem solving. Journal of Experimental Education. 34: 1-6.
- Scandura, J.M. (1968). The formulation of research on subject matter learning. Psychology in Schools. 5: 330-341.
- Scandura, J.M. and Wells, J.N. (1967). Advance organizers in learning abstract mathematics. American Educational Research Journal. 4: 295-302.
- Scandura, J.M., Woodward, E. and Lee, F. (1967). Rule generality and consistency in mathematics learning. American Educational Research Journal. 4: 303-319.
- Scheffler, I. (1960). The Language of Education. Illinois: Thomas Books.
- Shulman, L.S. and Keislar, E.R. (eds.), (1966). Learning by Discovery: A Critical Appraisal. Chicago: Rand McNally.
- Siegel, L. (1960). The instructional gestalt: a conceptual framework. Teachers College Record. 62: 202-213.
- Siegel, L. and Siegel, Lila C. (1967a). The instructional gestalt. Pp. 261-290 in Siegel, L. (ed.), Instruction: Some Contemporary Viewpoints. San Francisco: Chandler.
- Siegel, L. and Siegel, Lila C. (1967b). A multivariate paradigm for educational research. Psychological Bulletin. 68: 306-326.
- Sigel, I.E. (1965). Styles of categorization in elementary school children. Cited in Wachtel, P.L. (1968). Journal of Personality. 36: 202-212.
- Sigel, I.E., Jarman, P. and Hanesian, Helen. (1967). Styles of categorization and their intellectual and personality correlates. Human Development. 10: 1-17.

- Skinner, B.F. (1954). The science of learning and the art of teaching. Harvard Educational Review. 24: 86-97.
- Smith, B.O. (1960). A concept of teaching. Teachers College Record. 61: 229-41.
- Smith, B.O. (1967). Recent research on teaching: an interpretation. High School Journal. 51: 63-74.
- Smith, B.O., Meux, M.O., Coombs, J., Nuthall, G.A. and Precians, R. (1967). A Study of the Strategies of Teaching. Urbana, Illinois: University of Illinois, Bureau of Educational Research.
- Smith, K.U. and Smith, Margaret F. (1966). Cybernetic Principles of Learning and Educational Design. New York: Holt, Rinehart and Winston.
- Snedecor, G.W. and Cochran, W.G. (1967). Statistical Methods. Iowa: Iowa State University Press.
- Stolurow, L.M. (1965). Model the master teacher or master the teaching model. Pp. 223-247 in Krumboltz, J.D. (ed.), Learning and the Educational Process. Chicago: Rand McNally.
- Suchman, J.R. (1961). Inquiry training: building skills for autonomous discovery. Merrill-Palmer Quarterly. 7: 146-169.
- Suchman, J.R. (1964). The child and the inquiry process. In Passow, A.H. (ed.), Intellectual Development: Another Look. Washington, D.C.: Association for Supervision and Curriculum Development.
- Suchman, J.R. (1966). A model for the analysis of inquiry. Pp. 177-188 in Klausmeier, H.J. and Harris, C.W. (eds.), Analyses of Concept Learning. New York: Academic Press.
- Swezey, K.M. (1955). After Dinner Science. London: Nicholas Kaye.
- Swezey, K.M. (1955). Science Magic. London: Nicholas Kaye.
- Taba, Hilda. (1962). Learning by Discovery: Psychological and Educational Rationale. Paper presented to the American Educational Research Association, Atlantic City, N.J.

- Taba, Hilda. (1963). Learning by discovery: psychological and educational rationale. Elementary School Journal. 63: 308-316.
- Taba, Hilda and Elzey, F.F. (1964). Teaching strategies and thought processes. Teachers College Record. 65: 524-534.
- Tallmadge, G.K. and Shearer, J.W. (1969). Relationships among learning styles, instructional methods, and the nature of learning experiences. Journal of Educational Psychology. 60: 222-230.
- Thorndike, R.L. (1951). Reliability. Pp. 533-559 in Lindquist, E.F. (ed.), Educational Measurement. Washington, D.C: American Council on Education.
- Tilton, J.W. (1951). An Educational Psychology of Learning. New York: MacMillan.
- Travers, R.M.W. (1969). An Introduction to Educational Research. New York: MacMillan.
- Tuckman, B.W., Henkelman, J., O'Shaughnessy, G.P. and Cole, Mildred, B. (1968). Induction and transfer of search sets. Journal of Educational Psychology. 59: 59-68.
- Underwood, B.J. and Richardson, J. (1956). Verbal concept learning as a function of instructions and dominance level. Journal of Experimental Psychology. 51: 229-238.
- U.N.E.S.C.O. (1956). Unesco Source Book for Science Teaching. Paris: U.N.E.S.C.O.
- Verduin, J.R. Jr. (1967). Conceptual Models in Teacher Education. Washington, D.C. The American Association of Colleges for Teacher Education.
- Vernon, P.E. (1961). The Structure of Human Abilities. London: Methven.
- Victor, E. and Lerner, M.S. (1967). Readings in Science Education for Elementary School. New York: MacMillan.

- Wachtel, P.L. (1968). Style and capacity in analytic functioning. Journal of Personality. 36: 202-212.
- Wallach, M.A. (1962). Commentary: active-analytical vs. passive-global cognitive functioning. Pp. 199-215 in Messick, S. and Ross, J. (eds.) Measurement in Personality and Cognition. New York: Wiley.
- Wallach, M.A. and Kogan N. (1965). Modes of Thinking in Young Children. New York: Holt, Rinehart and Winston.
- Ward, W.C. (1968). Reflection-impulsivity in kindergarten children. Child Development. 39: 867-864.
- Waterhouse, I.K. (1968). Some implications of modern psychological theory for the learning process. Unsourced: Mimeo'd.
- Werner, H. (1957). The concept of development from a comparative and organismic point of view. Pp. 125-148 in Harris, D.B. (ed.), The Concept of Development. Minneapolis: University of Minnesota Press.
- Westbury, I. (1967-68). The reliability of measures of classroom behaviour. Ontario Journal of Educational Research. 10: 125-138.
- White, R.W. (1959). Motivation reconsidered: the concept of competence. Psychological Review. 66: 297-333.
- White, S. (1963). Learning. Pp. 196-235 in Stevenson, H. (ed.), Child Psychology. Part I of the 62nd Yrbk of the National Society for the Study of Education. Chicago: University of Chicago Press.
- Wickelgren, W.A. (1964). Cues that elicit analytic-deductive methods in concept attainment. British Journal of Psychology. 55: 143-154.
- Winch, R.F. and Campbell, D.T. (1969). Proof? No. Evidence? Yes. The significance of tests of significance. American Sociologist. 4: 140-143.
- Winer, B.J. (1962). Statistical Principles in Experimental Design. New York: McGraw-Hill.

- Witkin, H.A., Dyk, R.B., Faterson, H.F., Goodenough, D.R. and Karp, S.A. (1962). Psychological Differentiation. New York: Wiley.
- Witkin, H.A., Lewis, Helen B., Hertzman, M., Machover Karen, Meissner, Pearl B. and Wapner, S. (1954). Personality Through Perception. New York: Harper.
- Wittrock, M.C. (1962). Set applied to student teaching. Journal of Educational Psychology. 53: 175-180.
- Wittrock, M.C. (1963a). Verbal stimuli in concept formation: Learning by discovery. Journal of Educational Psychology. 54: 183-190.
- Wittrock, M.C. (1963b). Effect of certain sets upon complex verbal learning. Journal of Educational Psychology. 54: 85-88.
- Wittrock, M.C. (1963c). Set to learn and proactive inhibition. Journal of Educational Research. 57: 72-75.
- Wittrock, M.C. (1966). The learning by discovery hypothesis. Pp. 33-75 in Shulman, L.S. and Keislar, E.R. (eds.), Learning by Discovery: A Critical Appraisal. Chicago: Rand McNally.
- Wittrock, M.C. and Keislar, E.R. (1965). Verbal cues in the transfer of concepts. Journal of Educational Psychology. 56: 16-21.
- Wittrock, M.C., Keislar, E.R. and Stern, Carolyn. (1964). Verbal cues in concept identification. Journal of Educational Psychology. 55: 195-200.
- Woodruff, A.D. (1964). The use of concepts in teaching and learning. Journal of Teacher Education. 15: 81-99.
- Worthen, B.R. (1968a). Discovery and expository task presentation in elementary mathematics. Journal of Educational Psychology Monograph. 59, No.1. Part 2.
- Worthen, B.R. (1968b). A study of discovery and expository presentations: implications for teaching. Journal of Teacher Education. 19: 223-242.

- Wright, J.C. and Kagan, J. (eds.), (1963). Basic cognitive processes in children. Monograph of the Society for Research in Child Development. No.86, 165-183.
- Yando, Regina M. and Kagan, J. (1968). The effect of teacher tempo on the child. Child Development. 39: 27-34.
- Yonge, G.D. (1966). Structure of experience and functional fixedness. Journal of Educational Psychology. 57: 115-120.