AFFECTIVE OBJECTIVES
IN
AN INTEGRATED SCIENCE CURRICULUM

SARA A BROWN

Thesis submitted in fulfillment
of the
requirements for the degree
of
Doctor of Philosophy
of the
University of Stirling

Department of Education
April 1975

Acknowledgments

This study has involved the investigation of the attitudes to science of a very large number of pupils. This would not have been possible without a substantial grant awarded by the Scottish Education Department and the help and co-operation of 47 Scottish schools (see Appendix A: 14A and 15A). I am most grateful to everyone concerned.

I would also like to thank a number of people within the University of Stirling for the considerable help they have given me: *Terence Davis* who supervised the pilot work and research design for the main study; *Fiona Breakes* who marked most of the AH4 scripts and carried out some of the calculations; *Andrew Comrie* who marked a large number of tests of divergency and carried out some of the calculations; *Phyllis McAdam* and *Hugh Batten* who marked a large number of tests of divergency; *Anne Kruzycka* and *Rob Marshall* who gave considerable assistance with computer programming; *May Abrahamson* who typed many of the tables and appendices; and *Muriel Heath* who took such care and effort in producing a beautifully typed manuscript for this thesis.

Finally I must express my gratitude to *Donald McIntyre*, my supervisor, who devoted so many hours to reading and discussing my work, provided invaluable constructive criticism and gave encouragement and help far beyond what I could have expected; and to *my husband and sons* who stuck it out, urged me on and never complained.

Sara A. Brown

April 1975.

CONTENTS (VOLUME I)

Page No.

Page No.

CONTENTS (VOLUME II)

Page No.

Page No.

LIST OF FIGURES

LIST OF TABLES

LIST OF TABLES (Cont'd)

LIST OF TABLES (Cont'd)

LIST OF TABLES (Cont'd)

LIST OF TABLES (Cont'd)

LIST OF TABLES (Cont'd)

INTRODUCTION

The purpose of this study is to see firstly, whether or not the attitude objectives of a particular integrated science course are 'any good', and secondly, the extent to which they are being achieved. It cannot be claimed that this constitutes a complete evaluation of the objectives (and certainly not the science course) but the aim has been to look at them from a number of different perspectives with the following in mind:

> 'A curriculum has not been positively evaluated in this full sense until it has been shown to have clear objectives and appropriate means to achieve them; to have objectives which have been proved against all comers to be educationally respectable; to connect with the abilities of those pupils for whom it is designed; and to be more efficient than rivals in the field. Only then can it get its tick' (White 1971).

The *integration of the sciences* and the inclusion of *affective as well as cognitive goals in science curricula* have been given considerable attention by writers on science education and curriculum developers over the last decade or two. The 'integration' debate has been far ranging (and sometimes bitter) among teachers, administrators and philosophers. While some of this debate has been at a high level of rational argument, some has been at the tub-thumping level between 'bandwaggon integrationists' and the 'separate disciplines brigade'. No such pair of factions can be identified in the discourse on affective goals (since no one has taken the categorical position that 'pupils should not achieve favourable attitudes to science'). Unfortunately, this has led to a situation where much of the literature exhorting us to teach towards affective goals has been statement of personal opinions unsupported by adequate theory or empirical evidence, and which has not received the criticism it deserves (perhaps able critics have felt that they have better things to do than comment on trivia). One of the things that this study attempts to do is to sort out the substantial contributions to this area of science education from the general ragbag. These theoretical perspectives are then related to a group of attitude objectives 'in action'. While many integrated science curricula have been developed throughout the world over the last few years, the meaning of 'integration', the value of the integrated approach, and its implications for pupils' attitude development have by no means always been made explicit. It is hoped that the examination of these areas that appears in this thesis will help to clarify the situation.

The study centres round a group of five attitude objectives laid down as part of an integrated science course. The course is the first cycle of *Science for General Education, Curriculum Paper 7, Scottish Education Department, 1969,* and is intended for pupils of all abilities in the first two years of secondary school in Scotland. The ages of these pupils are from 12 to 14 years and the year groups are designated 'S1' and 'S2'.

The five objectives are that pupils should acquire:

'[1] awareness of the inter-relationship of the different disciplines of science

[2] awareness of the relationship of science to other aspects of the curriculum

[3] awareness of the contribution of science to the social and economic life of the community

[4] *Interest and enjoyment in science*

[5] an objectivity in observation and in assessing observations' (Curriculum Paper 7, page 16).

The *problem* has been conceptualized in two broad questions, one non-empirical and one empirical:

(i) What are the arguments presented for inclusion of these attitude objectives in this curriculum, and are these arguments educationally valid?

(ii) To what extent are these objectives being achieved among secondary school pupils in Scotland, and what factors influence that achievement?

The Scottish Education Department showed considerable interest in the second of these questions. Indeed, Curriculum Paper 7 itself (page 31), in discussing the evaluation of the effectiveness of the Integrated Science course in achieving specific objectives, suggests that 'There is here a wide field for further investigations including assessment of attitudes'. The outcome of this interest was the award to this author of a 3 year grant in 1971 that enabled an empirical study of the achievement of the attitude objectives to be carried out on a large scale (3000+ pupils). The work relating to this study is described in Chapters 5, 7, 8, 9, 10, 11 and 12 of this thesis, while Chapters 1, 2, 3 and 4 (and possibly 6) relate to the first of the two questions.

In Chapter 1 the different types of argument that are put forward for teaching towards affective objectives are reviewed, and the rationale for the inclusion of the attitude objectives in Curriculum Paper 7 is examined in terms of these various categories. The sorts of choices about appropriate purposes for the curriculum that the curriculum developer must make before selecting his objectives are outlined, and the extent to which such selection depends on empirical evidence or subjective judgments is discussed. This chapter addresses itself to the *political* decisions that are made about which broad areas of attitudes are the concern of the school. Hopefully these decisions are taken on the basis of rational argument and empirical evidence (where that is available), and can be defended in relation to what *ought* to be worthwhile knowledge and skills for pupils to acquire.

The second chapter moves away from judgments about what are worthwhile goals, and concerns itself with the information we have (mostly from social psychological theory and experiment rather than directly from the classroom) about the ways by which attitude goals may be achieved, and the implications of this information for the selection of affective objectives. Attention is given to the relationships between cognitive and affective learning, to the most effective ways of presenting cognitive material for attitude modification, and to the relationships between verbally expressed attitudes and other behaviour. This psychological information provides some guidance for the curriculum developer in conceptualizing the attitude that he has in mind and in operationalizing it as an affective objective. However, the actual formulation of objectives is hampered by the overlap and lack of precision of much of the terminology of this area, and later sections of Chapter 2 provide a discussion of levels of specification and classification systems which may help to systematize the selection of affective objectives for a science curriculum. Finally the extent to which the objectives of Curriculum Paper 7 relate to such systems is examined.

The attitude objectives of Curriculum Paper 7 are integral parts of an integrated science course. Chapters 3 and 4 examine the proposition that the justification for the inclusion of these objectives may be implicit in the overall rationale for the integrated course.

Chapter 3 is devoted to an examination of the various meanings of 'integration', those meanings that Curriculum Paper 7 ascribes to integration, and those attitudes to science that can be logically related to such meanings.

Chapter 4 explores the various categories of arguments that have been used to establish the value of an integrated approach and the emphasis laid on each of these categories by Curriculum Paper 7. The justifications for such courses are made in terms of either their anticipated outcomes or the constraints under which the learning takes place. For each category then, the extent to which such outcomes are likely to emphasise attitudes and such constraints to constrain achievement of attitude objectives is discussed. This leads to the formulation of a number of hypotheses reflecting the extent to which the Integrated Science course might be expected to be more or less effective than three separate science courses in achieving the attitude objectives. These hypotheses are investigated in the later, empirical part of the study.

Up to this point the concern has been with evaluation of the objectives to see if they have educational worth. Chapter 6 asks whether these five dimensions of *affective curriculum objectives* are related to any distinct *attitude dimensions* on which pupils can be seen to differ. This investigation uses factor analysis to explore the responses of pupils to five attitude scales each related to one of the five attitude objectives. These scales are the main instruments used in the part of this study concerned with the second broad (and empirical) question.

The construction of the five attitude scales, together with an account of the pilot study, is described in Chapter 5. The pilot study was carried out among pupils of both sexes, from three age groups and with one group following Integrated Science and one following separate science courses. From the results of the pilot study it was possible to formulate further hypotheses on affective achievement for the main empirical study that was to be carried out in 40 secondary schools.

The research plan for this main study of the assessment of the achievement of affective objectives is laid out in Chapters 7, 8 and 11. Chapter 7 identifies the research questions and discusses the relevant literature. The general research findings are then used to generate further hypotheses for this study. Chapter 8 describes the research structure and strategy. It provides details of the selection of criterion and independent variables, the sampling procedures adopted, the data gathering procedures and some preliminary studies of pupils' ability measures. Chapter 11 provides a detailed description of the statistical analyses of the attitude scale scores between-schools, within-schools, and within-teaching-groups.

Chapter 12 collects together results from all the analyses relating to each of the attitude objectives.

Two aspects of the empirical work that did not employ the five attitude scales developed in Chapter 5 are described in Chapters 9 and 10. Chapter 9 is concerned with a study of teachers' attitudes and evolved from the assumption that there should be salient attitude factors among teachers that will influence the attitudes to science of their pupils. In Chapter 10 an account is given of an attempt to measure, and to discriminate between, the 'scientific attitudes' (rather than the 'attitudes towards science') of pupils following the two types of science course, by means of a Cognitive Preference Test.

The final chapter provides interpretations of the results of the analyses and relates these to the research questions and the hypotheses derived from theory, other research findings and the pilot study. The closing section attempts to draw some conclusions from the two strands (empirical and non-empirical) of this evaluation, and so to make a judgment on the five objectives from the Integrated Science course.

The empirical study goes some way towards identifying the sorts of 'educational' variables that are, or are not, likely to be useful in explaining differences in attitudes to science among pupils. This has implications for the kinds of research questions that it is fruitful to ask, and for *who* it is that should identify what are the appropriate questions.

Two very general conclusions about research and development in this area emerge from the study:

a) research based only on what teachers think are important and interesting questions and not on any guiding theoretical framework, runs the risk of coming up with very little useful information,

b) curricula that are not related to some substantial theoretical structure, run the risk of having obscure or ambiguous goals which cannot be understood or have their worth assessed, and so are unlikely to be translated into the classroom procedures intended by the curriculum planners.

CHAPTER 1

THE CONCERN FOR TEACHING TOWARDS AFFECTIVE OBJECTIVES

Widely varying views have been expressed on the extent to which, and the ways in which, educators should concern themselves with affective goals in the secondary school. While there is a measure of consensus on the kinds of cognitive objectives with which the school curriculum should be concerned, there is no such agreement on what are appropriate kinds of affective objectives. Consequently choice of such objectives involves political judgments which must be based on the various arguments for and against emphasis on different types of affective goals.

These various types of argument that have been put forward for teaching towards attitude goals will now be examined. They have been grouped into seven categories, and after discussion of each category, the choices that confront the curriculum developer, and which relate to his particular selection of objectives, will be outlined.

Finally one secondary school science course (Scottish Integrated Science for 12 to 14 year-olds) will be examined to see the ways in which, and the extent to which, its selection of attitude objectives has followed the sorts of arguments and choices outlined in the first part of the chapter.

Arguments for Teaching Towards Affective Goals

1. *Ideal society - good citizens*

The first type of argument is concerned with assertions about the characteristics of an ideal, democratic society, and with derived conclusions about the appropriate role of a 'good', 'responsible' and 'useful' citizen. Such conclusions imply appropriate affective objectives.

For example Broudy (1971) discusses the shaping of educational goals in this area by the values of democracy. He sees as appropriate the formation of attitudes that involves the individual seeing the valuing process as being a *duty:*

> 'a) I have a duty to become informed on and involved in all great social issues, whether the impact on me is obvious or not. For the nature of a modern society is such that I may not be able at a glance to detect the impact on myself. Citizens no longer can pick and choose the issues on which they shall be involved; b) I have a duty to respect the rights of others which means that I shall automatically feel uneasy about taking action that ignores or over-rides these rights.'

Broudy presents a logical argument for the development of these attitudes based on his concept of the political, moral, economic and intellectual values appropriate to a modern technological, democratic society and distinguishing it from a totalitarian system. There appears to be no empirical evidence on the feasibility of such goals.

Miles has proposed a model (Figure 1.1) for curriculum design which has a similarly idealistic starting point, and which incorporates a system for the selection of both affective and cognitive objectives.

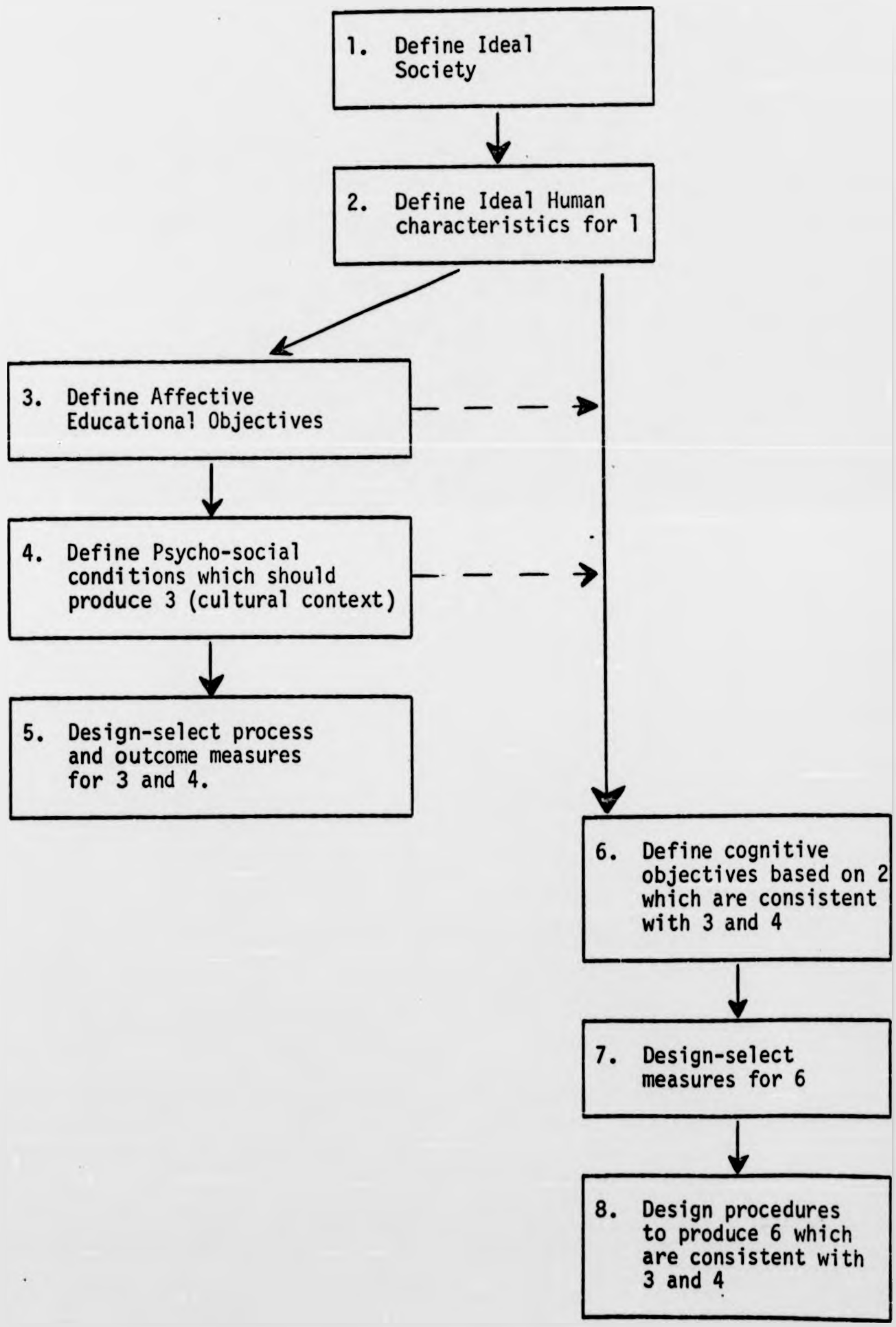


FIGURE 1.1: Educational Design Model (Miles 1972).

He suggests that the development of 'informal' over 'traditional' education has changed a situation in which

> 'little effort and few instructional decisions are made on the basis of affective outcomes..[and] a great deal more emphasis [is placed] on low level cognitive outcomes',

to one in which we

> 'place such affective goals as self-concept, self-directedness, curiosity, resourcefulness, creativity, responsibility and sensitivity in a position of highest priority, and thus plan much of the educational procedures to facilitate achieving these outcomes...[and] devote considerable emphasis to high cognitive skills, such as learning to learn, problem-solving, critical thinking, and social and communication skills'.

Miles argues that the school has always influenced pupils in all these areas but that in the 'traditional' approach the influence in the affective (and to a lesser extent the high cognitive) area was largely 'unplanned'. 'Informal' approaches have therefore, led to a higher ratio of 'planned influence on pupils' to 'unplanned influence on pupils' in the curriculum, and also to a larger proportion of the 'planned influence' being concerned with affective goals. He sees the behavioural objectives model of 'formulation of objectives' ————> 'selection of learning experiences' ————> 'evaluation of programme', as having been almost exclusively concerned with the cognitive domain. This he attributes, in part, to the preoccupation of learning psychologists throughout this century with cognitive learning.

His proposed model for instruction (Figure 1.1) incorporates the priorities, as he sees them, of the affective domain. He admits that it is impossible to define the 'ideal society' or 'ideal human characteristics' but he suggests that we can agree on some of these characteristics (tolerance, self-esteem, social responsibility). These can then lead directly to specification of affective objectives. Information on the psycho-social conditions necessary for attainment of these affective objectives, he argues, can be obtained from psychological and sociological theory, and research and development must provide tools for measurement of affective outcomes. He would then see this development as having precedence over cognitive goals which would have to be formulated consistently with the established affective aims.

This distinctive relationship between cognitive and affective objectives specified by the model is reflected in arguments of this category. The internal logic and completeness displayed by the model is, however, seldom encountered. Broudy is exceptional in providing explicit attitude objectives related to human characteristics, most arguments peter out after defining characteristics of the ideal society and its citizens. It is even less likely that the argument process will have continued far enough to define either the characteristics of the social context in which such objectives may be achieved, or the appropriate teaching methods for attaining the specified outcomes. Nor is it common for the cognitive objectives to be directly related to the affective as Mile's model requires.

It appears that these arguments do not conflict with the model, but fail to fulfil all its criteria. The model can, perhaps, be identified as an idealised form of this type of argument.

2. Desirable socio-political values and behaviours

Arguments in the second category are concerned with assertions about specific 'desirable', socio-political attitudes, values or behaviours. Unlike the previous category there is no utopian starting point of definition of the 'ideal society'. If the first category arguments are identified (in idealised form) with Mile s' model, then the second category arguments, in contrast, take as their starting point the second stage of the model and define what are seen as 'ideal human characteristics'. In other words, they identify 'good citizen' qualities in terms of what society *is*, rather than in terms of what it *ought to be*.

Like the other categories, this has been particularly stressed by science educators. Harmin, Kirschenbaum and Simon (1970) suggest that:

> 'We simply cannot afford to train a generation of students who know the *how* and *why* of scientific phenomena, but do not have a process for inquiring into the values issues raised by the topics they study.'

Some writers see education towards such attitude formation and value-systems as the solution, possibly the only solution, to many contemporary problems and needs of our society. Kuhn (1973) states that:

> 'more than ever, there is a pressing need for citizenry that has scientific literacy and the awareness of values necessary to make decisions or influence decisions about questions of population control, radio active fallout, pesticide usage, or industrial effluence *etc.*'.

Few authors explicitly advocate that teachers should teach 'values' (either their own or society's) as such. For example Harmin *et al.* (1970) insist that pupils must be encouraged 'to search for their own values, the only values that will ever mean anything to them'. Yet these same authors, in their exemplification of value-education by a study of lead poisoning in slum children, clearly expect the pupils undertaking this study to acquire the *particular* 'LEAD POISONING IS A BAD THING AND SOMETHING HAS TO BE DONE ABOUT IT' attitude. They suggest seven outcomes from this study, none of which relate to any other attitude as a possible objective *e.g.*

> 'they [the pupils] could write their congressmen and urge them to look at the problem and perhaps draft legislation to combat the wasteful and almost criminal neglect of human life which occurs with lead poisoning'.

Kuhn (1973) also emphasises the valuing-process and exhorts teachers to stress this 'vital component in the preparation of citizens to function well in a changing society'. He avoids negating this with a list of values to be taught, and instead some 'pedagogical tools of value education' are described. But no clear framework in which to use these tools is provided, and we are left with a list of suggestions that lack clarity of educational purpose:

> 'it is useful to sample student opinion before and after a unit on drugs. It is interesting to see if their opinions are affected...Students may wish to survey the attitudes of adults in the community on air pollution and compare their views...as a basis for social action...Students need to develop skills in the analysis of articles in the popular press...hopefully they will accept the values expressed as being those of the author and in a non-judgemental way consider them.'

It seems unnecessary to pretend that the intention is not to inculcate specific attitudes among pupils. There seem to be a number of attitudes in this area that are viewed (probably with justification) as 'scientifically warranted value judgments' that 'can be backed up all the way back to breakfast, *using* the resources of various sciences and technologies' (Scriven 1966).

Allport (1961) shuns any ambiguity of approach. He argues that if the school does not teach towards a system of values (which he views as 'matters of importance' rather than 'matters of fact') then the pupil will be at the mercy of the confusion engendered by the communication media. He argues that if a pupil does not experience values such as honesty and tolerance in his home they must be provided by the school. For the pupil who does see such values expressed by his parents, it may be that they will be viewed by him as 'old hat' and so rejected unless they are reinforced by the larger outside world (*i.e.* the school).

Several writers (*e.g.* Harbeck 1970) have pointed out the paradox in which the teaching of values is shunned while at the same time the main purpose of education is seen as passing on the culture of the society. Such a process necessarily involves the transmission of values, and as Krathwohl *et al.* (1964 page 56) point out we are already blatant in our indoctrination of pupils on what is 'good' and 'of value' in art, music and literature.

As in category 1 arguments, the priority is still given to affective objectives. It may well be a straightforward matter to relate cognitive objectives to the sorts of affective goals that are proposed here (though this is infrequently done). For example Kuhn's concern for pupils' appreciation of problems such as population and pollution control may provide fruitful areas for the development of related cognitive study.

warranted value judgments' that 'can be backed up all the way back to breakfast, *using* the resources of various sciences and technologies' (Scriven 1966).

Allport (1961) shuns any ambiguity of approach. He argues that if the school does not teach towards a system of values (which he views as 'matters of importance' rather than 'matters of fact') then the pupil will be at the mercy of the confusion engendered by the communication media. He argues that if a pupil does not experience values such as honesty and tolerance in his home they must be provided by the school. For the pupil who does see such values expressed by his parents, it may be that they will be viewed by him as 'old hat' and so rejected unless they are reinforced by the larger outside world (*i.e.* the school).

Several writers (*e.g.* Harbeck 1970) have pointed out the paradox in which the teaching of values is shunned while at the same time the main purpose of education is seen as passing on the culture of the society. Such a process necessarily involves the transmission of values, and as Krathwohl *et al.* (1964 page 56) point out we are already blatant in our indoctrination of pupils on what is 'good' and 'of value' in art, music and literature.

As in category 1 arguments, the priority is still given to affective objectives. It may well be a straightforward matter to relate cognitive objectives to the sorts of affective goals that are proposed here (though this is infrequently done). For example Kuhn's concern for pupils' appreciation of problems such as population and pollution control may provide fruitful areas for the development of related cognitive study.

3. *Social adjustment of the pupil*

Thirdly, arguments may be put forward that the pupil's social adjustment may be determined by the priorities given to certain affective goals.

Raths *et al.* (1966) suggest that many behaviour problems of pupils may be put down to a lack of values. The pupil who is unable to decide what is 'right', 'good' or 'worthy' will have an unclear relationship with modern complex society and will be apathetic, confused, irrational, over-conforming or over-dissenting. They class such value-related disturbances with emotional disturbances and physical handicaps. They propose that such problems can be eased in intensity and frequency by value-experiences and assert that the valuing process has been shown to lead to improvement in attitude towards learning, initiation, self-direction, perseverence and active participation on the part of the pupils. They view the development of values, (manifested in attitudes, interests, beliefs and activities) as a process whereby the teacher provides the pupil with alternatives from which he can *choose* freely, the pupil *prizes* this choice and is then seen to *act* on it repeatedly. They provide details of suggestions of what they see as appropriate teaching strategies to aid the pupil in this development process, placing particular emphasis on helping him to clarify his values. As Khan and Weiss (1973) point out, methodological short comings of the empirical studies (*e.g.* Raths 1962) on the effectiveness of such strategies in modifying behaviour, have provided only tenuous evidence in their favour.

Raths *et al.* do not see education in this area as fulfilling emotional needs. Such fulfillment they suggest must come from outside the pupil, they are concerned with valuing that must be done by the pupil. They assert that emotional needs must be fulfilled before the valuing process, since the child must face up to 'self' before it can answer value questions, but they do not fit fulfillment of such needs into their affective teaching framework. Beatty (1969), however, sees the pupil's self-concept as the pivot of learning, and the development of his values as identical with the development of his concept of himself as adequate. Further he suggests that suppression of normal development of feelings and emotion is an important, but little understood, source of distorted and ineffective behaviour. He suggests that if children are encouraged to display emotions and feelings, their self-concept (need for such a display) will be more in line with their perception of what society approves of, so leading to more consistent and constructive behaviour. From the teacher's point of view, he sees the expression of feeling by pupils as providing clues to whether or not learning is taking place and why learning in some contexts is unpleasant.

These arguments are not making explicit assertions about what 'good' social behaviour is. They are, nevertheless, although couched in the language of pupils' 'needs', concerned that the outcome be pupil behaviour that is *acceptable* to society. They do not define any 'ideal human characteristics' (and certainly not an 'ideal society'), although such characteristics are implicit in the removal of 'problems'. They define their affective objectives in terms of the development of values or self-concepts, and plot, quite clearly, the paths by which such objectives might be achieved. The link with cognitive objectives is, however, tenuous, the emphasis is not on *what* is learned (but the assumption is apparently made that the learning will become more effective).

This category views various personal 'needs' of the individual pupil as dominant. However, the concern is still with the pupil acquiring societal norms and 'The individual is said only to be able to develop "healthily" if he satisfies his individual need to fit securely into the pattern set by the rules and practices of his society' (Wilson 1971, page 29).

4. Rational, intelligent, valuing behaviour

Fourthly, affective objectives are seen as instrumental in enabling pupils to achieve what is generally accepted as rational, intelligent, valuing behaviour.

It has been suggested that in the past teachers have either ignored this area altogether or have pressed their own values on to pupils. The former is seen as undesirable (avoiding the issue), and the latter as both undesirable (indoctrination) and unproductive (pupils frequently reject these values). The solution is seen in providing the pupil with the opportunity to develop his own value-system.

Gribble (1970) suggests that:

> 'The business of education in the Affective Domain is to get children feeling positively towards what is valuable or true or right insofar as there are publically acceptable criteria for determining goodness or truth or rightness. These criteria are not contingently related to the specification of educational objectives in the affective domain - they are a necessary part of such a specification.'

He is critical of the Taxonomy of Affective Objectives (Krathwohl *et al.* 1964), and suggests that these objectives are 'described as if they were separable from the development of mind through knowledge', and are illustrated by test items that do not refer to standards of what is desirable and valuable. He argues that the cognitive objectives of the Taxonomy of Objectives, Cognitive Domain, (Bloom *et al.* 1956) are restricted so that 'false beliefs and specious procedures are not covered', but no restrictions (in terms of acceptable value-systems) are imposed in the affective domain, thus divorcing two domains that are logically indivisible.

He sees the criteria and standards of the category of Evaluation in the Cognitive Domain as being not merely linked with the Affective Domain, but as integral parts of the Affective Objectives.

As an example Gribble proposes that the objective: 'increased appetite and taste for what is good in literature' (Krathwohl *et al.* 1964 page 124) must presuppose criteria for evaluating 'what is good in literature' if it is to be an *educational* objective. It is suggested that by omitting the necessity for these criteria the affective domain is trivialised and confusion in the author's concept of education is revealed.

Smith (1966) is also concerned with cognitive-related values which he calls 'ratings'. He distinguishes between attitudes towards objects (distinguished by 'liking', 'enjoying', 'accepting', 'disapproving', 'aversion to', *etc*) and the rating of objects (ascribing to the objects' terms such as 'good', 'desirable', 'ugly', 'wrong' *etc*), as two different types of value.

Smith sees the teaching of ratings as leading to a situation where the pupil has learned that object X is rated in a particular way by 'competent persons', has understood the objective criteria on which such a rating is based, has developed the ability to use these criteria to rate X for himself and has eventually applied a similar strategy to rate objects other than X. This closely parallels the progression through the cognitive hierarchy of Knowledge, Comprehension to Application (Bloom *et al.* 1956).

Smith considers the case where the criteria for the ratings are controversial. He suggests that the teacher may simply leave the situation open, with different sets of criteria available and so different ratings possible, or he may attempt to arrive at a single set of criteria. In the latter case he suggests that the teacher's objective becomes concerned with the pupils developing a favourable attitude towards (liking and commitment to) a particular set of criteria. Since the object of the attitude is not *X*, but the *criteria used in rating X*, the objective functions as an affective (attitude towards certain criteria) means to a cognitive (use of those criteria for rating X) end, rather than as an affective end in itself.

Smith argues that it would be unwise to replace the teaching of ratings of X by direct teaching towards an acceptable attitude towards X. While he admits that the latter should lead to 'proper' ratings of X, he sees such an attitude as lacking in intellectual support. This, he suggests, would result in the individual being at a loss to defend or justify his attitude in confrontation with others.

In contrast to the three earlier categories, the decisions about affective and cognitive objectives and the procedures for attaining them are closely integrated, and the former sublimation of the cognitive to the affective is no longer apparent.

5. Affect conducive to cognitive functioning

The fifth type of argument concerns itself with affective objectives that are conducive to cognitive functioning within a particular discipline.

Bloom *et al.* (1971 page 226) and Scriven (1966) suggest that teachers who do not attempt this cognitive-related type of value-education are avoiding their responsibilities. Scriven argues that we can make judgments *that can be justified* about what are valuable 'cognitive-values' on the basis of a combination of *'the facts about the entity being evaluated'* with 'other *facts about the needs, wants, and ideals of the valuing agents*'. His view is that:

> 'Criticism and approval are a necessary part of the process of internal improvement of a science, and value judgments expressing them are important and complex, and hence much debated, but absolutely inescapable - except by the ostrich route'.

Thus, for example, the scientist and science teacher are not expected merely to present the facts about the consequences in scientific enquiry of being 'open minded' or not, of being 'objective' or not, of being 'sceptical' or not. They are also expected to examine the values of their society in this area (*e.g.* to what extent is effective, efficient scientific enquiry seen as universally desirable?) and then make a *decision* on what values should form the basis for affective objectives. 'Our goal, in value matters, should be the discovery of the solution to problems of selection and rejection that require our professional expertise, and the demonstration, to those who face the problem, of the validity of our solution'. (Scriven 1966)

If those in power accept such recommendations, Scriven sees those who 'teach well' teaching specific values of two kinds, and as exemplification suggests that:

> 'We will be teaching the general value of objectivity, of the scientific approach, as the most effective way of arriving at the truth. And we will be teaching how to apply this general method to socially and practically important issues'.

The kinds of objectives implied by these arguments have received much of the attention given to values in science education. Eiss and Harbeck (1969 page 3) point out that:

> 'Values have seldom been stated as goals of science instruction. In fact many scientists take pride in pointing out that there are no philosophic values in science...but...we find that many such values as honesty in reporting, openmindedness, and the usefulness of evidence in making decisions are very important'.

Selection of attitude objectives along the lines of these arguments follows a very different pattern from those of the first three categories. Mile s' model of category 1 requires that we use *philosophical* and *political* arguments to establish judgments about the ideal society, desired human characteristics, and so appropriate affective objectives. Categories 2 and 3 are also based, but to a lesser extent, on such arguments. All three categories, however, give priority to the affective objectives before the cognitive.

In contrast, the cognitive-related affective objective of this category stems from a value judgment based on *empirical* evidence of what the 'facts about the entity' and the 'facts about the needs, wants and ideals of the valuing agents' appear to be. The curriculum developer is thus expected to look at the knowledge and the society as they are, and pass this information on to the pupil (and this includes information on *conflicting* ideals, needs, and theories, where they exist). Scriven (1966) and Bloom *et al.* (1971 page 226) suggest that this type of value education may be achieved by:

a) teaching that clearly delineates objectively established facts from hypothesis, thus leaving pupils to make their own choices where possible;

b) developing pupil skills necessary to arrive at and test conclusions that the teacher thinks are true,

c) emphasising that many 'established' conclusions may turn out to be wrong, but at the same time providing the documentation that has led to their establishment.

If, as this category prescribes, the affective objectives are seen as a means to facilitate cognitive functioning, there is an implicit initial selection of cognitive objectives and the distinctive affective-cognitive sequence of the earlier categories disappears.

6. Affective routes to cognitive goals

Sixthly, there are the arguments that maintain that affective goals are instrumental to the achievement of cognitive educational goals. Thus, the child who is interested, satisfied, persevering will gain in cognitive performance as a result of this. Mager (1968) asserts that:

> 'The liklihood of the student putting his knowledge to use is influenced by his *attitude* for or against the subject; things disliked have a way of being forgotten' (page 11).

about the entity' and the 'facts about the needs, wants and ideals of the valuing agents' appear to be. The curriculum developer is thus expected to look at the knowledge and the society as they are, and pass this information on to the pupil (and this includes information on *conflicting* ideals, needs, and theories, where they exist). Scriven (1966) and Bloom *et al.* (1971 page 226) suggest that this type of value education may be achieved by:

a) teaching that clearly delineates objectively established facts from hypothesis, thus leaving pupils to make their own choices where possible;

b) developing pupil skills necessary to arrive at and test conclusions that the teacher thinks are true,

c) emphasising that many 'established' conclusions may turn out to be wrong, but at the same time providing the documentation that has led to their establishment.

If, as this category prescribes, the affective objectives are seen as a means to facilitate cognitive functioning, there is an implicit initial selection of cognitive objectives and the distinctive affective-cognitive sequence of the earlier categories disappears.

6. Affective routes to cognitive goals

Sixthly, there are the arguments that maintain that affective goals are instrumental to the achievement of cognitive educational goals. Thus, the child who is interested, satisfied, persevering will gain in cognitive performance as a result of this. Mager (1968) asserts that:

> 'The liklihood of the student putting his knowledge to use is influenced by his *attitude* for or against the subject; things disliked have a way of being forgotten' (page 11).

Rogers (1972) asks 'Do our young students *enjoy* some of the things they do in science? Does their enjoyment reach a higher level than fun and become a feeling of doing and thinking in science?'

Klopfer (1971) states that 'There is strong psychological evidence that students learn better, learn more, and remember longer when they find pleasure in the learning experience'. He does not provide references for the evidence that he has in mind. Meyer 1961 also believes that 'the creation of lasting interests, in some cases, would lead pupils towards scientific careers, but in all cases should help them to appreciate the scientific basis of our culture and to develop an adequate philosophy of life'.

These arguments reflect much popular opinion regarding the wisdom of taking an affective means to a cognitive end. However, whether or not the introduction of affective goals influences achievement of cognitive goals is largely untested in the classroom context. The next chapter will describe a psychological theory (and the attempts to test it) which proposes that if an individual's affect towards a particular psychological object is manipulated in some direction (*e.g.* more favourable) then his cognitive beliefs towards that object will be modified in the same direction. The empirical evidence supporting this hypothesis appears sound, but it involved hypnotising subjects, which suggests that its relevance for secondary school teaching is tenuous.

Alvord (1972) in a study of 'science achievement' and 'attitude-toward-school' of over 3000 elementary, junior high, and high school pupils, found that

> 'little more than 1 to 5 percent of an individual's science achievement score can be attributed to performance on the attitude-toward-school measure or *vice versa*...In other words, it can be inferred from the present research that working with pupils to create improved attitudes or interests toward school will not bring about resultant changes in achievement commensurate with the amount of time spent in improving such attitudes and interests'.

However, in general the evidence for relationships between affective and cognitive variables among pupils is equivocal. Gardner (1975) cites seven studies showing zero or negative relationships, and three studies revealing positive relationships. Of the latter group he criticises one for labelling items as 'pupil interest' when they may simply reflect parental expectations; one he cannot evaluate because there is no indication of whether 'Course Satisfaction' was determined before, or after, pupils received their achievement grade; and for the third he points to the much stronger relationship between the variables among pupils who have opted to continue with science than among the general population.

On the other hand Khan and Weiss (1973) cite five further studies showing positive affective-cognitive relationships, and four further studies showing zero or negative relationships. Unlike Gardner, Khan and Weiss seem to support the positive findings and they report Jackson's (1968) argument that zero or negative relationships may result from the use of instruments that are not sufficiently sensitive to reveal differences in the intensity of attitudes for large groups of pupils with neutral or mixed feelings. They also suggest that the studies would not have been expected to reveal positive relationships if those relationships had been non-linear. The significance of the results of these empirical studies for the selection of curriculum objectives is limited, since they reveal only *correlational* and not *causal* relationships, and therefore no affective-cognitive sequence can be assumed.

These various studies have been concerned with affective behaviour at a number of levels. It can be argued that it is *logically* necessary to achieve certain *minimal* affective objectives before achievement of cognitive objectives. The very low level affective objective of *Awareness* (Krathwohl *et al.*, 1964) corresponds to a behaviour where 'the learner will merely be conscious of something - that he take into account a situation, phenomenon, object or state of affairs' (page 99). Since the prerequisite of *knowing* about something is *being conscious* of it, we have a logical affective-cognitive sequence of learning at this level. However, at higher levels of behaviour there is no such obvious logical sequencing and the cognitive-affective relationship becomes problematic.

Like categories 4 and 5 this category contrasts with 1, 2, and 3 in that it does not choose affective before cognitive goals. If affective goals are a means to achievement of cognitive goals then logically the cognitive goals must be identified first.

Selection of goals will take into account 'human characteristics', insofar as such characteristics are relevant to the motivational nature of the goals, but no definitions of 'ideal behaviour' or 'ideal society' will be involved. The emphasis will be on selection of appropriate knowledge[1] and of affective

[1]'Appropriate knowledge' is not an objective, culture-free selection, it will reflect, to some extent, the society's ideals. However, choices made in this area are less likely to be idiosyncratic, arbitrary, or politically motivated than are the choices to be made in categories 1 and 2.

objectives consistent with effective acquisition of that knowledge by the pupils.

7. Meeting pupils' 'needs' other than social adjustment

The seventh type can barely be classified as an argument, but unfortunately it makes up much of what Scriven (1966) refers to as 'the vast magnitude of the relevant literature and the marked triviality of much of it'. Assertions are made that cognitive goals alone are inadequate and that the aims of the curriculum must meet pupils' needs such as those for excitement, satisfaction, and enjoyment. It is taken as self-evident that there must be affective objectives formulated in these terms. However, unless these pupil needs are related to the pupil's social adjustment (as in the second category of arguments) it is difficult to see how excitement, delight, or joy can be considered as *educational objectives*. They seem rather to be *means* whereby other attitude or cognitive goals might be achieved, more related to motivation than goals. It may be necessary to satisfy pupils' needs for enjoyment, excitement, and interest *in the subject* for learning to occur, but that is identical with a category 6 argument.

If meeting these 'needs' *is* the educational goal, it is necessary that an explanation be provided that indicates which of the many 'needs' that pupils have are the ones to be met and why. Pupils enjoy, are interested in, and are excited by many bizarre and cruel things: 'A young bully, for example, from *his* point of view may "need" to find victims' (Wilson 1971, page 7), 'children have many interests that are educationally undesirable - *e.g.* blowing up frogs with bicycle pumps' (Hirst and Peters 1970, page 37).

Many interests of pupils will be more acceptable than these, and some will be considered *educationally* valuable. If it is the latter that are to determine the curriculum they will have to be selected on some basis and identified specifically *e.g.* 'Pupils' interest in scientific experiments', or 'Pupils' need to carry out experiments'. For such examples a case would have to be made for 'satisfaction of the pupils' interest in/need to carry out, scientific experiments' as *primary goals* of the curriculum, rather than as conditions that may have to be satisfied if the pupil is to achieve other goals such as: the development of practical skills in science, knowledge of the procedures and methodologies of science, or the determination of a series of qualitative or quantitative results.

If they are *intermediate* goals related to one of these other goals, then the argument will belong to category 6; if not then, logically, it will fall into category 2 or category 3.

Areas of Concern - Selection of Objectives

Each of these arguments recommends a selection of affective objectives based on a particular assumption about what issues are the *appropriate concerns of the school.* The choice of what is appropriate is a political one, and the curriculum developer must decide whether or not it is a responsibility of the school:

a) to define the ideal characteristics of our society (which may or may not correspond to existing characteristics, and if not will involve the school in teaching towards goals specifically aimed at *changing* society) as category 1 arguments suggest;

b) to formulate its goals to reflect what it sees as society's requirements of pupil behaviour as in categories 1, 2 and 3;

c) to take responsibility for fulfilling certain psychological needs of pupils in order to promote their social adjustment and other less clearly defined attributes, as in categories 2 and 7;

d) to develop goals which contribute directly to, or facilitate, the pupil's cognitive learning (the acquisition of knowledge, understanding and associated rational behaviour) as in categories 4, 5, and 6.

Once the judgment on the appropriate area of concern has been made, the nature of the argument provides broad conceptual criteria for the choice of affective objectives. To varying degrees these criteria are themselves subjective and therefore their interpretation into more operational criteria is, again, a matter of political choice *e.g.* Broudy's *personal conception* of a modern democratic society, leading through logical argument to his choice of attitudes characterizing the good citizen; Raths *et als' judgment* on the necessity of the valuing-process for pupils' self-adjustment, and their interpretation of this, on the basis of psychological theory, into objectives of pupil behaviours of 'choosing', 'prizing' and 'acting on'; and Scriven's *'scientifically warranted value-judgment'* on cognitive-related values, leading directly to objectives. These examples provide criteria of varying subjectivity (Scriven's being the least subjective).

There is also variation in the degree to which operationalization of the criteria can be related to empirical evidence. Scriven's argument and selection (as described earlier) leans heavily on empirical evidence;

Raths' is supported by some limited empirical studies; Broudy's selection of objectives is unrelated to empiricism. In general, arguments in categories 5 and 6 (and, to a lesser extent, 2) will be able to use substantial relevant empirical work.

The Justification for Affective Objectives in the Scottish Integrated Science Course

Following this *general* discussion of the arguments for the use of affective objectives and the choices which have to be made before selection of objectives, Curriculum Paper 7 will be examined and the extent to which its affective objectives are related to such arguments and choices will be considered.

The Paper specifies two kinds of objective: general or summative, and specific or formative. The general objectives are divided into three groups:

a) Knowledge and understanding.

b) Attitudes.

c) Practical skills (page 16).

The concern for affective objectives is clearly related to category (b). However, there are no such distinctions made in the specific objectives that are laid down for individual sections of the syllabus.

There are five general attitude objectives, that pupils should acquire:

' 9. awareness of the inter-relationship of the different disciplines of science

10. awareness of the relationship of science to other aspects of the curriculum
11. awareness of the contribution of science to the economic and social life of the community
12. *Interest and enjoyment in science*
13. an objectivity in observation and in assessing observations' (page 16).

No explicit rationale for selection of these objectives is provided. The Working Party who drew up the Curriculum Paper make several statements which may be relevant to their choices:

> 'Science must somewhere be seen as a whole' (page 18).
>
> 'The corpus of knowledge in science is growing at an explosive rate, and the effect of this growth is to alter not only our physical conditions but also our morals, our ethics, and our whole cultural development ...The implications of this for science, as part of general education, are obvious. We have also included as our final aim our desire to expose pupils to this cultural aspect of science' (page 13).
>
> 'In all this nothing has been said about interest, enjoyment or satisfaction in the things of science, or in the possible commitment to the study of science ...Everything it has done in preparing a syllabus... has been influenced by the desire to foster in pupils an interest and an enjoyment in science' (pages 13-14).
>
> 'The Working Party has included as one of its stated aims a desire to develop objectivity in observation, and has structured some of the content of its syllabus to develop this' (page 12).

Such statements may define the position of the Working Party but they do not justify it. It is difficult to determine which category of argument they have in mind, though for the older age group (third and fourth year secondary) they provide an explanation for teaching towards attitudes in terms of what *is* done in schools rather than what *ought* to be done:

> 'Permeating everything which is done in school is... [an]...element which is concerned with the formation of attitudes...Headmasters and teachers, with whom we have individually and in groups discussed this particular set of young people, continually use phrases like "awareness of", "interest in" "committing themselves to", "involved in" "natural interests of" "active attitudes towards" "acquire a taste for" "satisfaction from work", "sense of purpose", "responsible attitude towards" "pupils who are co-operative", "willing to learn". This represents so small a sample of the different affective implications we have encountered that relevant attitude formations must be seen as one of the essential contributions of science to the general curriculum' (page 33).

Here the concern seems to be a) with motivation (probably corresponding to arguments in category 6), and b) with socially acceptable behaviour (probably corresponding to arguments in category 2).

If the specific objectives that are laid down for each of the 15 sections of the syllabus are examined, the very close relationship of affective concerns with the cognitive is clearly seen. In almost all cases the nature of the objective suggests that it was selected on the basis of an argument of category 5 or 6. Out of 209 objectives there is only one (*i.e.* Section 1.8 'Interest and enthusiasm for science', page 98) that is not either cognitive, or cognitive with an affective component such as:

1.1 'awareness that human senses are limited and unreliable'

2.5 'some ability to observe objectively, this time in changing situations over longer periods of time'

3.7 'awareness of the need for control and efficient use of energy resources'

5.12 'awareness of the need for patience in a long term project (*e.g.* crystal-growing)'.

It may be the case that the rationale for the inclusion of affective objectives in this course is not presented explicitly since it is implicit *either* in the rationale that is presented for teaching science in an integrated form, *or* in the meaning that has been given by the Working Party to 'integration'. These possibilities will be examined in Chapters 3 and 4.

CHAPTER 2

THEORETICAL PERSPECTIVES ON THE FORMULATION OF AFFECTIVE OBJECTIVES

In the last chapter the various kinds of arguments put forward for teaching towards affective objectives were examined. The sorts of *political* choices that have to be made, before selection of such objectives, were outlined. Having made such choices, the curriculum developer has identified broadly, the sorts of attitude that he hopes to develop or modify. His problem now is to select those affective objectives that a) it is feasible to expect achievement of by the pupils, and b) are expected to be instrumental in bringing about the desired attitude development or modification.

To some extent, as discussed on page 28, this selection will depend on political choices (*e.g.* the particular characteristics of behaviour that are acceptable to society) or empirical evidence (*e.g.* those affective attributes that are demonstrably conducive to cognitive functioning), but it must also take cognisance of the *ways by which attitude change may be achieved.* If we are to use the attainment of a set of affective objectives to achieve some broad attitude goal, then it is necessary that those objectives should not be at variance with, and if possible should directly reflect, the methods we can postulate as being effective means to reach that attitude goal.

Unfortunately, little substantial information is available on effective methods for attitude modification in the classroom. It is necessary to look to social psychological theories and evidence, some of which may have been developed and investigated in contexts very different from the secondary school.

Smith (1966) has summarized what he sees as the three general research findings (he provides no references) by saying that attitudes are modified by a) the cognitive learnings related to them, b) the way in which the cognitive material is learned, and c) the attitudes of the 'group' with which the pupil associates. In the restricted classroom situation where the attitudes of the 'group' are themselves the target, he sees manipulation of attitudes by (c) as irrelevant. In this chapter the evidence for the effectiveness of (a) will be considered by examining the theories and associated evidence on the relationship between cognitive and affective learning. Information available on the most effective ways of presenting cognitive material for attitude modification will be examined in an attempt to clarify (b).

Having scrutinized the relevant psychological information the curriculum developer has still to decide how he is conceptualizing the 'attitude' (*i.e.* as some sort of inherent attribute, as a verbally expressed attitude, or as some other pattern of behaviour) before he can formulate his objective. This decision will depend, in part, on whether or not (and for what purposes) he wishes to assess the attitude, and on the relationships (conceptual and empirical) that exist between inherent attitudes, expressed attitudes and other behaviour. Information on these relationships will be discussed later in the chapter.

Finally the appropriate level of specification for the objectives must be decided, and either suitable ones from other curricula must be selected, or been developed and investigated in contexts very different from the secondary school.

Smith (1966) has summarized what he sees as the three general research findings (he provides no references) by saying that attitudes are modified by a) the cognitive learnings related to them, b) the way in which the cognitive material is learned, and c) the attitudes of the 'group' with which the pupil associates. In the restricted classroom situation where the attitudes of the 'group' are themselves the target, he sees manipulation of attitudes by (c) as irrelevant. In this chapter the evidence for the effectiveness of (a) will be considered by examining the theories and associated evidence on the relationship between cognitive and affective learning. Information available on the most effective ways of presenting cognitive material for attitude modification will be examined in an attempt to clarify (b).

Having scrutinized the relevant psychological information the curriculum developer has still to decide how he is conceptualizing the 'attitude' (*i.e.* as some sort of inherent attribute, as a verbally expressed attitude, or as some other pattern of behaviour) before he can formulate his objective. This decision will depend, in part, on whether or not (and for what purposes) he wishes to assess the attitude, and on the relationships (conceptual and empirical) that exist between inherent attitudes, expressed attitudes and other behaviour. Information on these relationships will be discussed later in the chapter.

Finally the appropriate level of specification for the objectives must be decided, and either suitable ones from other curricula must be selected, or a new set must be developed. The problem at this stage may well be the vagueness of many affective objectives which results in a) an inability to understand what an objective is expecting in the way of pupil behaviour change, and b) some doubt as to the range of affective behaviours that are covered by a particular group of objectives. The curriculum developer would find useful some sort of classification system indicating meaningful terminology for affective objectives, and giving some indication of the scope of behaviours reflected in existing curriculum objectives and evaluation instruments.

The final sections of this chapter will provide some suggestions from educators for appropriate levels of specification for affective objectives, and a description of the classification systems available to the developer for reference during his formulation of objectives. The extent to which the attitude objectives selected for Curriculum Paper 7 are accommodated by these systems will be discussed.

No clear patterns for the formulation of affective objectives have yet gained wide acceptance and much of the use of the terminology is confused. A brief comment on this usage, therefore, appears before the main body of this chapter.

Terminology - Affect, Attitudes and Values

Terms such as affective, attitude, and value are frequently used interchangeably in educational literature.

Affect is usually taken to relate to feelings and emotions and Krathwohl *et al.* (1964) classify affective objectives as those:

> 'Which emphasise a feeling tone, an emotion, or a degree of acceptance or rejection' (page 7).

Such a definition is broad enough so that while the emphasis must always be affective it is possible that a cognitive component is also present. Indeed, except in the case of, objectives of the type: 'Pupils will enjoy X', it is difficult to see how the cognitive element can be avoided.

There is no consensus of opinion on the definition of attitude. Vernon (1953) acknowledges this lack of agreement and suggests a working definition of attitude that 'implies a personality disposition or drive which determines behaviour towards or opinions about a certain type of person, object, situation, institution or concept'. Scott (1968) suggests that the looseness of convention within social psychology and the variety of purposes for which the construct of attitude is used have led to a situation where 'it is unrealistic to expect a single, final definition of "attitude" to emerge within the foreseeable future'. He sees one view of attitudes as attributes having affective, cognitive and possibly evaluative components together with an 'action tendency'. As such, attitudes are regarded as a sub-class of motives distinguished from other such sub-classes by their cognitive component.

Scott suggests that values may be considered as a sub-class of attitudes which may 'include the belief that the focal object is desirable or undesirable'. On the other hand we have already seen that Raths *et al.* (1966) consider attitudes to be manifestations of values. Oppenheim (1966 page 109) suggests that:

> 'For ease of understanding, social psychologists make a rough distinction among these different levels, calling the most superficial one beliefs, the next one attitudes, a deeper level, values or basic attitudes, and a still deeper level, personality'.

Lack of clarification of appropriate theoretical constructs has led to considerable overlap of the three concepts. Krathwohl *et al.* (1964), as described later, had to formulate a different terminology to solve the problem. For the purposes of this chapter we will not try to distinguish between the terms but will attempt to restrict the use of the word attitude to that of a noun, affective as an adjective and value as a verb.

Relationships Between Cognitive and Affective Learning

Krathwohl *et al.* (1964 page 54) adopt the widely accepted view that the achievement of cognitive objectives is, in general, an effective and necessary way to affective achievement. They quote several statements conceptualizing attitudes as having both cognitive and affective components which tend, for an individual, to reinforce each other. Hence, the expectation that a pupil's cognitive and affective behaviour will co-vary closely with each other. They see teachers in the classroom deliberately taking well-trodden cognitive routes to unfamiliar affective goals, although sometimes unfortunately:

> 'our preference for approaching affective achievement through the attainment of cognitive objectives tends to focus attention on these cognitive goals as ends in themselves without determining whether they are actually serving as means to an affective end' (page 57).

Cognitive and affective achievement may sometimes be mutually supportive but there are numerous instances where thorough, careful study (frequently forced) of subject matter has led to very negative attitudes on the part of pupils. Eiss and Harbeck (1969) list ten factors that often cause pupils to lose interest in science and nine of these are concerned with the pupil's cognitive learning. Unfortunately there seems to be no conclusive evidence in the classroom setting, about the significant factors affecting the attainment of affective objectives through cognitive learning.

A convenient way to conceptualize attitudes (Hollander 1971, Katz 1960, Wagner 1969) is to consider them as having three major components: cognitive (beliefs), affective (feelings) and behavioural (predisposition to respond). Rosenberg (1960) has considered the consistency of the affective and cognitive components within a single attitude. He proposes that inconsistency between these two elements must lead to a change in one or other. Thus he predicts that a pupil undergoing an irreversible change in his cognitive beliefs (*e.g.* information supplied by the teacher and seen as important by pupil), will show a corresponding change in affect. This is pertinent to the view of cognitive objectives as a 'means' to affective 'ends'. He finds the apparent evidence for this is extensive but that it is mainly correlational and

> 'most of these studies, however, do not provide for a precise check of whether, and to what extent, the communications designed to alter cognitions actually do so'.

However, he does refer to two studies (Carlson 1956 and Peak 1959) which confirm his prediction. These studies consisted of a pretest of cognitive and affective responses towards an attitude object ('Negroes being allowed to move into white neighbourhoods' and 'making good grades' respectively), an intervening cognitive manipulation designed to change the individuals beliefs about the object, and a post-test of the cognitive and affective responses. Both studies showed that a shift of the cognitive component produced a shift in the affective component in a direction maintaining the consistency between the two. His theory also involves the prediction of the converse *i.e.* that

> 'if a person somehow undergoes an "irreversible" change in his *affect* towards an object his beliefs about that object will show corresponding change'.

Such a prediction is more difficult to confirm, since manipulation of affect without manipulation of the cognitive element is very difficult. Rosenberg achieved this by inducing changes of feeling about an attitude object ('the abandonment of the United States policy of giving economic aid to foreign nations') while his subjects were in hypnotic trances. After being awakened from hypnosis the subjects had measures of cognition towards the object administered. These demonstrated that a cognitive shift had occurred in a direction maintaining the cognitive/affective consistency, confirming the prediction. One strong objection to considering this work as relevant to the classroom is that change of affect by such highly manipulated means may not produce results that are generalizable to the usual ways that affect is changed.

Rosenberg (1956) has also investigated affective/cognitive co-variation in connection with *stable* attitudes. He found that stable positive affect and stable negative affect towards an attitude object are associated with beliefs assigning positive values and blocking negative values towards the object, and beliefs assigning negative values and blocking positive values towards the object, respectively. Further, intermediate affect was associated with beliefs relating the object to less important values than those associated with extreme affect.

Rosenberg and Krathwohl both suggest that in many cases the affective and the cognitive are being manipulated at the same time, and it is impossible to say which goal is the 'means' and which is the 'end'. Having suggested discovery-learning as an affective means to a cognitive end, Krathwohl then goes on to use the discovery - learning of Suchman's (1962) Elementary School Scientific Inquiry program as an exemplification of simultaneous affective and cognitive achievement.

Festinger's (1957) theory of cognitive dissonance is the most extensively researched in this area (Sherwood, Barron and Fitch 1969 refer to a survey that reveals almost 400 separate studies). The theory proposes that dissonance is an uncomfortable state, and dissonance is said to exist when an individual holds two cognitions one of which implies the opposite of the other. Cognition here is interpreted very broadly and encompasses knowledge, beliefs and attitudes. For example, suppose a pupil is fearful that scientific discoveries will destroy our society, yet has seen on television a programme reporting on a scientific invention that is saving lives. There will exist dissonance between the pupil's unfavourable attitude towards science and the favourable information. It is proposed that the pupil will be motivated to reduce the dissonance and achieve consonance. One way to do this is to increase the number of consonant cognitions. If this unfavourable attitude is fairly firm he may add consonance to this cognitively by regarding the useful aspects of science as being trivial compared with the devastation caused by the explosion of nuclear bombs. He may reduce the dissonance behaviourally by joining a society dedicated to the cessation of all scientific research that could be used in the development of any type of weapon. He could selectively choose to read only writings that criticize science on the grounds of its expense, danger and inhumanity. On the other hand if the information he received from the television programme is seen as important and if other cognitions favourable to science are added to it (in, say, his school science class), then the pupil may reduce dissonance by changing his attitude. Whether the pupil's attitude will change or be retained will depend on the dissonance that will be created by the change (cognitions consonant with the old attitude will be dissonant with the new). This implies that achievement of cognitive objectives in conjunction with affective objectives will lead to attitude change provided the pupil does not have excessive cognitions that would, in that event, become dissonant with his new attitude. One such cognition arises if more than slight coercion has been used. Suppose the pupil has been forced to behave in a way dissonant with his unfavourable attitude to science (severe punishment will be inflicted if pupil does not write an essay extolling the contribution of science to society). This forcing will, in itself, reduce the dissonance ('I don't believe in what I am doing but I am justified in behaving this way because they are forcing me to') and so diminish the likelihood of attitude change.

Whether the pupil adopts a favourable or unfavourable attitude to science he will always have some cognitions dissonant and some consonant with his attitude. Dissonance will never be completely removed, indeed it is suggested (Sherwood, Barron and Fitch 1969) that people may

> 'have some tolerance for a certain level of dissonance, further reduction of which is unnecessary or not particularly satisfying.
>
> ....Mild stimulation, tension, or dissonance may actually be pleasurable'.

Once again the paucity of evidence on the relationship between cognitive and affective learning in the classroom context must be emphasised. Psychological experiments have provided support for the generalizations of the cognitive-affective relationships described above, but it seems likely that the pattern of co-variation of cognitive and affective achievement in science courses will depend very much on the precise nature of the cognitive and affective objectives used. Thus, while certain elements of science *knowledge* may lead to increased *interest* in science, at this time it is not clear what those elements are. If we are concerned with an attitude which has a more emphatic cognitive component than does 'interest in science' (*e.g.* 'awareness of the inter-relationship of the different disciplines of science') then we may be in a more favourable position to predict what cognitive material will be instrumental in the attainment of the attitude.

The Presentation of Cognitive Material for Attitude Modification

It may be possible to identify the cognitive material that is likely to be instrumental in the achievement of any particular affective objective. However, the way in which the cognitive learning occurs may affect the extent of the attitude change, and may even prevent any change occurring. Information on this can influence the formulation of the objectives.

For example is a programme that *aims for pupils to synthesise for themselves 'favourable attitudes' from information provided,* likely to be more or less successful in developing such attitudes than a programme that *aims for pupils to develop those 'favourable attitudes' after being given an explicit description of them* (in addition to the information)? Will an objective that demands that the pupil exhibit the attitude and provide support for it in public, be more or less successful in the desired attitude change than one which does not require the public demonstration?

Logically it is necessary that the pupil must *attend* to the cognitive material if it is to influence his attitudes, but what further kinds of behaviour should be aimed for if effective attitude change is to be achieved?

Once again we have to look to psychological experiments and theory rather than to classroom evidence. Hovland, Janis and Kelley (1953) approach attitude change from the stand point of the identification of the conditions favourable for the 'learning' of attitudes. Their conclusions rely on extensive experimental work, mostly among university students. They are concerned with a) the communicator of new information (the teacher), b) the nature of the communication (lesson content and strategy), c) the audience (the pupils), d) the response of the audience (active participation and duration of effects). Most of their conclusions are more pertinent to the development of instructional materials and strategies than to the formulation of objectives. However, two points they make would seem relevant to attitude formation in the present context.

Firstly, a series of experiments (Janis and King 1954, Kelman 1953, King and Janis 1956) investigated the factors that determine whether an individual, induced to behave in a way conforming to a particular attitude, will adopt this attitude or not. Comparisons between 'active participant' students (who had to deliver talks conforming to a particular attitude), with 'passive control' students (who read and listened to the same material), suggested that the role-playing of the former leads to greater attitude change (in conforming direction) than that of the latter. King and Janis (1956) then showed that the magnitude of the opinion change depended on the improvisation demanded of the active participants, those required to read a script and then give the talk without it exhibiting a greater change than those allowed to read the talk from the pre-read script. Kelman's (1953) study in which school pupils were allowed to write essays conforming to a view put forward in a talk (under one of three incentive conditions), essentially confirmed the 'active-improvisation' results.

The second point is on the question of whether a conclusion should be given explicitly to pupils or whether they should be allowed to draw their own conclusions. Hovland and Mandell (1952) observed a significantly greater shift of opinion change among students with whom conclusions have been explicitly drawn (in this case 'devaluation of currency') than among those left to derive the conclusions independently. It is pointed out however, that this treatment may not be effective if the attitude change performs an ego-defensive function, here implicit presentation (*e.g.* as in psychotherapy) may be more effective.

All these experiments were concerned with attitudes not likely to be considered as deviating extensively from the normal attitudes of the group. Similarly, the attitudes associated with school science curriculum objectives are unlikely to advocate 'way-out' or unacceptable positions for the pupils.

These results suggest that behaviour in which the pupil complies with a set task (role-playing) would be an effective means towards attitude development. This is particularly interesting in view of the fact that course objectives specifying this type of behaviour rarely appear (though such behaviour is, in reality, a frequent outcome). Krathwohl *et al.* (1964) suggest that 'behaviour instilled in the learner at this level may never become more internalised and more self directed' (page 120). Hovland *et al.* (1963) also accept that 'under certain conditions role-playing and other means of producing verbal conformity may *interfere* with acceptance', but the evidence of attitude change induced by role-playing is extensive (Carlson 1956, Culbertson 1957, Festinger and Carlsmith 1959, Harvey and Beverly 1961, Janis and Gilmore 1965).

Hughes (1971) used: a) exposure to persuasive communication, b) observation of role-playing, and c) participant role-playing, in an experiment on changes in attitudes to science and science teaching among 184 student teachers of elementary science. He found that the three treatments combined led to higher attitude scores than no treatment. There was, however, no evidence for significant differences between the attitudes of three sub-groups who (1) read the communication (R), (2) read the communication and observed role-playing (RO), or (3) read the communication and participated in the role-playing (RRP). (The attitudes were assessed immediately following the treatments and again nine days later.) Hughes suggests that the failure of the observation of, and participation in, role-playing to create a further significant difference over the persuasive communication treatment alone may be caused by the relatively high level of acceptance of the communication's ideas that the students displayed prior to the role-playing. The verbalisation in the role-playing was not sufficiently effective to raise the attitudes from this high level to an even higher level (the mean attitude scores for the three treatment groups did however increase from R to RO to RRP).

Hughes concludes that the strategy whereby the learner is provided with explicit persuasive material 'oriented toward the cognitive domain of the individual's', is likely to be more effective in influencing attitudes than is one in which the individual is expected to synthesise for himself a 'favourable' attitude from all the information presented to him.

As Hughes points out:

> 'by allowing the choice of the content and source of the communication to be controlled by the investigator, this technique makes a closer approximation to a classroom situation. In a class...the persuasive communication would be a reading assignment and the role-playing would be similar to a traditional "class discussion".'

He seems reluctant to abandon the role-playing element despite its lack of significant contribution. It may be, if his hypothesis regarding the cause of the non-significance is correct, that results in a situation with school pupils as subjects might provide support for the value of role-playing as the superior technique. We would expect a wider variation in motivation and comprehension levels among pupils following a required course in school, than among students following a chosen course at an institution of higher learning. Thus while some of the pupils, like the students, may have reached a relatively high level of acceptance of the ideas of the communication after reading it once, others may need the reiteration of the arguments together with new information that could emerge from the role-playing procedures. It would seem to be worthwhile to carry out an experiment similar to that of Hughes, with pupils in the classroom context, retaining the role-playing element. The results of such an experiment would provide considerable guidance on the desirability of formulating objectives in terms of role-playing as a means to the development of particular attitudes.

Krathwohl would like to limit objectives specifying such compliant behaviour to areas such as health and safety, and states that 'of course, the teacher strives to reduce the incidence of acquiescent behaviour to an absolute minimum'. This would seem a reasonable point of view if we consider objectives as 'ends' in themselves - behaviour that is merely compliant is not an attractive goal. But Krathwohl is concerned with a hierarchical classification of objectives where the lower levels are explicitly described as 'means' towards high level 'ends':

> 'The learning of the more difficult and internalised objectives must be in the form of a "loop" which begins with a simpler and more overt behaviours gradually moves to the complex and more internalised behaviours and repeats the entire procedure in new areas of content and behaviour until a highly internalised, consistent, and complex set of affective objectives is finally developed'.

The demonstrable effectiveness of the provision of explicitly stated conclusions (and possibly role-playing) for attitude development, suggests that emphasis on this form of acquiescent behaviour may be fruitful, and might well be employed on each 'loop'. The boundary between compliance and coercion is however, tenuous and, as we saw earlier in the discussion of cognitive dissonance, the greater the coercion the less likely is the attitude to change.

The Concern of Affective Objectives with Inherent Attitudes, Expressed Attitudes or Other Behaviour

Curriculum developers, in formulating affective objectives have been noticeably inexplicit in:

i) describing the attitudes with which they are concerned,

ii) specifying their purposes in teaching towards these attitudes,

iii) distinguishing between their interests in pupils' verbally expressed attitudes and pupils' behaviour (other than verbal expression),

iv) stating whether appropriate evaluation should be concerned with grading individual pupils or comparing the effectiveness of different curricula or judging the competence of different teachers.

Krathwohl (1965), Harbeck (1970) and Popham and Baker (1970) all make the point that the affective objective is concerned primarily with what the pupil *does* while the cognitive objective focuses on what he *can* do. The former, they suggest, may well be more important in the long term. (It is, however, the latter that school assessment is normally based on, since that is what can be most easily and reliably measured). This would imply that the 'attitudes' frequently referred to in affective objectives must have a theoretical construct that is identified with what the pupil 'does do' (*i.e.* his customary behaviour). The term attitude is not normally synonymous with behaviour in either psychological or common usage, though it may be described as a predisposition to certain behaviour. The bulk of attitude assessment uses psychometric methods where the subject is asked to verbally express his attitude. It is suggested (Popham and Baker 1970 page 37) that this method elicits the responses that the pupil thinks the teacher will approve of even when anonymous, self-report devices are used.

They assert that:

> 'If for example, it is possible to observe how a pupil behaves when he is not aware that he is being observed, one has a much better estimate of how he really feels than if the learner is clearly behaving for the teacher's benefit.'

This may or may not be true depending on how one conceptualizes 'really feels'. If it is as some sort of inherent attribute that dictates how the individual will behave, then clearly behaviour will be a valid measure. However, the rationale for differentiating behaviour which indicates how a pupil 'really feels' from other behaviour which does not, is far from clear. De Fleur and Westie (1963) describe two conceptions of attitude. Firstly, the 'probability' conception in which responses of a particular type towards an attitude object can be predicted as likely to occur. For example a pupil may tend to respond positively to items on a science-interest inventory. He may be said to have a positive attitude towards science-interest items, and his responses in this context are predictable. He may, however, be observed to behave in a negative way when confronted with a science society to join, a science book to read or a scientific expedition. He may then be said to have a negative attitude to science-interest activities. A lack of consistency is explained by saying that these are two different types of response, and there is no single attitude or 'science-interest' involved. Such an approach does nothing more than side step the question of the relationship between pen-and-paper attitude measures and other behaviour.

Secondly, De Fleur and Westie (1963) describe the latent-process attitude. This is conceptualized as some attribute, hidden within the individual that regulates behaviour and expressed attitudes. When behaviour and expressed attitudes do not correspond we have to decide which is the more reliable predictor of the latent attitude. To decide whether we will try to provide an answer to this last very difficult question, or fall back on the 'attitude-is-what-is-measured-by-this-response' strategy, we require information on the extent of the inconsistency between expressed attitudes and other behaviour, and on the possible sources of these inconsistencies.

A rather gloomy conclusion from an extensive review by Wicker (1969) of empirical research in this area states that:

> 'it is considerably more likely that attitudes will be unrelated or only slightly related to overt behaviours than that attitudes will be closely related to actions. Product-moment correlation coefficients relating the two kinds of response are rarely above 0.30, and often are near zero. Only rarely can as much as 10 per cent of the variance in overt behavioural measures be accounted for by attitudinal data'.

His review centres on 33 studies concerned with attitudes to job, minority group members and miscellaneous objects, none of which closely parallel the classroom situation. In all cases verbal expressions of attitudes were taken before observation of overt behaviour.

Cognitive dissonance theory predicts that if the verbal expression of attitude occurs after the overt behaviour then we might expect the attitude to change to reduce dissonance and so be more highly correlated with behaviour. Knox and Inkster (1968) were able to confirm this prediction. They found that betters were significantly more confident that their horse was going to win if they expressed their opinion after they had placed their bet than if they had expressed it before. Similarly Mann and

Abeles (1970) found voters adopted a more confident attitude about their candidate after voting for him than before. Wicker (1969) quotes six studies demonstrating the greater expressed attitude-behaviour consistency when the behaviour precedes the expression of attitude (and four studies that do not).

Campbell (1963) has suggested that the inconsistencies may be more apparent than real in that they may result from *situational* pressures. It is easy to express interest by marking a questionnaire but it involves more effort to demonstrate interest by participating in a scientific activity. He proposes that there is a situational threshold that is lower for expressing positive feelings on a questionnaire than that for generating the momentum to activate interests in science. In Campbell's terms inconsistency would only be demonstrated if pupils expressing negative feelings on the questionnaire were then seen to participate in activities assumed to demonstrate an interest in science. He sees many of the research results as evidence of these thresholds and not as evidence of inconsistency.

Inscho and Schopler (1967) suggest that many workers may have seen high positive correlations between attitudes and behaviour and consider them so self-evident that they have not bothered to report them. If this is the case, such workers must have been blissfully ignorant of the concern that has been shown for at least forty years (see La Pierre, 1934) about the ignorance of the attitude-behaviour relationship.

Festinger (1964) has reviewed research and theory on the relationship between attitude *change* and behavioural *change*. He found only three relevant studies none of which showed any obvious relationship. Since the relationships between attitudes and behaviour are not as simple as they might be, it seems appropriate to look at other factors that may influence these relationships.

Wicker (1969) gives a comprehensive account of other factors that researchers have suggested must be taken into account when predicting overt behaviour from attitudes. Wicker implies that many of these factors arise from guess work and that

> 'clearly the greatest need in the attitude-behaviour area is to operationalize and to test the contributions of the factors which have been offered as reasons for attitudes-behaviour inconsistency'.

The following list is a summary of the factors discussed by him, with examples concerned with science education provided by this author.

1. The behaviour inconsistent with the expressed attitude may be consistent with some other attitude which is more strongly held. For example, pupil says that he is interested in science but takes no part in out-of-school science activities. Reason - he is even more interested in football and spends all his spare time on that.
2. The motivation to behave in a particular way may be stronger than the motivation associated with the expressed attitude. For example, a pupil wishes to make experimental observations honestly and objectively but he 'fiddles' his readings to achieve 'right' answer. Reason - his motivation to get a good mark from the teacher is greater than his motivation to be objective in his experiments.
3. The inconsistency may be due to the inability (intellectual or social) of the individual to behave in an appropriate way. For example, a pupil shows considerable enthusiasm about the formation of a science

club at school, yet he offers no contribution when asked to make suggestions for scientific activities for the club. Reason - he does not understand what level of scientific activity is feasible for a group of young pupils.

4. The individual may not have a high enough activity level to behave in a way consistent with his attitude. For example, a pupil signs a petition calling for the cessation of the use of scientific research establishments for the development of weapons, but he refuses to canvas support for this or go on a march to a chemical warfare research station. Reason - he is intrinsically apathetic and loath to act to achieve his goals.
5. The association between attitude and behaviour may be much stronger if the expressions of attitude and behavioural responses are given in similar situations. For example, if a pupil expresses an interest in science on a questionnaire given to him in a science lesson then he is very likely to agree to participate in an interesting science activity in that same lesson. Reason - his expressed attitude to science and his behaviour are both restricted to the context of this science class and other variables become less important.
6. The inconsistency may be due to the presence (or absence) of certain people at the time of the expression of attitude or occurrence of behaviour. For example, a pupil may indicate on the confidential research questionnaire that he finds science extremely boring, yet he visits the science museum once a month. Reason - his father is fascinated by science and the pupil, partly through parental pressure and partly through a desire to please his father, will not admit this boredom.
7. Norms of the peer-group may force behaviour not consistent with the attitude. For example, a pupil shows concern, in an essay, for the pollution of rivers by scientific industry, but he is unwilling to join

an active campaign to visit the industries and put forward the environmentalist's case. Reason - his peer group are sneering at the anti-pollution movement and he is unwilling to deviate from their social norms.

8. Opportunities for behaviour consistent with the attitude may not be available. For example, a pupil expresses a dislike of science yet proceeds to take three sciences at 'O' level. Reason - the school observes certain qualifying conditions for 'O' level courses and this is the only course the pupil qualifies for.

9. The stimulus provided for the expression of attitude may be much more general than that provided for the overt behaviour. For example, a pupil expresses a dislike for science yet avidly watches television programmes on 'Black Holes of Gravity'. Reason - his feelings and beliefs about science in general are as a boring, repetitive and unimaginative school subject, these are quite different from his conception of Black Holes as new, interesting and exciting.

10. Unforseen extraneous events may lead to behaviour inconsistent with previously expressed attitudes. For example, a pupil shows enthusiasm for the formation of a radio club to meet after school one evening a week, yet he never turns up to the club once it is formed. Reason - the club meets on a Thursday night and the bus company has decided not to run any buses from the area of the school to the area to where the pupil lives after 5 p.m. on Thursdays.

11. Anticipated consequences of behaviour may modify that behaviour so that it is inconsistent with the attitude held. For example, a teacher considers that science in school is best taught in an integrated form, yet she organises and teaches the science course as separate subjects. Reason - the promoted posts are in the separate science subjects and she expects that consideration for such posts will not be given to

those favouring integration.

It can be seen that these are little more than speculations. The reasons supplied with each example are no more than suggestions, and almost any of the examples could be explained by almost any of the factors.

There has been no empirical investigation of the relationships of the variables discussed here. Fishbein (1967) has attempted to formulate a theory of behavioural prediction using a complex multi-attitude - multi-method approach, and concludes that:

> 'the most important determinants of behaviour may be other variables than an individual's beliefs about, attitude toward or general behavioural intentions toward, a given object. Indeed, this approach clearly indicates that behaviour towards an object may be completely determined by situational or individual difference variables rather than any variable associated with the stimulus object *per se*'.

If this is the case, then observation of behaviour cannot be considered as a valid measure of, say, an inherent attitude to science. It is even possible that we may have a more valid measure by using the attitude scale approach, which elicits written expressions of attitude from pupils and requires them to express both feelings and beliefs (no distinction is made here between the cognitive and affective components of the attitude).

We cannot make generalizations on whether the curriculum developer in formulating affective objectives is setting goals for overt behaviour or expressed attitudes or inherent attitudes. However, we may be able to make some inferences in individual cases.

An affective objective may be concerned with the immediate behavioural outcomes from the teaching. For example we might hope that the pupil would acquire an appreciation for adopting an attitude of open-mindedness, or objectivity in observation, or concern for others, in his laboratory work. The objective would then probably be concerned with his overt behaviour in the laboratory (though unfortunately at this time we have no methods for recognizing reliably what a pupil *does* to show that he is open-minded, approaches his observations objectively, or is concerned for others). It would be unwise to rely on attitude scales to assess achievement of this objective, since we would essentially be 'grading' the pupil and he may well be motivated to fake his responses in order to achieve a high 'mark' for *saying* he is open-minded or objective.

On the other hand, suppose we have developed a new curriculum with an objective of making science more 'interesting' to pupils. While it may be desirable that pupils are interested in science, it is not at all clear that it is useful to grade individuals on 'interest'. We may well be more concerned to compare *groups* of pupils following different curricula, or taught by different teachers or methods than to assess individual pupil interest levels. In the absence of established behaviours that can be used as criterion measures of interest, it would seem that the most satisfactory measure would be the anonymous self-report device rejected by Popham and Baker. It is suggested, therefore, that if the attitude scale is administered in a situation that is seen by the pupil as involving neither reward nor punishment, then the expressed attitude is the most satisfactory indicator of interest that we presently have. This does, however, assume that we are concerned with 'interest' in the sense of a pupil feeling positive, in a general way, about the course. The teacher may be viewing interest in terms of the pupil electing to take a further course in science, or joining the science club, and in that case overt behaviour (enrolling in the course, joining the club) will provide the criterion measure.

Teachers are frequently hesitant in accepting pen-and-paper measures of attitudes since they feel inherent attitudes are not being elicited. It has already been pointed out that there is no single accepted hypothetical construct for an inherent attitude, and until a theoretical formulation of such a construct is provided tests having construct validity cannot be devised.

The Level of Specification of Affective Objectives

Initial formulations of goals for curriculum innovations, frequently stress affective objectives. These objectives are usually specified at an intermediate level rather than at the highly specific level needed for the development of instructional materials and specific lesson plans. It is not possible to identify all the situations of behaviour covered by this type of objective, as it is for the 'mastery objective', but it is possible to specify some. Harbeck (1970) in the context of science education suggests that these would represent some 'minimal acceptance of the values of science'.

Attitudes and value-systems are demonstrated by the complex inter-relationship of numerous observable behaviours. These behaviours in isolation may be trivial and not indicative of an attitude.

An objective:

> 'to appreciate the economic implications for the community of scientific and technological progress,'

may be broad, ambiguous and difficult to evaluate. It is not easy to translate into instructional methods for the classroom, but it may well be an important and relevant objective. The much tighter objective:

> 'to appreciate the economic implications for the community of scientific and technological process as evidenced by the naming, from memory, of five inventions of the last fifty years that are now in general use as home appliances,'

leads easily to instructional procedures and evaluation techniques, but may well be meaningless in the development of the desired attitude.

Hirschlein and Jones (1971) suggest that the high level of objectivity demanded by instructional objectives may be inappropriate and inhibiting in the affective domain. Eiss and Harbeck (1969) suggest that a middle path which loosely defines the type of behaviour required may avoid the vagueness of the first example and the restrictions of the second. Our objective might then be written

> 'to appreciate the economic implications for the community of scientific and technological progress by suggesting examples of scientific inventions of the last fifty years that he considers have had a) a favourable and b) an unfavourable, impact on the community and to present clear arguments for his views.'

However, this is still at a very general level for the teacher who has probably very little idea of how such an attitude may be developed, and the cognitive demands (on teacher and pupil) are much greater than for the 'instructional' objective.

Classification of Affective Objectives

Taxonomy of Affective Objectives - Krathwohl et al.

The lack of precision and clarity of many affective objectives in school curricula makes it difficult to know what was, or was not, intended by the writer, and it is sometimes impossible to make inferences about the behaviour expected of pupils. In these circumstances comparisons of objectives of different curricula cannot be made reliably and considerable problems confront the evaluator who wishes to determine whether or not there are evaluation instruments already developed that are suitable for his particular objectives.

The Taxonomy of Educational Objectives: Affective Domain (Krathwohl *et al.* 1964) was developed to try to deal with these problems by producing a classification scheme with clear and meaningful terminology.

> 'Here it was hoped that placing the objective within the classification scheme would locate it on a continuum and thus service to indicate what is intended' (page 4).

In addition, an attempt was made to construct a scheme capable of revealing any underlying order among affective objectives.

The Taxonomy is seen by its authors as providing

> 'a bridge for further communication among teachers and between teachers and evaluators, curriculum research workers, psychologists, and other behavioural scientists' (page 23).

They also express the hope that any order revealed among the objectives may lead to a better understanding of how affective behaviours develop.

The Taxonomy deals with objectives involving attitudes, values, interests and social-emotional adjustments. These terms have a far wider range of meaning and so lead to even greater confusion in communication than do the corresponding terms in the Cognitive Domain. For example an objective of 'interest in science' can describe behaviour in which the pupil is aware that science exists and is not unwilling to attend his 'required' science lesson, or it can describe behaviour in which he goes out of his way to indulge in scientific activities and feels highly positive about science and a possible future career in that area. A 'favourable attitude to science' can mean anything from a willingness to say that he likes science, to a stated commitment to science with a clear conceptualization of its value in relation to other aspects of the environment. The confusion is exacerbated by the overlap of these terms. Frequently 'interest in', 'appreciation of', 'favourable attitude to', 'ability to value' are synonyms for a positive affect towards science.

The plethora of meanings for these terms precluded their use by Krathwohl in forming a hierarchical structure for the Affective Taxonomy. The characteristics of affective objectives that such a hierarchy should encompass were seen as:

> 'the emotional quality which is an important distinguishing feature of an affective response at certain levels of the continuum, the increasing automaticity as one progresses up the continuum, the increasing willingness to attend to a specified stimulus or stimulus type as one ascends the continuum, and the developing integration of a value pattern at the upper levels of continuum' (Krathwohl 1965),

and it was clear that the 'simple to complex' structure of the cognitive hierarchy would not be appropriate. This process of learning was seen as best described by the process of Internalization. This, it is claimed, refers to

> 'the inner growth that occurs as the individual becomes aware of and then adopts attitudes, principles, codes, and sanctions which become inherent in forming value judgments and in guiding his conduct' (Krathwohl 1965),

and led to the following category and sub-category system (Krathwohl *et al.* 1964 pages 34-35, examples are provided by this author).

1.0 Receiving: The pupil attends to a stimulus.

1.1 Awareness: Implying that the student is aware of various stimuli *e.g.* 'the pupil becomes aware of the various activities in the laboratory that constitute his science course'.

1.2 Willingness to Receive: Implying an ability to distinguish between, and a willingness to attend to, different stimuli *e.g.* 'is willing to consider the particular purposes of the various activities of the science lesson'.

1.3 Controlled or Selected Attention: Implying an ability to differentiate between stimuli and to control or select attention to one type *e.g.* 'looks for examples of crystalline substances among a large number of solid samples arranged in front of him'.

2.0 Responding: The pupil actively attends to a phenomenon.

2.1 Acquiescence in Responding: Implying compliance with instructions *e.g.* 'completes the assigned experiment'.

2.2 Willingness to Respond: Implying a voluntary exhibition of behaviour *e.g.* 'carries out simple experiments in chemistry at home'.

2.3 Satisfaction in Response: Implying a voluntary and emotional response *e.g.* 'gains considerable personal satisfaction from reading science magazines'.

3.0 Valuing: The pupil exhibits the consistency of behaviour that results from the ability to value the object.

3.1 Acceptance of a Value: Implying that worth is consistently ascribed to the object but their belief is, as yet, tentative. For example 'recognizes the need for the layman to have an understanding of science and technology'.

3.2 Preference for a Value: Implying identification with, and pursuence of a value *e.g.* 'prefers to conceptualize scientific information in terms of general principles rather than by memorisation of facts'.

3.3 Commitment: Implying a deep involvment in a belief at a high level of certainty *e.g.* 'holds firm conviction of the "integrated" nature of science'.

4.0 Organization: The start of the construction of a value system.

4.1 Conceptualization of a Value: Implying a quality of abstraction (associated with the high level cognitive behaviours of analysis and synthesis) that permits an individual to see how the values he holds are inter-related *e.g.* 'develops a rationale as to the responsibility of the scientist for the use of nuclear weapons'.

4.2 Organization of a Value System: Implying that complex values are brought into an ordered consistent relationship *e.g.* 'judges the ideal distribution of responsibility among scientists, politicians, and the general public for the conservation of natural resources'.

5.0 Characterization by a Value or Value-Complex: The individual responds consistently to value-laden situations according to his own unique characteristics and his philosophy of life.

5.1 Generalized Set: Implying that the individual has an internally consistent set of values which predispose him to act in a consistent way in response to highly generalized phenomena *e.g.* 'changes his "models" of the physical world and society in the light of new evidence that shows cause for revision'.

5.2 Characterization: Implying a philosophy of life that tends to characterize the individual *e.g.* 'develops a philosophy of life based on rationality and individual freedom'.

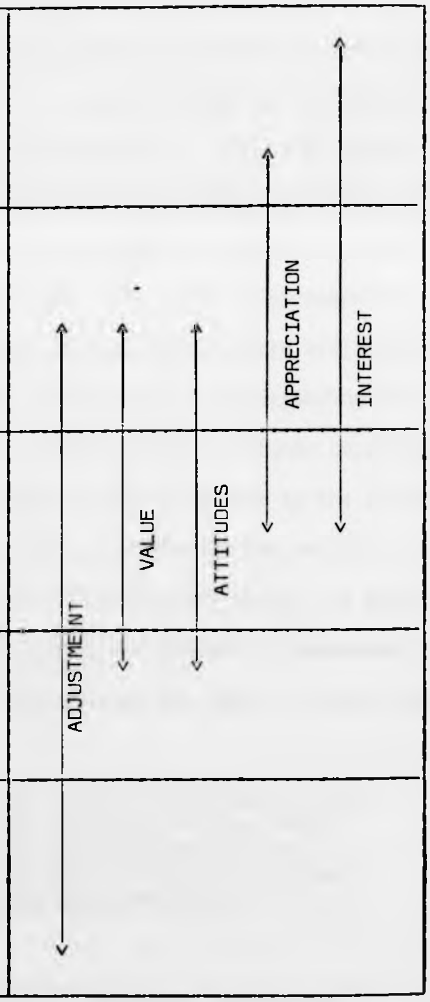
Krathwohl *et al.* (1964 pages 36-38) compare the terms used in this category system with the common usage in educational objectives of the terms interest, appreciation, attitudes, value and adjustment. The range and overlap of these terms is compared with the Taxonomy continuum as shown in Figure 2.1. They emphasise the resulting confusion in the middle range ('Willingness to Respond' to 'Preference for a Value'), and the absence of any common usage term that uniquely describes its area of use.

Empirical validity of the affective taxonomy

Lewy (1968) has attempted to investigate empirical validity of the classification scheme of the Affective Taxonomy. He looks at two aspects

FIGURE 2-1 THE RANGE OF MEANING TYPICAL OF COMMONLY USED AFFECTIVE TERMS MEASURED AGAINST THE TAXOMONY CONTINTUUM (REPRODUCED FROM KRATHWOHL ET AL, 1964, PAGE 37)

1.0 Receiving	1.1	Awareness
	1.2	Willingness to Receive
	1.3	Controlled or Selected Attention
2.0 Responding	2.1	Acquiescence in Responding
	2.2	Willingness to Respond
	2.3	Satisfaction in Response
3.0 Valuing	3.1	Acceptance of a Value
	3.2	Preference for a Value
	3.3	Commitment
4.0 Organization	4.1	Conceptualization of a value
	4.2	Organization of a value system
5.0 Characterization by a value complex	5.1	Generalized Set
	5.2	Characterization



of the model validity:

a) Descriptive - can we demonstrate the existence of empirical referents for its constructs?

b) Dynamic - do such referents display the structure and inter-relationship postulated by the model?

He constructed tests of about 80 items in each of the three fields (Mathematics, Music and Reading) at the four levels of Receiving, Responding, Valuing and Organization (insurmountable difficulties prevented formulation of test items at the fifth level of Characterization). Using a multiple-group factor analysis of the responses of 200 to 300 College and High School students he isolated four oblique factors. Two raters classified the 261 items according to the categories of the model (they were in agreement on 74% of items, adjacent categories were assigned on 25%, 1% were assigned to non-adjacent categories). He identified each factor as corresponding to a single taxonomic category such that 80% of the items had a higher correlation with the factor associated with their category (as determined by the raters) than with the other factors. The inter-factor correlation had ranges of 0.4 to 0.8 for Mathematics and Reading, and of 0.7 to 0.9 for Music. He argues that this demonstrates that the factors account for different components of the variance for Mathematics and Reading but not for Music. He concludes:

> 'The fact that empirical referents for the constructs of the model identified by the raters and by multiple-group factor analysis in two of the three tests indicates that for all practical purposes, the categories of the Affective Taxonomy can be regarded as possessing descriptive validity'.

He studied the dynamic validity in connection with the 'hierarchical structure' and 'inter-field generalizability'. In the former case he reports both internal and external evidence. The internal evidence he provides is, firstly, that:

> 'behaviour representing the Receiving category was exhibited on average by 66% of the students and in each test the percentage of students displaying a given behaviour decreased from lower to higher level taxonomic categories'.

(although one reversal was encountered in Music: Organization/Valuing). Unfortunately he does not make explicit what he means by 'exhibiting a behaviour'. He gives no sample test items (except two that were rejected: 'Reading poetry is a feminine activity' and 'A person who excels in Mathematics excels usually in logical thinking in general') and no indication of what the students did with the items. Was a Likert scale involved? Or an accept/reject procedure? Were all the items 'positive'? How was a mean score for each taxonomic level computed? Concurrent validation employed a Thurstone scale as a criterion variable in one case, how was this constructed? Secondly, Lewy looks at the relationship (postulated as inverse) between the hierarchical distance apart of two taxonomic variables and the positive correlation between scores on them. He reports three 4 x 4 correlation matrices exhibiting a 'simple structure' of hierarchically ordered variables, with the highest positive correlations in the diagonals next to the main diagonal and the lowest in the top right hand and the bottom left hand corners, thus verifying the hypothesis of an inverse relationship. Using a scaling method to carry this further and establish a quantitative measure of differences of difficulty levels within the taxonomy he finds that 'the four taxonomic levels in the tests of Mathematics and Reading reveal greater differences in their difficulty levels than they do in the test of Music'. However, he does not explain what 'difficulty' means in the context of a model of affective behaviours, it appears to be defined in terms of mean scores on taxonomic variables.

External evidence for the hierarchical structure of the Mathematics test was provided from a comparison of boys and girls responses in which the boys exhibited more positive affect (an established 'fact') than the girls, with scores more equal at the lower, than at the higher, taxonomic levels. This difference with taxonomic level is hypothesised on the assumption that:

> 'low level variables do not represent a high degree of involvement in, or internalization of, a particular phenomenon and hence differences in response to the variables are less likely to occur here than they are at higher levels.'

An examination of Music scores on a similar assumption, with girls expected to display more positive affect than boys, did not show the expected pattern of differences with the taxonomic level; equally large differences between girls and boys were demonstrated at all levels. A comparison of differences between scores of students (boys and girls) grouped according to the extent of their voluntary participation in activities related to Mathematics and Music, revealed no pattern providing evidence for the hierarchical structure. Lewy provides no extra evidence for the generalizability of the model across subject areas. The similarity of the results in Reading and Mathematics, inter-rater agreement and 'simplex structure' of the correlation in the Music test, and the support for a hierarchical structure of four categories in Reading and Mathematics and three categories in Music, lead him to the conclusion that the hypothesis of generalizability across subjects has been supported.

Overall he concludes:

> 'from analyses of the data, we conclude that the constructs of the model have empirical referents among affective educational objectives, and that the hierarchical structure of these referents corresponds to that claimed by the model.'

Internalization - a theoretical base for the taxonomy

The description of the Taxonomy category system has been set out in terms of behaviours and implied conditioning processes. Each category and sub-category appears to pre-suppose and include the one which precedes it, but the extra demands made by each successive sub-category cannot, in any immediately obvious way, be described in terms of any single general construct or dimension. It does not seem necessary (in view of the purposes of the Taxonomy) that they should be. The authors, however, claim that 'internalization' is such a construct. They state that:

> 'the more we carefully studied the components, however, the clearer it became that a continuum might be derived by appropriately ordering them' (page 27).

They view internalization:

> 'as a process through which there is at first an incomplete and tentative adoption of only the overt manifestations of the desired behaviour and later a more complete adoption' (page 29).

Their choice of this concept was based upon their view of its usefulness in defining a meaningful continuum of behaviour for the hierarchy and its accommodation and clarification of a large number of behaviours specified in existing affective objectives.

No precise definition of internalization is given. It is seen as having much in common with socialization, the difference being that socialization is concerned with overt behaviour conforming to the group's values, while internalization will result in either conforming or non-conforming behaviour in accordance with the individual's values.

Internalization implies that something from outside moves inside. The authors talk of 'external control by the environment' yielding to 'inner control' as we go up the hierarchy. What is the nature of these controls? On the same page (page 30) they talk about 'inner growth' as progression from emotionless perception, to active emotional response, to emotionless and routine internalized behaviours. What has been internalized?

To answer these questions the authors have looked to psycho-analytic theory. They envisage 'values' as being progressively internalized as we move up the affective hierarchy, and suggest that this may form part of the development and differentiation of the pupils superego and conscience (pages 31 to 39). They use considerable space (pages 39 to 43) in an identification of the levels of the Taxonomy with the Moral-Character study (Peck and Havinghurst 1960) stages of character development. These stages were proposed using the Freudian theory of personality development.

It is difficult to see how this concept of internalization will help 'to define operationally the kinds of task a teacher faces in this domain' (page 18). If the 'values' start by being external they cannot move inside the individual. They may be regarded as stimuli which evoke responses in the individual, these responses may then be viewed as manifestations of the individual's value-system. But nothing has been incorporated into the individual. As Morshead (1965) suggests it seems that the authors

> 'were not yet quite convinced that modification of behaviour could be adequately explained without postulating the existence of a mind allegedly located some place in the body'.

No matter what meaning is given to internalization in this context, the term implies a process rather than a construct for differentiating various processes or behaviours. Krathwohl *et al.* do not demonstrate how such a process relates to the various kinds of additional characteristics which differentiate successive sub-categories. They do, however, attempt to relate the Taxonomy to a three-process theory of attitude change (Kelman 1958). Kelman's theoretical framework proposes the three processes of 'compliance', 'identification', and 'internalization'. The first two concern the relationship of the teacher (agent of influence) with the pupil (recipient of influence), and the third the relationship between the change advocated and the individual pupil's value system. Krathwohl (page 33)

> 'uses the term "internalization" more broadly than does Kelman, referring to a process including a number of stages of internalization, where he refers to an internalized system as an end product'.

Kelman's compliance consists of a pupil apparently changing his attitude in order to achieve a favourable reaction or to avoid a punishment from the teacher. The change must be observable to the teacher as a behavioural change in the pupil. There may in reality, be no attitude change at all (Wagner 1969), the pupil may privately disagree with his verbally expressed opinion and so the 'change' may only last for the duration of the encounter with the teacher.

Identification refers to a pupil adopting the attitude of his teacher (or his peer group) in order to achieve or maintain a satisfactory relationship. He actually believes in the opinions he adopts but they are tied to, and dependent upon, the external source and so are not part of his own value-system. Kelman emphasises the relationship with the teacher:

> 'to the extent to which such a relationship exists, the individual defines his own role in terms of the role of the other. He attempts to be like or actually to *be* the other person'. (Kelman 1961)

Internalization in Kelman's scheme occurs when influence on attitudes is acceptable to the pupil since it is consistent with his own value-system. He can now perceive the new attitude as desirable because of the intrinsic worth of that attitude.

Krathwohl *et al.* try to identify affective objectives of the various levels of the Taxonomy with the different behavioural products that Kelman predicts for the three processes. They suggest that compliance 'corresponds' to the early part of the Taxonomy's internalization continuum (perhaps as far as Willingness to Respond). It may be that compliance is all that is necessary for the achievement of such objectives, but these objectives could be achieved, for example, through an identification process.

The outcomes of the identification process might be summarized as a) a change in pupil behaviour, and b) perception by the pupil of the new behaviour as desirable. The middle group of categories (Willingness to Respond as far as Commitment to a Value) can be identified with such outcomes, though there is no reason to suppose that the processes involved in achieving these objectives are necessarily those of identification.

Indeed, Krathwohl implies that pupil moves from Willingness and Satisfaction in Response through his *own* assessment of worth to Preference for, and finally Commitment to, Values. The attitude change is concerned with satisfaction which may, or may not, be dependent on inter-personal relationships.

Only for the process of internalization in Kelman's theory can the parallel with the Taxonomy levels of (Commitment, Organization of Value-Systems, and Development of a Value Complex) be clearly seen. This is possible because, unlike the other processes, internalization describes an intra-personal rather than an inter-personal process.

There would appear to be no one-to-one relationship between the three processes of attitude change and the affective objective levels, and Krathwohl has had to distort Kelman's theory to establish a relationship with the Taxonomy.

It is suggested that the idea of internalization is inadequate as a theoretical base for the Taxonomy continuum. Achievement of objectives of successive levels of the heirarchy would lead to a pupil progressing from compliant behaviour to behaviour characteristic of and controlled by, his own personal value system. The continuum is, therefore, characterized by the change from minimal to maximal autonomy on the part of the pupil. It may be that 'pupil autonomy' would be a less ambigious construct than 'internalization' to relate to the Taxonomy.

Even if a satisfactory, appropriate construct could be identified it is not clear how useful it would be to those concerned with affective objectives. Morshead (1965) is dubious of its value and wonders 'why did the authors of the Affective Domain find it necessary to construct a metaphysical support for their venture when it can stand very well without such a crutch?'

Klopfer's categorization of affective behaviours in science

Klopfer (1971) has attempted to classify affective behaviours that *science* pupils in secondary schools are expected to exhibit. He, after Krathwohl *et al.*, includes those 'objectives which emphasise a feeling tone, an emotion, or a degree of acceptance or rejection'. However, although he acknowledges the preoccupation of many science educators with this area, his scheme for 'Attitudes and Interests (H)'

> 'does not pretend to be a complete taxonomy of the affective domain as it pertains to the student's learning in science. While it would be most desirable to have such a taxonomy, the present lack of reliable knowledge and the primitive level of discussions about the affective domain in science education make it unlikely that an affective-domain taxonomy for science can be constructed at this time. About the best that now seems possible is a categorization of aimed-for and hoped-for attitudes and interests that are frequently stated by science teachers and curriculum builders' (page 576).

Klopfer asserts that 'it is already amply clear, however, that a student's attitudes and interests are always associated with cognitive elements'. The six sub-categories of his scheme reflect this view.

The first sub-category (H1) is 'Manifestation of favourable attitudes towards science and scientists' and the pupil is expected to exhibit this behaviour by speaking, writing and acting

> 'in ways which show that he places a positive value on the role of science in furthering man's understanding and that he gives due acknowledgment to scientists for their past and future contributions in this quest' (page 577).

Secondly, he proposes 'Acceptance of scientific enquiry as a way of thought' (H2) and suggests that the pupil will accept

> 'the processes of scientific enquiry as a valid way to conduct his thinking, his behavior in approaching a problem or novel situation will be sufficiently consistent for competent observers of his actions to describe him as "behaving like a scientist".' (page 577)

Klopfer warns, however, that the pupil must do more than carry out the procedures of scientific enquiry (Acquiescence in Responding on Krathwohl's scheme). He must demonstrate some personal conviction of the value of the enquiry process, such that he can be seen to be operating at least at Krathwohl's *Acceptance of a Value* level, and possibly at the *Commitment* or *Generalized Set* levels.

The third sub-category (H3) is 'Adoption of "scientific attitudes".' (page 578.) Here Klopfer is concerned with pupils exhibiting behaviours such as 'honesty, open-mindedness, self criticism, willingness to suspend judgment, and commitment to accuracy'. These Klopfer describes as 'professional standards, to which adherence by practitioners of scientific inquiry is expected by the scientific community'.

Fourthly, Klopfer suggests 'Enjoyment of Science learning experiences' (H4) which he justifies by saying 'There is strong psychological evidence that students learn better and remember longer when they find pleasure in the learning experience' (page 578).

'Developments of interests in science and science-related activities' (both the activities which the pupil does voluntarily on his own, and those related to ongoing science and the social interactions of science) forms the fifth sub-category (H5) and the concern here is with essentially transitory scientific interests.

The sixth is specific, vocational, long term, and probably only applicable to a relatively small group of pupils: 'Development of his interest in pursuing a career in science or science-related work' (H6).

Klopfer has developed a further main category of behaviours 'Orientation (I)' which is also concerned with attitudes to science as we have been conceptualizing them. He sees this category as complementing the 'Attitudes and Interests' category by providing essential cognitive elements and general perspectives on the relationship between science and society.

There are five sub-categories of 'Orientation' and

> 'a key term in four of these is "recognition", "realization", or "awareness". What is implied and intended by these terms is a certain sensitivity on the student's part to the *relationships* between science and other large areas of human endeavour and other ways of thought. Relationships are emphasized because these are the primary focus of the student's orientation, which enables him to perceive the enterprise of science and his study of science in a more meaningful manner' (page 579).

The first sub-category (I1) 'Relationships among and distinctions between various types of statements in science' emphasises the concern that the pupil should be aware of these relationships whenever he is carrying out a scientific enquiry or viewing science in a larger perspective.

The 'Recognition of the limitations of scientific explanation and of the influence of scientific enquiry on general philosophy' of the second sub-category (I2) is also seen as relevant to the pupil's view of science in the larger perspective, and Klopfer considers that

> 'almost every student can acquire some awareness of the relationship between the kind of thinking which he practices in science and alternative ways of construing the world' (page 579).

The third sub-category (I3) 'Historical perspectives: recognition of the background of science' focusses on the 'evolutionary character' of science

> 'The student's recognition that the past, present, and future development of science is a product of its own history and a reflection of the general culture of its time, gives him a historical perspective on the scientific enterprise' (page 579).

Fourthly 'Realization of the relationships among scientific progress, technical achievement and economic development', (I4) concerns the

> 'obvious influence of science on society [which] is seen in the changes in man's daily life brought about by technological applications of scientific principles and ideas'

and the reciprocal

> 'financial support that is available from public and private agencies often determines which research problems scientists investigate' (page 580).

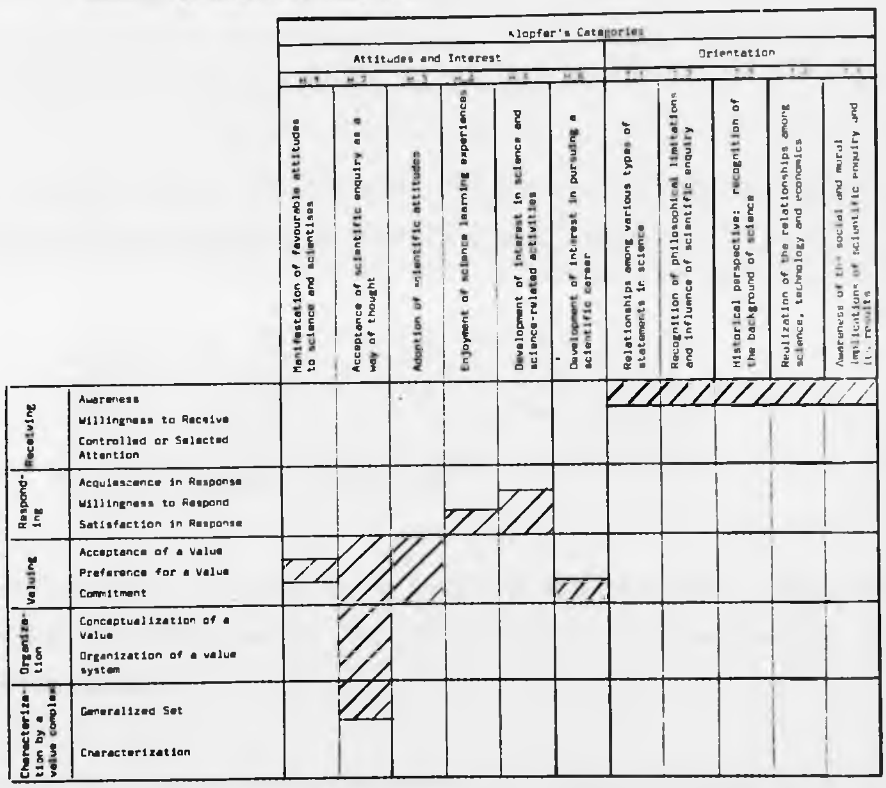
Lastly, the fifth sub-category (I5) 'Awareness of the social and moral implications of scientific enquiry and its results' concerns itself with issues such as population control, genetic engineering, organ transplantation and warfare research, and the subtle, fundamental influence such ideas and results of research have on society.

Two virtues of Klopfer's scheme are a) the way in which he has been able, by using a science-specific system, to emphasise those attitudes that relate to *science as a system of enquiry* and not just as a conceptual structure of a particular area of knowledge; and b) the way in which he has taken objectives, normally expressed in wooly terms (such as 'interest in science' 'scientific attitudes', and the 'relationship of science to other human endeavours or to society'), and carefully delineated the various areas of concern in an attempt to avoid ambiguity. The result is a scheme which appears to reflect faithfully the affective concerns of *science* education, and provides categories of objectives that are useful in that they are specific about such matters as the nature and context of the pupil's 'interest in' or 'enjoyment of' science, or about the particular aspects of 'human endeavour' or 'society' that the pupil is expected to relate to science.

The ways in which Klopfer's science categories broadly correspond to Krathwohl's subject-independent categories are outlined in Figure 2.2. Klopfer's first category (H1 - 'Manifestation of favourable attitudes towards science and scientists') is looking for

> 'positive expressions of feelings and, when occasions arise, actions supportive of science and scientist are wanted' (page 577).

FIGURE 2.2 RELATIONSHIP BETWEEN KLOPFER'S (1971) CATEGORIES OF AFFECTIVE BEHAVIOURS IN SCIENCE EDUCATION AND KRATHWOHL ET AL'S [illegible] TAXONOMY OF AFFECTIVE OBJECTIVES



This implies voluntary behaviour identifying and pursuing a particular value, and so appears to correspond to Krathwohl's Preference for a Value level.

It has already been stated by Klopfer that H2 - 'Acceptance of Scientific enquiry as a way of thought' may correspond to any of three of Krathwohl's levels the highest of these being Generalized Set.

For 'Adoption of Scientific Attitudes' (H3) there is again the implication of possible correspondence with several Krathwohl levels.

> 'The science student conducting inquiries is usually expected to imitate the scientist at work; and it is hoped that the habits of thought the scientist then displays will become a part of the students repetoire as well. If this has occurred, it will at some time be indicated in the student's actions and responses in novel situations' (page 578).

These behaviours may be only the the Acceptance of a Value level, though the reference to 'novel situations' implies Preference for a Value behaviour, or even Commitment.

'Enjoyment of Science learning experiences' (H4) appears to relate directly to the voluntary and emotional response of the Satisfaction in Response level.

H5 - 'Interest in science and science related activities' corresponds to a behaviour level at least as high as Willingness to Respond:

> 'A general criterion for a student's interest is that he does it voluntarily and without regard to the requirements of the science course...Some examples of behaviour which show that the student has interests in science activities ...are reading about new developments in solid state

> physics, watching a television program on cancer research, or circulating a petition for preservation of a wildlife refuge' (page 578).

If such behaviour is consistent and frequent then it may probably be inferred that there is Satisfaction in Response.

'Development of Interest in pursuing a scientific career' (H6) is explicitly described as behaviour at the Commitment level:

> 'If this interest is developed by a student his behavior in relevant situations...will show a commitment in the direction of careers or jobs in which science is involved' (page 578).

Klopfer's 'Orientation' categories (I1 to I5) are identified with

> 'The student's acquisition and understanding of some significant cognitive components that underlie or accompany general attitudes and interests in science' (page 577).

He appears to place all these behaviours at the low affective level of Awareness (see quotation page 76) but the *cognitive* processes implied may well be complex.

Affective Objectives of Curriculum Paper 7 - Their Relationship to the Classification Systems

Curriculum Paper 7, in the context of testing (page 27), refers to its reliance on Bloom's (1956) and Krathwohl's (1964) Taxonomies. This is reflected in the wording of some of the objectives. The first affective objective *'awareness of the inter-relationship of the different disciplines of science'* is apparently conceived at the lowest Krathwohl level of *Receiving - Awareness.* This awareness is

> 'almost a cognitive behaviour. But unlike *Knowledge,* the lowest level of the cognitive domain, we are not so much concerned with the memory of, or ability to recall, an item or fact as we are that, given appropriate opportunity, the learner will merely be conscious of something' (Krathwohl 1964, page 176).

If we now try to locate the objective along the Klopfer dimension we have the problem that the sorts of relationships with which the objective is concerned have not been specified. If they are related to scientific enquiry and are concerned with processes such as observation, interpretation and theory formulation, then the objective can be located in the first sub-category of 'Orientation' (*i.e.* 'Relationships among and distinctions between various types of statements in science') since Klopfer uses such processes as exemplification. If the relationships referred to are those among elements in a conceptual structure of that part of knowledge that is science, then it is not clear where the objective should be placed. It may be that this objective is more closely related to one of Klopfer's non-affective behaviours 'Application of scientific knowledge to new problems in a different field of science' (page 575). Similarly the second affective objective, *'awareness of the relationship of science to other aspects of the curriculum',* may be related to Klopfer's 'Application of scientific knowledge to new problems outside of science (including technology)' (page 576). However, there is no indication in Curriculum Paper 7 that 'Application' was what the writers had in mind for these objectives. The second objective might correspond to Klopfer's I2 category 'Recognition of the philosophical limitations and influence of scientific enquiry'. However, again the description of the category does not indicate any concern for relationships between conceptual structures of science and other disciplines.

This does, perhaps, identify an area omitted in Klopfer's system. He shows considerable concern for relationships where science exerts an influence on some aspect of contemporary society and history, or where society and historical forces influence science, but he has not provided categories concerned with relationships among, or distinctions between, the various elements making up the conceptual structure, or structures, of knowledge.

The third affective objective, *'awareness of the contribution of science to the social and economic life of the community'*, like the first two, corresponds to the lowest Krathwohl level. On the Klopfer scheme it appears to conflate the last two sub-categories, I4 'Realization of the relationships among science, technology and economics', and I5 'Awareness of the social and moral implications of scientific enquiry and its results'.

'Interest and enjoyment in science', the fourth affective objective conflates three Klopfer categories - H4 'Enjoyment of science learning experiences', H5 'Development of interests in science and science related activities', and H6 'Development of interest in pursuing a career in science'. As indicated in Figure 2.2 this implies behaviour at the three Krathwohl levels of Willingness to Respond, Satisfaction in Response and Commitment.

The fifth and final affective objective *'an objectivity in observation and in assessing observations'*, is clearly identified with Klopfer's H3 category 'Adoption of scientific attitudes'. As Figure 2.2 shows the behaviour may represent a Krathwohl level as high as Generalized Set, the minimum acceptable level will be Acceptance of a Value.

The first two of Klopfer's sub-categories have no referents in the objectives of Curriculum Paper 7. It may be the case that these were thought a) to be unimportant, or b) not feasible for the pupils to achieve, or c) that they would be achieved automatically. As the general arguments put forward for this curriculum (see Chapter 4) show (a) is unlikely. No case for (b) has been given, and the assumption has been made that categories such as the 'Adoption of scientific attitudes' are feasible. Much of the motivation for inclusion of affective objectives in new curricula has stemmed from the realization that affective goals were *not* automatically being achieved, (c) is therefore also unlikely. It appears then that objectives such as those of H1 and H2 were not considered.

The third Klopfer sub-category has a referent among the objectives, and a second objective can be related to any of the fourth, fifth or sixth sub-categories. The latter is indicative of the lack of precision of that objective. Unless some distinctions are made among enjoyment of and the various interests in science then the level of behaviour for which the curriculum writers were looking is indeterminate.

These categories are all concerned with the pupil's personal view of the relation between himself and science, an emphasis which is probably appropriate for the pupil just starting science when motivation and introduction into effective ways of working are important.

Three or four of the 'Orientation' sub-categories have referents among the objectives (one of these referents conflates two sub-categories). One or two sub-categories do not have referents, but it may well be that the cognitive demands of such objectives would be unrealistic for pupils of this age group. It is suggested that the difficulty in placing the first, second and third affective objectives of Curriculum Paper 7 unambiguously among the Klopfer Orientation categories is due to the lack of description of the *cognitive* behaviours that the writers had in mind for these objectives.

Curriculum Paper 7 has not provided a theoretical framework which may be used to clarify the purposes and meanings of the various attitude objectives. A comparison of these objectives with the framework mapped out by the work of Krathwohl and Klopfer, has revealed that there are some distinctions that need to be made, some behaviours (affective and cognitive) that should be specified, and some gaps that might be filled, if this is to become a comprehensive list of affective goals for science that may be readily related to classroom procedures.

CHAPTER 3

THE MEANINGS OF 'INTEGRATED SCIENCE' AND RELATED 'ATTITUDES TO SCIENCE'

The introduction of the Integrated Science course (Scottish Education Department 1969, Curriculum Paper 7) for the first two years of secondary education incorporated the first appearance of explicit attitude objectives for a national science course in Scotland. As discussed in Chapters 1 and 2, no explicit rationale for emphasis on attitudes was provided, although brief justifications were made for the inclusion of some of the objectives. It is possible, however, that such a rationale is implied in the rationale for the Integrated Science course as a whole.

The purpose of this chapter, together with Chapter 4, is to examine the ways in which a rationale for an integrated science course may be developed, the particular rationale presented by Curriculum Paper 7, and the extent to which an integrated course might be expected to be more, or less, appropriate for the achievement of attitude objectives than a separate science subjects course. Following this general discussion, various hypotheses will be developed concerning the extent to which the Curriculum Paper 7 course would be expected to lead to different achievement of various attitude objectives than would the alternative Biology, Chemistry and Physics courses for this age group (Scottish Certificate of Education Examinations Board, 1968, 1969a, 1969b).

The particular attitude objectives that we will be concerned with are those laid down for the 'First Cycle' (S1 and S2 science) in Curriculum Paper 7, page 16.

i.e. That pupils should acquire:

'9. awareness of the inter-relationship of the different disciplines of science

10. awareness of the relationship of science to other aspects of the curriculum

11. awareness of the contribution of science to the economic and social life of the community

12. Interest and enjoyment in science

13. an objectivity in observation and in assessing observations.'

This course is one example of the widespread interest shown in recent years in integrated approaches. Unfortunately, the *meaning* of 'integration' has not always been made clear, and the *value* of such approaches has not necessarily been self evident. These two aspects will be considered in this chapter and in Chapter 4 respectively.

The concern of this chapter, then, is to examine the various meanings ascribed to integration and the extent to which Curriculum Paper 7 is concerned with integration in these senses. The possible ways in which the various meanings could relate to 'attitudes to science' will be discussed, and the extent to which Curriculum Paper 7 is concerned with such attitudes will be considered.

The Meaning of Integration

Pring (1971a) observes that 'In any knowledge whatsoever there must be some sort of integration, of seeing otherwise unrelated events as events of a certain kind, of structuring our experiences by means of concepts'. In discussing integration then, we must be aware of the particular context or the 'scope' (Blum 1973). Are we considering integration within a traditional school subject discipline, or across a range of disciplines or fields of study ? Pring (1971a) sees the meaning of the *word* integration as logically implying a 'unity of parts that are in some way transformed'. In other words integration of physics, chemistry and biology would provide a curriculum with something *extra* and *different* from that provided by an interdisciplinary study of the three components in parallel. Blum (1973) has a much broader interpretation of the word integration and in addition to the dimension of 'scope' adds one of 'intensity'. Along this dimension we progress from 'co-ordination' (independent subject programmes taught simultaneously), to 'combination' (with major units organized round headings taken from the different disciplines), to 'amalgamation' (a particular 'issue' forming the unifying principle).

If the ideas that form the basis of an integrated science curriculum are to be translated into practical recommendations for teaching, it is necessary that a precise description of what it is about science that is unified, and what it is that is diversified, be provided. Various conceptions of 'unity' will now be considered.

Unity of all knowledge

The meaning of integration that implies the greatest scope is that which claims the unity of all knowledge (Pring refers to this as the 'strong thesis'). Rutherford and Gardner (1969-70 page 47) suggest that 'As a matter of faith, rather than as a conclusion based on evidence, most scientists and, indeed, most individuals, believe that in some sense the natural world is of a piece'. Such a view would see our present system of describing our universe in terms of a number of different disciplines as indicative of the limitations of our present knowledge and procedures, rather than as evidence of any inherent disunity in knowledge.

Blum (1973) points to scientists from Aristotle to Einstein who 'believed in the unity of the universe and tried to discover the unifying laws of nature'.

Pring (1971a) suggests that in addition to the 'uniformity of knowledge' aspect, some curricula view integration as a 'balance between different kinds of knowledge'. He quotes Schools' Council Working Party Paper No. 11 (1967), the Crowther Report (1959) and the Newsom Report (1963) which are concerned with balance between parts and total pattern of the curriculum. These reports, however, do not provide any criteria that specify what would constitute this balance or pattern.

The appeal of this holistic view of knowledge is pervasive and 'it is worth noting that there frequently lies beneath this search for integration an emotional attachment to unity which might arise from, or might even beget, the belief that knowledge is essentially one and undivided, and it is this sense of unity which must be reflected in the curriculum of our schools' (Pring, 1971a).

However, the view of knowledge as a unity may have a more substantial basis; for example, claims for this sort of unity of knowledge might be made in terms of a reductionist theory. Here it would have to be argued that some discipline had priority and others had to be reduced in turn to this fundamental form. Thus in hierarchies such as Compte's we have mathematics as some sort of natural logic in terms of which we can describe all the findings of physics; findings of chemistry can then be deduced from principles of physics; the characteristics of biological organisms can be seen as complex physico-chemical systems; psychological characteristics can be seen as organizations of biological systems; sociological phenomena can be seen as aggregates of psychological systems; and so on. Such a system may be superficially attractive but Schwab (1964a) regards it as 'tyrannical and largely unexamined' (particularly when used to determine curriculum *sequence* in schools).

Popper's (1972) view is that while the reduction of chemistry to physics may well be demonstrated in the future, this may be less easy for biology in view of the apparent basic difference between 'living' and 'non-living' material. He critizises the view that the possibility of reduction is settled *a priori*, and insists that it must be settled by substantial theory.

Curriculum Paper 7 provides some statements that imply integration viewed in terms of the inherent unity of all knowledge and life. For example they are concerned with science as part of a whole general curriculum (page 10), as an explanation of 'ourselves and the natural phenomena with which we are surrounded' (page 11), as a control of 'our morals, our ethics and our whole cultural environment' (page 13), and as providing a general training in 'thinking' (page 12).

Unity of the disciplines of science

On a less global basis Curriculum Paper 7 is concerned that the 'unity of science should be made clear' (page 18). This implies a particular view of science as a field of experience distinct from other such fields, and without any clearly distinguishable component disciplines. Pring (1971a) refers to this as the 'weak thesis'. However, the basis on which the distinctions between fields or between disciplines are judged is not made clear. This unified approach implies a particular conceptual structure ('[this] approach to science. . . teaches some fundamental concepts', page 25), a context in which the pupil will integrate his own experiences by his own enquiry ('science is admirably suited to discovery methods', page 22), and a uniformity in the procedures and methodology of the scientists ('an attempt has been made to expose pupils to many aspects of the work of the scientist, the apparatus at his disposal, the experimental methods he uses, the different processes of thought by which he arrives at his conclusions and the language which he uses to communicate these conclusions to others', page 10).

There is no description of the conceptual framework on which this unified science is supposed to stand. No account of a general process of enquiry in science is given. The generalizable characteristics and methodologies of scientists are not enumerated. We are not told on what basis the topics for study are selected.

Unless explicit arguments in this area are provided it is unlikely that teachers will know in what sense the writers of the Curriculum Paper are viewing the unity of science. Arguments of this type would have to be made at two levels:

a) epistemological arguments about the structure of knowledge,

b) arguments about the desirable structure of the curriculum involving considerations other than the philosophical.

For example Hirst (1967) argues that 'All knowledge is differentiated into a limited number of logically distinct forms or disciplines', but asserts that:

> 'To proceed from saying that there is a given structure in the knowledge we wish to be mastered to saying that this must be the structure of the curriculum is to be guilty of a simple logical fallacy... the characteristics of the means must not be taken for the characteristics of the ends. I see no reason why the curriculum should not be fully topic-organized provided it is understood that the development of understanding involves the mastery of conceptual structures which are not reflected in the topic-organization'.

Hirst's logical forms of knowledge are distinguished by a) their use of concepts of a particular kind, and b) their distinctive types of test that they use for their objective claims (Hirst and Peters 1970, page 62). Whitfield (1971) suggests that a crucial distinguishing feature of the different forms is the degree of objectivity and publicity of their tests and criteria for truth.

Hirst (1965) sees the characteristic concepts of a particular form of knowledge as being related to each other in a limited (though possibly complex) number of ways, thus forming a characteristic network of relationships between the concepts. The decision on where the dividing line between disciplines is to be found requires examination of the concepts, the relationships between the concepts (the logical structure), the testing procedures used for testing statements emanating from the logical structure, and the criteria of truth used in testing.

Hirst distinguishes between physical sciences and human sciences. Holroyd (1972) and Lawton (1973) both point to ambiguity in the placing of biology. Does Hirst place it with physical sciences concerned with 'abstract theoretical concepts' but 'employing concepts of what is seen, heard or smelt' and using empirical tests of observation by the senses? In which case 'physical science' might be equivalent to the common-usage 'natural science'. Or does biology fall with the social sciences into the form of knowledge associated with understanding 'our own and other peoples minds'? The human sciences are concerned with additional empirical concepts connected with the *intentions* of living things (*e.g.* 'deciding', 'enjoying', 'wanting') and the basis on which we can make objective judgments in this area are, as yet, inadequately understood.

Hirst emphasises that each of the forms of knowledge borrows from other forms. For example science uses mathematics, but such knowledge is a *tool* for science not an integral part of it, since the validity of such knowledge is established by mathematical, and *not* scientific procedures and tests.

Phenix (1964) takes a rather different perspective from Hirst in classifying human Realms of Meaning. He is concerned with the need of the individual for understanding over a number of general areas if he is to function adequately in society. One of these areas or Realms of Meaning - Empirics - includes physical science, biological science, psychology and the social sciences. Empirics are concerned with those truths which may be framed on an *experimentally* verifiable conceptual framework.

As already intimated, Hirst sees no reason to limit our *classifications* of knowledge to those of the distinct 'forms'. These 'forms' will provide the fixed reference points but he sees considerable importance in 'fields' of knowledge such as geography and engineering. Fields are not distinguished by their particular conceptual frameworks but by their subject matter. The various 'forms' contribute to the *particular* theoretical or practical interest of the 'field'. If we are to see integrated science as a 'field' it would be necessary to specify the ways in which it straddles the different forms *i.e.* the *particular subject matter* to which the distinguishable 'forms' of science knowledge would contribute.

It is very difficult to decide on logical grounds where boundaries in science should be drawn. It may be that it is *methodological* distinctions that should be investigated.

Rutherford and Gardner (1969-70) suggest that for some science educators the concept of integration is based on general agreement among scientists on:

1) Appropriate language for stating scientific problems.
2) How to collect and analyse data.
3) Correct application of logic.
4) Use of theory and models.
5) Need for verification.
6) Need for critical appraisal by colleagues.

These, it would be argued, are more significant than the particular facts, laws and theories that comprise the separate sciences. However, if a curriculum were to be based on such an interpretation it would be necessary to be explicit about the nature of the 'use' of theory, or the 'verification', or the 'appropriate language'.

Holroyd (1972) has looked at the way the scientist uses his theories and observations, and concludes that the distinction between enquiry in science and other enquiry is that theories must be 'testable by experiment and that its practitioners attempt refutation rather than verification'.

On the other hand Schwab (1964a, 1962) uses physics and biology to exemplify *different* syntactic structures. By the syntax of a discipline he is referring to its patterns of procedure and methodology of reaching its goals. He sees biology as verifying the consequences of a large number of hypotheses *i.e.* seeking knowledge in 'bits and pieces'. Physics he views as a verification of the structure of existing knowledge, as developing broad comprehensive theories concerning large areas of subject matter. These different goals, will, he suggests, require different sorts of evidence and so different pathways of enquiry.

Robinson (1969) reports on a study he made of major concepts of methodology of six philosophic works by three biologists and three physicists. He found:

> 'The biological sciences were considered to utilise predominantly correlational procedures. Selected sense data are conceptualized into statements that attempt to explain the organization, history, and apparent directiveness of living systems. Many of the concepts within the explanatory statements provide sufficient conditions, but not the necessary conditions demanded by the exact procedures.'

In contrast the physical sciences were concerned with constructs that are:

> 'linked logically with other constructs and empirically to sense data through rules of correspondence......the spontaneity, relative independence and irreductibility of sense data. . .are brought under methodological control by the establishment of correspondences between sense data and constructs and the subjection of constructs to the demands of metaphysical principles through the circuit of verification. The circuit of verification includes inductive and deductive procedures.'

However, Schwab emphasises the changing structures of the disciplines and in another article (Schwab 1964b) he develops descriptions of two generalizable, syntactical structures for the natural sciences. Firstly, *syntax of stable enquiry,* which is concerned with problems arising from the prevailing substantive structure of science. This he characterizes by:

a) Formulation of a problem.

b) Search for data to suggest possible solutions.

c) Reformulation of problem to include solutions.

d) Determination of data necessary to solve the problem.

e) Plan of experiment to elicit data.

f) Execution of experiment and accumulation of desired data.

g) Interpretation of data by means of guiding substantive structures of previous knowledge.

Secondly, the *syntax of fluid enquiry*, he sees as a response to the demand for increasing validity of existing substantive structures. This follows no systematic pattern (and as such would be very difficult to build into the curriculum) but must deal with perceived inadequacies and weaknesses in the existing conceptual structures, and produce new or modified ones that can be seen as adequate and feasible.

Schwab (1962) is anxious that instruction should a) lead the pupil to see science as enquiry (*i.e.* integrated in the sense of describing generalized procedures and methodologies of scientists), and b) enable the pupil to enquire into science (*i.e.* integrated in the sense of taking part in the generalized process of scientific enquiry).

It has already been pointed out that Curriculum Paper 7 may imply an underlying, specific model of scientific enquiry but this has not been clearly described. Much of the writing on integrated science lacks this sort of specification of the processes of science. For example Jacobson (1969-70) suggests that 'there are different modes of enquiry in different sciences' (page 69) but suggests that 'fundamental concepts and basic science process skills can be developed from within a wide range of content fields from meteorology to physiology.' He may be right, but that will depend on what is *different* about the modes of enquiry and what the *fundamental concepts* and *basic process skills* are. The modes of enquiry are not described; the fundamental concepts are exemplified by the conservation of matter and energy (this example is used with such regularity one wonders whether there are any others); the basic science process skills are exemplified by model formation in solar system astronomy, genetics, atmospheric physics, electricity and atomic physics. Model formation is a vague term and one which is applicable to other areas of knowledge (plastic arts, metaphysics, social sciences). Closer specification is needed - are we concerned with models in the sense of the 'black-box hypothesis' (Schwab 1964b) which is inaccessable to immediate observation (*e.g.* model of matter incorporating a neutrino), or in the sense of a 'glass-box hypothesis' which poses a problem that can be answered empirically (*e.g.* model of the solar system incorporating a new planet), or in some other sense?

Interdisciplinary study

A third possible interpretation of integration is that of *interdisciplinary* enquiry. Whitfield (1971) distinguishes integrated science as a blending of subjects and deliberate synthesis of material, from interdisciplinary science as a collaboration between subjects and a viewing of a topic or theme from logically different viewpoints (with the learner left to synthesise in any way he pleases). Curriculum Paper 7 implies an interdisciplinary approach when it states that where integration 'is not easily possible, we have considered it better to deal with the content as a single discipline' (page 18), and 'physics, chemistry and biology should have an equal share of the time allocated' (page 17). This interdisciplinary pattern of viewing individual topics from more than one perspective is (unlike the 'strong' and 'weak' theses) reflected in the subject matter of nine of the fifteen syllabus sections (two sections are designed for a 3-perspective approach, seven section for a 2-perspective approach), while the other six are single subject sections. Such an approach might correspond to Blum's 'combination' (see page 88), and since no synthesis of established disciplines is involved in this approach, epistemological arguments about the nature of the concepts and processes of science are not necessary. However, it is necessary to specify the basis on which the subject matter is selected, if we are not be be left with what Whitfield (after Toynbee) calls 'one damned thing after another'.

Hirst and Peters (1970) have stressed the importance of the inter-relationships between the different forms of knowledge (with the proviso that 'these inter-relations must not be thought to weaken in any way the claims for independence'). They suggest that where we have topic- or theme-centred curricula with objectives from different disciplines, there are great demands on the knowledge and ability of teachers to bring out the interrelationships. And 'if the objectives from the different domains are not being adequately related to the structures within each of these, little is likely to be achieved.'

Whitfield (1971) also warns of the daunting tasks a) of choice of topic which will group together objectives having a significant educational relationship to each other, and b) of providing pupils with the necessary initiation into the distinctive ways of knowing and testing of the various disciplines that will enable them to inter-relate the knowledge and apply it meaningfully to the topic.

Attitudes Logically Related to Meanings of Integration

It has been suggested that Curriculum Paper 7 implies, at different points, each of the three meanings of integration that have been discussed. The attitudinal emphases that are logically implicit in these three meanings adopted by the curriculum planners will now be examined.

If the meaning of integration in the 'strong thesis' sense is accepted, a view of science is implied in which science is an integral part of the conceptual structure of all knowledge. Thus, any development of scientific knowledge would be expected to affect all other areas of our experience. This implies attitudes which reflect knowledge of, and feeling towards, firstly, the pervading influence of science on our culture, ourselves and our environment, and secondly, science as part of the unified structure of knowledge.

An attitude of the first type is exemplified in Curriculum Paper 7. The planners are concerned that pupils should acquire an *awareness of the contribution of science to the economic and social life of the community* (page 16). Further, they explain that:

> 'we have included as our final aim our desire to expose pupils to this cultural aspect of science. While realizing that there is much which will be quite beyond the young people with whom our report deals, and conceding the great difficulty of building this aim into a syllabus, we feel that there will be many opportunities for discussing such cultural implications with pupils during class time' (page 13).

The second type of attitude (towards science as part of unified knowledge) is reflected in two of the objectives of Curriculum Paper 7, *i.e.* that pupils should acquire an *awareness of the inter-relationship of the different disciplines of science* and an *awareness of the relationship of science to other aspects of the curriculum* (page 16).

If we consider integration in the 'weak thesis' sense we imply a view that sees science as a discipline which differs from other forms of knowledge, but which is without component disciplines which are clearly distinguishable. The logical attitudinal implications are narrower than for the 'strong thesis', and the concern is limited to the area *of the inter-relationship of the different disciplines of science* (page 16).

The meaning of integration that is synonymous with 'interdisciplinary study' implies no particular attitude towards science other than as a collection of established disciplines.

CHAPTER 4

ARGUMENTS FOR AN INTEGRATED APPROACH TO SCIENCE AND THE IMPLICATIONS FOR ACHIEVEMENT OF ATTITUDE OBJECTIVES

As explained in Chapter 3, this chapter will be concerned with the *value* of presenting science in an integrated form. The various sorts of argument that are put forward on the value of an integrated approach will be reviewed, together with an account of the emphasis laid by Curriculum Paper 7 on the different areas of argument. These arguments are generally in terms of a) the sorts of *outcomes* that are anticipated from such a course, and b) the sorts of *constraints* under which the course must be implemented. The extent to which such *outcomes* may be expected to emphasise attitudes, and such *constraints* may be expected to constrain the achievement of attitude objectives will be considered.

Arguments for an Integrated Science Curriculum

The views on the structure of science that have been described in Chapter 3, have a bearing on the content and structure of the science curriculum that would be chosen in any particular situation. However, the decision on whether to adopt this curriculum or that curriculum will also depend on other factors such as the current demands and constraints imposed on the education system by society, current theories of learning and psychological development, and the availability of appropriately trained teachers.

If a new educational practice is under consideration it is necessary that we ask certain questions, the answers to which will form the basis of the rationale for the practice. Peters (1962) has suggested that these may be questions of *technology* (what are our *requirements*?), questions of *policy* (what *ought* we to be teaching towards?), questions of *theory* (what *is* the case?) or questions of *philosophy* and *metaphysics* (what are the *rules* of the game?). O'Conner (1957) sees the statements that form the basis of any educational practice (*i.e.* answers to the questions) as composed of three types: *empirical* claims (answers to questions of theory and partly questions of technology), *value judgment* claims (answers to questions of policy and partly questions of technology), *metaphysical* claims (answers to questions of philosophy and metaphysics).

The arguments for Integrated Science will be examined and, where possible, classified according to these three types. However, though it may be important to be aware of the extent to which the rationale is based on reliable empirical evidence, on value judgments, on substantial logical argument, or on myth, it is also of interest to see what types of factors are most frequently influential in determining the form of the curriculum. For this purpose a classification system (Figure 4.1) has been developed. This is a modification of four groups of determinants of the structure and content of the curriculum identified by Schwab, Dunkel and Tyler (Schwab 1962). They saw these determinants as i) the social milieu (corresponding roughly to the first three categories of Figure 4.1), ii) learner factors, iii) teacher factors, and iv) subject factors, (these last three correspond roughly to categories 4, 5 and 6 on Figure 4.1).

1.	Outcomes demanded by society, *e.g.* provision of scientists, informed lay population.

2.	Resource constraints, *e.g.* accommodation, equipment, time, teachers.

3.	Political constraints, *e.g.* 'O' grade, common-core course.

4.	Conditions for effective learning, *e.g.* pupil security, motivation, interest.

5.	Conditions for effective teaching, *e.g.* teachers' background, interest, competence.

6.	Constraints imposed by subject, *e.g.* unified nature of science.

FIGURE 4.1 Classification of arguments for an Integrated Science Curriculum

Arguments in terms of outcome demands made by the social milieu

These demands are initiated by the needs of the social structure. They are largely determined by economic problems (*e.g.* the need for trained manpower, the pool of unemployed graduate scientists, distribution of energy resources), environmental-economic problems (*e.g.* pollution), political problems (*e.g.* the 'cold-war') and political ideologies (*e.g.* equal opportunities for all pupils). Identification of these demands depends on *value judgments* by politicians, economists and administrators about what our *requirements* are. Like the problems and ideologies which determine them, these demands are transitory and provide our instrumental aims for science education at this point in time. They may be idealistic in the sense that they may not take cognizance of the practicalities of the classroom or the structure of the knowledge to be imparted.

Schwab (1962) has argued that the current basic demands on science education are for:

i) trained scientists,

ii) a political leadership informed on scientific matters,

iii) a general public informed on scientific matters.

Curriculum Paper 7 implies a similar view, for example it is concerned that pupils should acquire 'a little of the vocabulary and grammar of science', 'an awareness of the culture which is science', 'an ability to solve problems and think scientifically' (page 11); and that pupils should be exposed to

> 'many aspects of the work of the scientist; the apparatus at his disposal, the experimental methods he uses, the different processes of thought by which he arrives at his

> conclusions and the language which he uses to communicate these conclusions to others' (page 10);

and that

> 'Knowledge is growing at an explosive rate and the effect of this growth is to alter....our morals, our ethics and our whole cultural environment....we feel that there will be many opportunities for discussing such cultural implications with pupils during class time' (page 13).

Whitfield (1971) examines three sources of educational objectives; one of these 'masters' of our system is 'society' and here it is citizenship which is emphasised. He argues that society has infinite variety and, though the education system must serve society, it would be dangerous to base curricula on social needs as perceived by some person (or persons). In reviewing the justifications made for integrated or interdisciplinary studies Whitfield observes that the 'topic' approach is advocated since topics are closer to everyday life, and the 'thematic' approach is advocated to inter-relate science education with aspects such as socialization and moral education through themes such as the 'family', the 'town', 'housing' or 'VD'. However, the precise rationale for choosing 'these' rather than 'those' topics or themes is seldom given.

Several other authors lay considerable emphasis on these broad outcome demands by society. Kirk (1973) considers that the strongest argument for integration is:

> 'that there are many highly controversial issues - race, sex, pollution, crime, for example - which are basically interdisciplinary issues. To be dealt with at all adequately they require the bringing together, of integration, of many different types of evidence - historical, economical, psychological, literary, moral or statistical'.

Rutherford and Gardner (1969-70) summarize one group of arguments for integrated science as being

> 'based on the claim that the process of education should prepare the student to cope with the world he encounters ... to learn how to seek information effectively... how to get answers for themselves about the world they experience'.

Jacobson (1969-70) suggests that integrated science 'may give students a more complete view of the nature of science as a human enterprise'. The superiority of the integrated approach for achieving such outcomes is frequently assumed. The question is, however, moot. It does seem reasonable to say that problems such as pollution, energy resources and racialism do not respect subject boundaries and therefore can only be solved by the integrated approach; or that scientists follow a generalizable pattern of working that can only be made clear by an integrated approach. However, Pring (1971b) describes a thesis with which the integrationist must deal if he is to take up this position. This thesis takes the position that 'enquiry' into some problem is not a *general* method, it must 'involve the meanings revealed at different levels within one or other of the disciplines'. These distinct disciplines (which may be the established school subjects) are characterized by the types of concepts with which they are concerned, their typical conceptual frameworks, their principles of verification (Hirst and Peters 1970), and 'the use of symbols which already dictate... what moves are correct or at least permissable'. Before enquiry into the problem the pupil must be 'initiated into the different modes of understanding' and this can only be achieved by exposure to the individual disciplines (subjects) within the curriculum.

The lack of an acceptable unique 'scientific method' at this time (see pages 94-97) militates against a course that sets out an integrated approach in order to display the general pattern of working for all scientists. However, it may be that the differences in methodologies of, say, the biologist and the physicist are too sophisticated or subtle to be important at the level of secondary school science. If that is the case it is still necessary to specify the characteristics of the generalizable methodology that is to be presented. Fiasca (1970) argues that 'concentration' of processes, skills and principles from various science disciplines, through an integrated course can lay bare the 'epistemology of the working scientist'. We do not yet have agreement on what that epistemology is.

It was interesting to find that in interviews with the science staff of 50 secondary schools in Central Scotland (McIntyre, Brown, Davies and Drever unpublished), very little emphasis was placed by teachers on these demands of society when stating what they saw as the advantages of an integrated course. In schools where they were mentioned at all, integrated science was seen as providing an appropriate background for those who would do no more science, and as a course more relevant to the pupil's understanding of his environment than a separate subjects course.

However, despite the teachers' lack of emphasis on this area, it has been suggested that Curriculum Paper 7 *is* concerned with such outcome demands. The Paper suggests (implicitly rather than explicitly) that such outcomes will be more effectively achieved through an integrated, rather than a separate subject, science course. If, for example, society is demanding that we have a political leadership, informed on scientific matters, then that leadership will be expected to have a realistic perception of the potential of scientific solutions for national problems. If we demand an informed general population who will support useful scientific enquiry, we need emphasis in science courses on the development of attitudes such that pupils appreciate the advantages to the community that acrue from science, while at the same time being aware of its limitations and dangers. Curriculum Paper 7 has such an attitude outcome among its objectives: an *awareness of the contribution of science to the economic and social life of the community* (page 16). As discussed earlier, it was the planner's intention that the integrated course be concerned with broad issues where science is seen as part of the social and cultural, as well as the economic, life (page 13).

On the other hand, the alternative separate subjects courses (Scottish Certificate of Education Examinations Board, 1968, 1969a, 1969b) appear to favour attitudes concerning the subject's importance (through technology) to the rather narrower economic sphere. The Chemistry syllabus is concerned that pupils acquire 'some appreciation of the part science has to play in the world economy' (page 5); and the Physics syllabus aims to encourage 'a sympathetic understanding of their environment and science', and attempts to show that Physics 'is a human endeavour relevant to modern technology, vital to our economy and having tremendous social impact' (page 7). The Biology syllabus places no explicit emphasis on the impact of science on society.

It seems, therefore, at least plausible that the integrated course could be expected to achieve attitude outcomes concerned with the broad influences of science on society more effectively than the separate science subjects courses.

Curriculum Paper 7 is concerned with the demand from society for trained scientists. It sees the need for training pupils to think scientifically (page 11) and for familiarizing them with the procedures and methodologies of the working scientist (page 10). If the integrated course is to be justified in terms of it being more effective in achieving such outcomes than are the alternative courses, then we would expect increased emphasis to be placed on the development of 'scientific attitudes', or the 'styles of thinking which scientists are presumed to display' (Gardner 1975).

Scientific enquiry is popularly characterized by attitudes such as curiosity, critical appraisal, open-mindedness, lack of bias and objectivity. Curriculum Paper 7 lays down one attitude objective in this area: that pupils should acquire *an objectivity in observation and in assessing observations* (page 16). There is no indication, however, of whether or not the curriculum planners considered that integration was a direct aid to 'objectivity'. If 'objectivity' is a characteristic of scientific enquiry, we would expect it to be manifest in any type of science course, separate subjects or integrated. The separate subject syllabuses are, however, less specific in their objectives than is the integrated course.

The general aim of 'a training in objective observation' (page 5) is laid down for the 5-year Chemistry course, but no indication of what level of achievement may be expected from first and second year pupils is given. The Physics syllabus is 'designed to assist teachers to encourage pupils in developing their own attitudes of enquiry' (page 5), but in the first two years it is only concerned that the pupils gain 'confidence in handling experimental situations' (page 13). The Biology syllabus makes no comment in this area.

The only basis for suggesting that the Integrated course is likely to be the more effective in developing 'objectivity' in pupils, is that the separate subjects syllabuses, insofar as they are concerned with such characteristics, are less explicit than Curriculum Paper 7 in statement of objectives. This might lead to more conscious emphasis on 'objectivity' in the integrated classroom, but the argument is tenuous.

Arguments in terms of resource constraints imposed by the social milieu

As well as demanding certain outcomes from teaching, society imposes *constraints* on *resources* of equipment, accommodation, time, technical and other ancillary services, worksheets and books, and teachers. It does this by the finance it makes available, by the status it gives the teachers and by the rules it lays down for the use of the resources.

Whitfield (1971 page 227) observes that integrated curricula can, superficially, cope with deficiencies of time allocation and imbalances between subjects. Fiasca (1970) sees integration as a way of avoiding time lost by unnecessary duplication of learning experiences that are common to the various disciplines. Curriculum Paper 7 is also aware of time limitations:

> 'For the time allocation which we have accepted, a division into three separate subjects would imply less than two periods per week for each. We consider this an inadequate time in which to develop a proper understanding of a subject if it is to be seen as a separate entity' (page 18).

However, apart from suggesting (page 23) that station experiments will make good use of the equipment available (station experiments are not, of course, specific to an *integrated* course), they appear anxious to avoid the impression that the Integrated Science Course is proposed in order to paper over cracks in the resources.

On the other hand teachers from the sample of 50 schools placed considerable emphasis on the advantages of an integrated course in departments where there was a shortage or imbalance of staff across the three science subjects. The *particular* resource problems of a given school will influence the favour with which integration is viewed, and evidence of such constraints is readily available.

The Integrated Science course has been justified, in part, by its more effective use of time. It has been suggested that within the same time allowance a separate subjects course would not lead to adequate understanding, and there would be unnecessary repetition of material among the disciplines and in the form of revision at the start of the infrequent lessons. While the integrated course is seen as relaxing some of the constraints on understanding and amount of material covered, there is no indication of how it might influence any constraints there might be on the achievement of attitude objectives.

Arguments in terms of political constraints imposed by the social milieu

The teacher and pupil in our educational system are controlled by an authority structure which incorporates at national level the Scottish Education Department, the inspectorate, the national assessment system for 'O' and 'H' grades, and the teachers' unions; at the local level we have the policies of Directors of Education, Education Committees and science advisers; at school level we have Head Teachers, their deputies and assistants; and at science department level we have Principal Teachers and Assistant Principal Teachers.

Teachers may argue for a particular curriculum because they see it as complying with the views of the inspectorate, or the science adviser, or the headmaster. They may justify it in terms of being in line with current national educational policies such as mixed ability classes, a common-core course, or equal 'O' grade opportunities for all pupils. They may see it as the most appropriate foundation for 'O' grade and later courses, or they may simply accept that 'this is the syllabus and must be followed'.

Curriculum Paper 7 is conscious of such constraints. It sees its remit as being the production of an 'introductory syllabus. . .to suit the needs of a common-core course to be taught to unselected classes' (page 9). Beyond this they found

> 'it was possible to produce a content in physics and chemistry, identical in wording for both the examination syllabuses and our integrated course. That biology is not also identical is only because of the form in which the biology syllabus was published by the Examination Board; the biology content of the integrated syllabus is all drawn from the examination syllabus and is material agreed as suitable for teaching to first and second year pupils' (page 19).

And at the school administration level they consider that 'For the headmaster it will surely be much easier to timetable one science teacher per set, than it would be to co-ordinate the efforts of three' (page 18).

Bernstein (1971) argues that introduction of an integrated curriculum will disturb the authority structures. In a separate subject curriculum[1] the established power heirarchy is isolated into the individual science departments. Junior teachers show allegiance to Assistant Principal and Principal Teachers through socialization into strong subject loyalties. The advent of the integrated curriculum will, Bernstein suggests, undermine this structure. This will be brought about by new work-based relationships between staff at junior levels in *different* departments which will weaken the heirarchies and the 'private property' aspects of the former individual subjects.

He also argues that boundaries between pupils, and between pupils and staff will be weakened. The situation in which D stream pupils do not have access to certain subjects and A stream pupils do not have access to certain other subjects, will have changed, and all pupils will have a common work task. Bernstein sees this as the implementation of a particular educational policy since 'The less rigid social structure of the integrated code makes it a potential code for egalitarian education'.

Musgrove (1973) looks at the influence of integration on the authority structure from a different perspective. He views individual subject specialization as an instance of 'division of labour' and suggests that 'respect for the autonomy of subjects is neither intellectually nor socially divisive; and that it is a vital defence against centralized autocracy'. He sees the underlying reason for integration as a desire for control at the top (that is, *above* the level of Head of Department). He sees this as

---

[1] Hamilton (1971) has argued that the Scottish separate biology, chemistry and physics courses exemplifies Bernstein's 'collection code' while the Scottish Integrated science course exemplifies Bernstein's 'integrated code'.

manifested by the power that primary school Heads wield over their staffs, and this he believes to be closely connected with the integrated nature of the curriculum in the primary school.

Hamilton (1971) studied the implementation of the Integrated Science course in two Scottish secondary schools. He suggests that there is a fundamental dilemma for the schools in that they are required to present a course that is appropriate for the first two years of a four year 'O' grade course, and at the same time is suitable for the common-core course for mixed ability classes:

> 'Today, comprehensive schools are expected to retain this academic/intellectual function while at the same time paying tribute to new patterns of organization, new boundaries of knowledge and new conceptions of education'.

Simultaneous introduction of the two types of political constraint provides a conflict situation which is very difficult to resolve. Insofar as the Integrated Science course is designed as part of the 'O' grade science courses, pupils will be developing attitudes towards the science subjects which see them as associated with specialist groups, defined by the syllabus and terminal examination, and inaccessable, in the long term, to the less able pupils. Insofar as the course is designed for pupils in mixed ability classes following a common-core curriculum, it might be expected to aim for development of attitudes that see science as affecting and involving *everyone*. Attitudes of the former type will clearly impose constraints on attitudes of the latter type, and the development of attitudes such as *an awareness of the contribution of science to the social and economic life of the community* (Curriculum Paper 7, attitude objective, page 16) will be inhibited. However, the separate science subjects are directed specifically towards 'O' grade and do not claim to be designed for the more egalitarian, common-core purpose, and so would be expected to be even less likely than the integrated course to lead to development of such an attitude.

It is therefore suggested that if Curriculum Paper 7 is making a valid justification in saying that it is designed for the common-core course for mixed ability classes, then it is to be expected that it will enable pupils to acquire *an awareness of the contribution of science to the economic and social life of the community* to a greater degree than will the separate subjects course.

Arguments in terms of the conditions which pupils require for effective learning

Many of the major arguments for integration have been based on its assumed value in fulfilling certain pupil needs. The concern here is not for the past or future needs of the pupils, but for those he has while the integrated curriculum is responsible for his learning. The arguments cover such areas as the pupil's motivation, security, opportunity to be creative, level of interest, and achievement commensurate with levels of ability and development.

Fiasca (1970) suggests that learning will be more effective if pupils are exposed to the content of the disciplines combined to make a 'coherent whole' thus 'concentrating the learning experiences'. This he sees as also reducing the chances of the pupil missing any inter-relationships that may exist between the component disciplines. He does not provide details of any evidence or theory on which he bases these judgments.

He also sees integration as useful in enabling the pupil to make contact with a large number of specialized fields of knowledge, fields which are becoming too numerous to be dealt with individually.

Schulman and Tamir (1973) suggests that any opinion about what sort of knowledge is most learnable, most motivating, most readily retained and transferred to new situations by pupils, is 'tracable to the choice of psychological theory that has directed the educator's conception of what is to be taught and learned'.

Curriculum Paper 7 makes statements about the value of Integrated Science to the pupil, but whether this can be linked to any identifiable psychological theory is doubtful (the only possible explicit reference they provide is that to Bloom *et al.*: Taxonomy of Educational Objectives 1956).

Pupil motivation is not considered as a separate issue in the Paper but is conflated in the group 'interest, enjoyment or satisfaction in the things of science, or in the possible commitment to the study of science' (page 14). The document hopes that 'if our recommendations are followed there will be found much satisfaction...by the pupil in mastering the ideas put before him' (page 14).

Whitfield (1971 page 227) suggests that an integrated approach in thematic form 'may provide for greater pupil motivation' in that it can deal with matters of immediate concern to pupils (sport, cosmetics, sex, money).

Kirk (1973) criticizes such views as being without empirical support. He sees them, as based on 'hearsay and anecdote', and pupil motivation as dependent on the effectiveness of the teacher, not on the organization of the knowledge.

With regard to pupil 'interest' Curriculum Paper 7 states that 'Everything has been done in preparing a syllabus, and in considering the materials to support science teaching, has been influenced by the desire to foster in pupils an interest and enjoyment in science'. Presumably this conception of interest covers two aspects: a sort of intrinsic interest in the academic subject, and also in the sense of the curriculum being concerned (in some way that separate subjects are not) with areas relevant to the interests of pupils. Whether or not an integrated curriculum is inherently more interesting (*i.e.* interest in the first sense) seems to be argued at the level of personal experience only. Bernstein (1971), in looking at the relevance to pupils' interests aspect, suggests that while boundaries between commonsense, everyday knowledge and the knowledge imparted in the classroom are very strong in a separate subject code, for the integrated code the boundaries between school knowledge and everyday realities are broken down. Once again Kirk (1973) argues that the relevance of the curriculum does not, in fact, depend on how the knowledge is classified (integrated or not) but on how the teacher or curriculum developer chooses to structure the course or frame the instruction. He does admit, however, that while individual subjects may choose to be concerned with areas of deep concern to the pupil, their inherent narrowness may only allow 'partial insights or partial explanations' of problems.

The Curriculum Paper is concerned (page 18) that pupils who are used to having only one teacher in primary school shall not lose their sense of security on reaching secondary school and being confronted with a large number of teachers. Teachers (interviews cited earlier) were also concerned about this, and felt that pupil security and adequate teacher-pupil relationships could only be built up on the basis of contact three times a week rather than once (as is the case for separate subjects at that stage). There is no evidence to indicate whether or not contact with one teacher rather than three is beneficial to the pupil (he probably meets ten other teachers from other departments). It is interesting that a frequently presented argument that the teachers made for a *separate subjects* course was the benefit to be gained by the pupil from a variety of teachers.

Learning and transfer seem conflated by the Curriculum Paper under 'Thinking'. They see the Integrated Science course as providing more time and opportunity for thinking than the separate subjects course where frequently:

> 'these thinking situations arise fortuitously, and rely on thought processes established in a haphazard way by the pupils during their struggles with the knowledge they have been acquiring. The Working Party believe that it is possible to give some training in thinking by creating situations in which it is necessary (*i.e.* problem situations). [Integrated] Science is well placed to provide such a training since it can call upon so many practical as well as theoretical techniques. By basing the syllabuses firmly on the idea of discovery methods the pupil is required to react continuously in a thinking situation; he learns by hypothesising and discussing, by experimenting, by measuring and by reassessing his hypothesis in the light of experimental results' (page 12).

These descriptions of generalized scientific abilities assume that transfer of training can take place across science subject boundaries. The extent to which such transfer is possible might be determined by empirical investigation (this has not been done in this context). Hirst (1965) is dubious about the extent to which generalizable abilities exist and cautions that 'such abilities must necessarily be characterized in terms of the public feature of knowledge, and...the particular criteria for their application in diverse fields' (page 120). Whitfield (1971 page 25) suggests that Hirst 'picks on highly specific skills and knowledge of conventions in which transfer could only take place through an agency such as telepathy'. He argues that there is transfer of complex abilities (such as hypothesis formation and testing) between subjects, particularly if they are as closely related as the natural sciences; and if this were not the case 'much educational endeavour cannot be but a waste of time'. Kelly (1967) argues that transfer is an unnecessary and vague concept and should not be introduced. His position is that if 'learning' and 'transfer' are distinct concepts, we should be able to separate a learning situation into the part that will influence subsequent learning and the part that will not. But, he argues, all testing of 'learning' involves a 'new' situation *i.e.* testing for transfer. To fail to transfer may be due to a pupil failing to know, or failing to realise that he is expected to make use of his prior knowledge *i.e.* a failure to learn. Curriculum Paper 7 does not differentiate between learning and transfer either but it may be helpful here to follow Schulman and Tamir (1973) in contrasting several psychological theories, on the basis of their view of transfer, in an attempt to identify the psychological basis on which the Curriculum Paper places its integration.

Schulman and Tamir suggest that Gagné is concerned with *vertical* transfer in his learning heirarchies. Gagné (1970) builds a simple to complex structure (actually he starts from the complex end) in which concepts are developed from distinctive facts, principles are formed from associations of these concepts, use of these principles then results in some particular problem-solving capability. His model is concerned with 'intellectual skills or strategies' and not with 'verbalizable knowledges' (content).

While *lateral* transfer is not seen as impossible it is, in general, not achieved between different groups of concepts at the low heirarchy levels but at the high 'problem' levels. The Schools Council Integrated Science Course (1971) has been developed on the basis of a Gagné model. However, Curriculum Paper 7, though concerned with development of problem solving capabilities (page 13), provides no evidence of an underlying process-heirarchy. Indeed, content or subject matter (verbalizable knowledge) are given considerable emphasis (pages 19 - 21, 63 - 102), and any structure is based on the content of individual sections rather than on the course as a whole.

Ausubel (1968) is seen as viewing transfer in terms of subject-matter rather than processes. His theory of Meaningful Verbal Learning requires a framework of *advance organizers* to which the learner can relate new elements. This framework of statements is supplied by instruction from the teacher, or, in Ausubel's terms, has been learned through *reception* rather than *discovery*. Meaningful learning then consists of absorbing the new elements of knowledge into an established complex of generalizations. Provided the learning task is intrinsically meaningful, provided it is potentially meaningful for the 'intellectual capacities, ideational content and experiential background' of the particular learner, and provided the learner has a set to relate substantive information to the existing structure then Ausubel is satisfied that 'the outcome should be meaningful and the advantages of meaningful learning (economy of learning effort, more stable retention and greater transferability) should accrue irrespective of whether the content to be internalized, is presented or discovered, verbal or non-verbal'.

Curriculum Paper 7 gives no indication that its subject matter was selected on any such systematic basis as that implied by advanced organizer frameworks and subsequent learning elements. Indeed it seems to have been more like bits of the physics and chemistry 'O' grade syllabuses, something appropriate from Biology to give balance, and a few other things like 'the earth' and 'introducing science'. It hints that there might be some underlying conceptual structure in the integration

'The approach to science envisaged in this report is, however, not designed to require the learning of many unassociated facts. Instead, it teaches some fundamental concepts and the ways of scientific thought by means of pupil investigation and participation' (page 25).

However, no such structure is described.

The pupil participation mentioned above and the observation that

'It has seemed to us that science is a subject which is admirably suited to discovery methods; indeed the very process of discovering appears to us to be as important as the knowledge discovered' (page 22)

suggest that the Curriculum Paper may be following Bruner's (1960) view of learning as a process of discovery. He sees within this process the lateral transfer from subject to subject of broad principles and processes. This view perhaps reflects a possible argument for integration (proposed by

Pring 1971b) in which the structure of the subjects are seen as 'end points' of knowledge that have been 'worked out' over the ages, but which do not indicate the *processes* by which this was done. An integrated curriculum can allow the pupil to reach these same knowledge structures by means of his own curiosity and enquiry (*i.e.* an integrating *process*, or discovery).

However, Curriculum Paper 7 has felt the need to temper Bruner's approach. It recommends guidance in the form of worksheets setting up a system of 'stage-managed heurism'. It recommends on the one hand, that wherever possible the pupil should be able to exercise selection of his own method and approach, yet on the other hand, that worksheets giving guidance in method and approach be used throughout the first year and for a large part of the second. They point to lack of skills and experience on the part of the pupil necessitating support from the teacher. This denies any commitment to Bruner's position that the pupil *can* (and will only learn meaningfully if he *does*) discover. They view the worksheets as 'programmed procedures of investigation' but without the provision to the pupil of 'the final and correct answer before he begins his investigation' (page 22). There is little correspondence between this situation and the problem/ discovery situation of Bruner which may take the form of (summarized by Schulman and Tamir 1973):

'a) goals to be achieved in the absence of readily discernible means for achieving those goals,

b) contradictions among sources of information of apparently equal credibility,

c) the quest for structure or symmetry in situations where such order is not readily apparent'.

There is little evidence that the Curriculum Paper is supported by any of these psychological theories of learning, but there is ample indication that it has rested heavily on the classification of educational objectives by Bloom *et al.* (1956) and Krathwohl *et al.* (1964) both for objectives (page 16) and testing procedures (page 27). This Taxonomy was developed from an extensive review of the literature on objectives and has produced an outstandingly useful subject-independent check-list for curriculum developers. By its very nature it cannot emphasise the objectives specifically relevant to science and certainly cannot support an argument for or against integration of science. It provides a structured list of behaviours observable to the teacher (*i.e.* testing material) but it is not a theory of learning. Since it is based on neither a theory of knowledge nor psychological theory, nor on the needs of society, it is unlikely to form a useful basis on which to argue for a particular science curriculum.

Closely related to the arguments for integration in terms of pupils' learning are those in terms of pupils' level of development. Curriculum Paper 7 (page 17 and 22) is conscious of the limitations of 12 to 14 year-olds, but it does not provide an explicit argument for an integrated approach in these terms. Ausubel (1963) suggests that for some of this group at least, the 'general laws and methodological cannons' of science will only have meaning to the pupil if he can relate them to more tangible experiences. He sees the more abstract principles of scientific enquiry as having to give way to concrete - empirical explanations (demonstration, practical experiments). Thus the *logical* sequence of proceeding from *basic concepts* (such as those of physics and chemistry) to *complex phenomena* (such as those of biology and geology) cannot be followed because of the level of abstraction of those basic concepts. The pupil is intellectually more ready for the 'complex' everyday experience he has (which has no regard for subject boundaries *i.e.* integrated) than he is for the 'simple' laws of physics.

Curriculum Paper 7 lays explicit emphasis on pupil attitudes of interest, enjoyment and enthusiasm for science, and implies that such attitudes will develop, to a large extent, as a result of the weak boundaries between the knowledge of the science classroom and their everyday experience (*i.e.* the relevance of the course). The separate subjects emphasise illumination of the *subject* rather than satisfaction for the *pupil*. If the Curriculum Paper's claims are valid then it must be hypothesized that the integrated science course will lead to greater *interest and enjoyment* of science than will the separate science subject courses.

Arguments in terms of the conditions that the teacher needs to operate effectively

Within the demands and constraints of the social milieu there are a number of perceptions that the teacher has of his own adequacy in the context of a particular curriculum, of the satisfaction and comfort he would experience in that context, and of the advantages that would accrue to him in terms of making his teaching tasks easier. Such factors would be expected to be influential in any arguments that he might put forward for, or against, that particular curriculum.

Whitfield (1971 page 232) has pointed to the problems of integrated science arising from hostile attitudes of teachers who have had specialized training in separate science subjects. He suggests that integrated study is contingent upon: 'a willingness of teachers at present in separate compartments to merge their activities'.

The notion of integrated science assumes transfer of training across science subject boundaries. If we accept that some such transfer can take place, then it is necessary to see why the teacher is more likely to be able to bring this about in an integrated than in a separate subject curriculum.

On the basis of a review of the relevant literature Stephens (1963) has compiled a list of 'general rules' for teaching that is designed to maximize transfer between school subjects. He suggests that, firstly, the feature to be transferred should be brought out, secondly meaningful generalizations should be developed, thirdly a variety of experiences should be provided, fourthly practice in application to other fields should be given, and lastly practice in transfer should be encouraged.

Suppose, for example, the teacher wishes to encourage transfer of the principles of model-building encountered in the solution of physics problems to the solution of biological problems. Firstly, the model-building in the context of the physics problem will have to be brought into focus to ensure that it will not be missed by the pupil. Secondly, the teacher will be more likely to help the pupil to transfer if he gives him the opportunity to recognize a *general* activity of model-building than if the experience is limited to one (perhaps rather esoteric) example. Thirdly, the general activity will be exemplified in a number of areas of physics. Up to this point the teaching may be equally effective in a pure physics or in an integrated course, but the former limits the number of different contexts in which model-building can be exemplified. However, the next recommended step, that of giving the pupils practice in model-building in another field (biology), is probably not appropriate for a physics course. This may well be considered a task for the biology teacher and, as such, may, or may not, be carried out. The greater scope of content in an integrated course will enable the third and fourth recommendations (above) to fuse, and so impose fewer restrictions on the teacher's choice of examples. Lastly, the teacher will organize for a series of successful experiences of transfer, since, for example, it is expected that the pupil who has had practice in transferring the principles of model-building will be more likely than the pupil who has not had this practice, to transfer the principle of the conservation of energy to various fields. Logically we would expect that an individual teacher would have more opportunity to exemplify transfer when teaching across the broad subject base of an integrated course than in a separate subject course.

Bernstein (1971) sees the integrated code as a means whereby teachers can enter into social relationships with teachers from other departments based on a 'shared, co-operative, educational task'. This may make co-operation with other teachers easier and avoid 'a type of organizational system which encourages gossip, intrigue and a conspiracy theory of the workings of the organization, as *both* the *administration* and the *acts of teaching* are *invisible* to the majority of staff'. Bernstein warns that this more relaxed atmosphere for the teacher will be at the expense of some of his privacy in teaching, and will be threatened if the integrating ideologies are not shared by all the staff involved.

Kirk (1973) has argued that social relationships, in this case between teachers, are independent of the organization (integrated or collection) of the curriculum, and depend on the personalities and behaviour of the teachers concerned *i.e.* 'tender-minded democrats' or 'martinets'.

Kirk is also critical of arguments (Stenhouse 1968, James 1968) which suggest that progressive teaching methods follow 'inevitably from the adoption of an integrated curriculum'. These arguments suggest that because the integrated curriculum is concerned with overarching problems or topics the teacher will have to abandon 'instruction' for a 'discussion' mode. Kirk (after Pring 1971b) is of the view that even though there are no 'right' or 'wrong' answers to the problems under consideration, this will not prevent a teacher, who so wishes, working out some 'acceptable answer' which will then be imparted by instructional procedures. He concludes that 'The mode of curriculum organisation is logically quite independent of the methods used by the teacher to bring about the learning'.

Curriculum Paper 7 does not claim any logical connection between integration and particular teaching methods. However, it does suggest that this particular course in combination with the worksheets will enable the teacher to deal effectively with mixed ability teaching groups.

> 'At the same time the teacher must be freed from the need to deal with the whole class at once, so that he can give his attention to the individuals or small groups, working at the different levels and rates. It has seemed to us that this is best accomplished by the use of worksheets covering the various activities in which pupils are expected to participate and we have prepared a set of these for use with the syllabus' (page 23).

The Curriculum Paper argues that the integrated course will cut down the time the teacher needs for revision, and will facilitate his assessment of pupils:

> 'Seeing a class at most once a week will require much more time to be spent on revision of work already done than would be the case when the class is seen three times a week. It is therefore probably more economical of a teacher's time if he takes a single class through

> the subject matter of a combined science course, rather than try to take three classes for his own specialism'.
>
> 'If science is taught as three separate disciplines at this stage the time available for each teacher to assess his pupils is really quite inadequate' (page 18).

They also are optimistic that a course that is concerned with generalizable methodologies of science will improve teacher/pupil communication:

> 'We would hope that all of the methods open to the scientist will be explored and practiced so that communication between pupil and teacher...will be as effective as possible' (page 11).

It is clear that the writers of the Curriculum Paper were uneasy that the teachers would perceive themselves as inadequately trained for the integrated course. Their mild rebuke states:

> 'As teachers...we still, too often, see ourselves as chemists or physicists or biologists first and scientists second. The Working Party has felt, nevertheless, that the curriculum should be framed for the pupil's benefit rather than the teacher's convenience' (page 18).

However, they implicitly sympathize with the teacher's feelings of anxiety, stress, inexperience, and lack of interest and knowledge, in teaching in an area outside his own speciality, when they describe their intention as being

> 'to help the non-specialist in any subject to have an adequate and safe first passage through what may be new ground, both in subject matter and teaching method' (page 9).

Arguments used by Scottish science teachers (interviews cited earlier) laid emphasis on how teaching the integrated course broadened their experience, and on how much more interesting they found it was to teach integrated science to one group than, say, the same chemistry course to three different groups one after the other.

A situation in which a teacher is able to teach competently, is interested in and enthusiastic about his subject, and in which he feels secure, would seem more likely to lead to high levels of pupil attainment and interest than one in which the teacher feels inadequate, insecure, and uninterested. However, Curriculum Paper 7 makes few claims in terms of teacher's feelings of satisfaction, and a separate sciences course may well provide a less stressful context where the teacher is only concerned with the subject for which he is trained. Statements from the science teachers indicated a division of views about which type of course they considered most beneficial to themselves. There is, therefore, no reason to expect that either the Integrated Science course or the separate science subjects, (insofar as they are designed to supply the conditions the teacher needs to operate effectively) will be more effective in developing pupil interest.

To the extent that an integrated course facilitates the teacher's task of transfer of training between science subjects, we would expect inter-relationships between the science disciplines to be more readily established than in a separate subjects course. However, the distinction is not so clear when we try to interpret 'awareness of inter-relationships'. The nature of the relationships must be identified (and this Curriculum Paper 7 does not do) if we are to differentiate between:

a) inter-relationships between different *forms* of knowledge (Hirst and Peters 1970),

and b) inter-relationships between different *elements* of the structure of a unified knowledge.

The latter, it is suggested, would involve highly complex (and largely unestablished) theorizing. It seems unlikely that the curriculum planners intended introduction of such sophisticated material into these curriculum objectives. A discussion with two members of the Working Party that drew up Curriculum Paper 7 led to the formulation of 'Position Statements' which, they felt, expressed the intention of the attitude objectives of the Paper (see Chapter 5). The statements, corresponding to the objective concerned with relationships between the sciences, showed concern for techniques, procedures, and concepts that are common, or borrowed, among the disciplines *i.e.* inter-relationships in the sense of (a) above. It is possible that awareness of such commonality or borrowing can only be developed among those pupils who are already aware of the existence of the different subjects. The teacher may be expected to encourage the latter awareness more effectively in the separate science courses. We are left with an equivocal situation where the conclusion that the integrated course provides the teacher with more opportunities for emphasising and exemplifying the relationships between the sciences, conflicts with the conclusion that it constrains the teacher from establishing the prior distinctions between the sciences that may have to be made.

Arguments for integration in terms of the constraints imposed by the nature of science

Much of this argument has already been discussed under 'Meanings of

Integration'. Curriculum Paper 7 predicts that the integrated course will make clear 'the unity of science', (page 18) provide 'equal time for all three' sciences (page 17), teach 'fundamental concepts and ways of scientific thought' (page 25) and 'expose the characteristic methodology of the scientist' (page 10). If the implied methodologies and structures can be adequately described, these are powerful arguments.

Jevons (1969) has a somewhat more pragmatic approach. He suggests that the way we have carved up knowledge in the past has been to some extent arbitrary and not on epistemological grounds. He sees the decisions that have determined the extent and intellectual pattern of the traditional subjects as being based on an 'intellectual coherence' *i.e.* a system of looking at many things in one way. However, he observes that 'some old course structures are no longer representing the patterns of coherence that are the most significant and most educationally valuable in the present states of the subject-matter and society' (page 116). A second form of 'intellectual coherence' is looking at one thing in several different ways, and may well be more appropriate. He does, however, warn that patterns of integration are not uniquely determined.

> 'It is not difficult to draw diagrams showing how all the main topics are derived from one or two key concepts - matter and energy say - with a lot of lines showing the total interconnectedness of everything with nearly everything else. Exercises of this kind are certainly interesting and valuable to teachers, but how much they help students is debatable...and in any case teaching cannot be done according to such schemes, since time has one dimension less than a block diagram' (page 150).

Again and again (*e.g.* Fiasca 1970, Kerr 1966) we find arguments for teaching science as a process of enquiry. these arguments leaning heavily on Schwab's syntactical structures. It may be that Curriculum Paper 7's view of the nature of scientific enquiry is implied by some of its objectives. These are concerned with procedural schemes and attitudes which may be perceived as typical of 'good' enquiry.

The following examples are taken from the specific objectives laid down for the various sections of work for first year pupils (page 98):

Section 2-6 'some familiarity with the formation of hypotheses concerning animals and ability to test these hypotheses experimentally'.

Section 4-9 'some familiarity with the process of reasoning inductively, in constructing a kinetic model, and of testing the predictions from the model experimentally'.

Section 5-9 'ability to form hypotheses concerning solubility and to test these experimentally'.

Section 7-6 'ability to generalize from particular observations in simple electrical circuits'.

There seems, therefore, to be concern that the pupil will become familiar with the procedures of collection of factual material and hypothesis formation and testing, but no single procedural *order* is stressed. Thus both the scheme of 'induction' (Section 4-9, Section 7-6) where the scientist collects facts with an entirely open mind and then constructs generalizations (and so laws) to govern these facts, and that of 'hypothetico-deduction' (Section 2-6, Section 5-9) where the scientist on the basis of few or no observations 'hunches' a hypothesis, deduces *expected* facts and checks them with *observed* facts, are illustrated.

However, as well as specific objectives *explanatory notes* are provided for each section. If we look at the notes corresponding to the Section 2-6 we can see that it is not, in fact, the hypothetico-deductive scheme that is being proposed:

Section 2-6 'the investigation should continue for a period of time and thus allow opportunities for making accurate observations at intervals, recording these simply, in order to formulate hypotheses' *i.e.* inductive scheme (page 64).

The heavier weighting given to the inductive scheme is difficult to understand in view of the weight of opinion in favour of the hypothetico-deductive system (Bradbury 1969, Holroyd 1972, Koestler 1964, Medawar 1967, Popper 1963) as a more appropriate description of the ways in which scientists work. Curriculum Paper 7 appears to exemplify what Jevons (1969, page 30) has referred to as the 'standard naive philosophy of the scientific method'.

It may be that the writers of the Curriculum Paper consider that the pattern of investigation in science appropriate for pupils of this age group should not necessarily mirror that of the scientist (Ausubel, 1963, claims that 'the goals of the science student and the goals of the scientist are not identical'). However, such a distinction is not made explicitly, and there is, in fact, an implicit correspondence between the two since the course is seen to teach 'ways of scientific thought by means of pupil investigation and participation' (page 25) and 'to expose pupils to many aspects of the work of the scientist' (page 10). The characteristic attitudes of the 'good' scientist[1] are exemplified in the Curriculum Paper by the objective that

---

[1]Gauld (1973) has argued that such characteristics popularly perceived as objectivity, open-mindedness, lack of bias, scepticism, curiousity with a critical, questioning and rational mind, do not characterize the real practicing scientist. His review of the evidence leads him to conclude that 'If all scientists were controlled by some of these attitudes there is a strong possibility that the progress of science would be seriously inhibited...some concern [should] be expressed over the prescription that development of these attitudes be encouraged in pupils at school'.

pupils should acquire *an objectivity in observation and in assessing observation* (page 16).

The emphasis in Curriculum Paper 7 on scientific processes is mirrored in the separate subjects courses (indeed, much of the material is lifted straight from the separate syllabuses into the Paper). The 'objectivity' objective appears only in Curriculum Paper 7, but as argued previously, no explicit claim has been made that the integrated course is likely to be more effective in achieving this than are the separate sciences. Curriculum Paper 7 has provided no argument to refute the contention that if 'scientific enquiry' is a generalizable methodology, then it can equally well be demonstrated in the separate sciences as in integrated science.

If Curriculum Paper 7 is concerned with integrated conceptual structures of scientific knowledge when it talks of the 'unity of science', then there is a need to be explicit about what these structures are. As it stands, the syllabus appears to regard science as a collection of separate subject strands (sections or parts of sections) to be taught by one teacher. As such, it seems unlikely that there would be any differences (relating to the nature of science) between the integrated and the separate science subjects courses in achieving those objectives concerned with awareness of relationships among the sciences and other school subjects. In an earlier section it was argued that attitudes such as these were logically related to meanings of integration in the 'strong' and 'weak' theses sense. However, though the text of Curriculum Paper 7 implies integration in the 'strong' and 'weak' theses sense, as observed earlier, the organization of the content reflects only 'interdisciplinary study'.

Summary and Hypotheses Concerning Attitude Objectives

The reasons that have been given to justify the introduction of integrated science curricula have been reviewed, and the extent to which these reasons emphasise, or constrain, pupils' attitudes to science have been discussed. Curriculum Paper 7 has been examined and its justifications for integrated science described, its attitude objectives related to such justifications, and its content and organization of subject matter briefly considered.

On the basis of these considerations it is tentatively suggested that:

1. No significant differences are hypothesised in the levels of achievement of the following attitude objectives between pupils following the Integrated Science course of Curriculum Paper 7 and pupils following three separate science subject courses: that pupils should acquire
   a) *awareness of the inter-relationship of the different disciplines of science,*
   b) *awareness of the relationship of science to other aspects of the curriculum,*
   c) *an objectivity in observation and in assessing observations.*

2. It is hypothesised that pupils following the Integrated Science course will achieve the following attitude objectives to a greater degree than will pupils following the three separate subjects course:
   a) *awareness of the contribution of science to the economic and social life of the community,*
   b) *Interest and enjoyment in science.*

These hypotheses have been cautiously proposed. They could only be presented with confidence if we had reason to assume that those involved in teaching the Integrated Science course not only follow the syllabus, but also accept and implement the various facets of the course (often only *implicitly* dealt with in Curriculum Paper 7) that have been considered in this chapter.

CHAPTER 5

THE CONSTRUCTION OF SCALES TO ASSESS THE ACHIEVEMENT OF ATTITUDE OBJECTIVES

The last four chapters have been concerned with an examination of the nature and formulation of the attitude objectives of Curriculum Paper 7, of the justifications for their inclusion in the Paper, and with the relationships between them and the 'integration' of the course.

Chapter 6 is directed towards an exploration of the relationships between these attitude objectives and dimensions of attitudes to science exhibited by pupils. Chapter 7 and the later chapters relate to an empirical study of the assessment of achievement of the objectives among S1 and S2 pupils in Scottish secondary schools (Chapters 7, 8 and 11 describe the plan of the study). For each of these purposes some means of assessment of criterion attitude variables is required.

This chapter concerns itself with the construction of such an instrument for assessment of attitudes, corresponding to the 5 attitude objectives of Curriculum Paper 7, by pen and paper methods. The pilot study of 323 pupils which provided data for the item analysis and reliability measures for this instrument will be described, and the results of the assessment of the attitudes of these pupils will be used to develop hypotheses for the main empirical study with which the later chapters are concerned.

Construction of the Attitude Scales

The purpose of this empirical study was to evaluate the achievement of each of five attitude objectives. There was no reason to believe that these objectives *necessarily* corresponded to distinctive dimensions of attitudes possessed by pupils (the extent to which such correspondence exists will be explored in Chapter 6). Consequently it was felt that factor analytic methods (such as those used by Omerod 1971 or Skurnik and Jeffs 1971) for attitude scale development would be inappropriate. It was decided that five separate subscales should be drawn up and combined to form a single questionnaire. The objectives range from those which emphasise intellectual (highly cognitive) attitudes to those which emphasise emotional attitudes but it was expected that there would be substantial correlations between some subscale scores.

Position statements

Meetings were held with two members of the Working Party that wrote Curriculum Paper 7 in order that they might expand and define the objectives. The outcome of these meetings was a list of *position statements* that, the two Working Party members agreed, reflected their intentions for the objectives. Two statements were constructed for each objective: an 'A' statement identifying an appropriate 'favourable' attitude, and a 'B' statement identifying an appropriate 'unfavourable' attitude.

The first attitude objective *(awareness of the inter-relationship of the different disciplines of science)*, it was revealed, was included in the Paper in order to allay the fears of those who felt that there was some danger of pupils reaching the end of S2 with little idea of what constituted the separate science disciplines that make up the S3 course. For this reason certain sections of the syllabus had been deliberately chosen as separate physics, chemistry and biology components in the hope that pupils would become aware of the separate subjects but at the same time would be able to look back to the earlier sections for the inter-relationships. The position statements developed for this objective were:

> 'A: It is sometimes convenient for us to divide science into three sections: biology, chemistry, and physics. Although each of these three sections stands alone at some points, there are many overlapping concepts and techniques which link two or all three.'
>
> 'B: Physics, chemistry, and biology are separate, independent subjects.'

For the second attitude objective, *awareness of the relationship of science to other aspects of the curriculum*, the following position statements were put forward:

> 'A: I can see that other subjects such as geography, art, music, homecraft, physical education, and agriculture depend in many ways on science for their development. Science in turn depends on mathematics.'
>
> 'B: Science is of little use to any other subjects. Artists, musicians, geographers, P.E. teachers, farmers, and homecraft teachers have no need of it. Science can carry on quite successfully without mathematics.'

The Working Party had been anxious that this course would not attempt to coerce pupils into regarding science as making only 'good' contributions to society. They felt strongly that the pupil must be given the opportunity to recognize and examine the 'bad' aspects as well. The third attitude objective, *awareness of the contribution of science to the social and economic life of the community*, was characterized by:

'A: Our lives are affected both favourably and unfavourably by progress in science. Everyone should be made aware of the nature of science and what it attempts to do.'

'B: Understanding of science by the general public would contribute nothing to the community nor to the advancement of science. Science and the community can operate independently of each other.'

The objective *Interest and enjoyment in science* led to:

'A: Science is an interesting, rewarding and enjoyable subject. I would enjoy doing scientific work.'

'B: Science is dull and uninteresting. It is only suitable for "brainy folk". I would not enjoy doing scientific work.'

(The members of the Working Party used the word 'swots' rather than 'brainy folk'.)

For the final attitude objective *(an objectivity in observation and in assessing observations)* the following position statements were developed:

'A: A good scientist depends on observations of phenomena that are empirically verifiable, reproducable and free from the experimenter's idiosyncracies. He is intellectually honest and is ready to alter his position on the basis of further evidence.'

'B: A good scientist offers true explanations of observable phenomena. He accepts the results of the other good scientists and always takes their side.'

Choice of attitude measure

It was decided (see Chapters 7 and 8) that the empirical study of attitudes related to the Curriculum Paper 7 objectives should involve a very large number of pupils (3000+). In view of this, direct observation of behaviour of the pupils would have been impractical and pencil and paper measures of attitudes appeared the obvious choice (in any event behaviour has many determinants other than one specific attitude and so is unlikely to be a reliable indication of a particular attitude - see Chapter 2).

There were a number of qualities that it was necessary, or at least desirable, for the attitude measure to have:

1) It must be suitable for administration to large numbers of pupils.

2) The response procedures, vocabulary and meanings of statements must be understood by young pupils (12 years old) of a wide range of abilities.

3) The response procedures must allow for 'don't know/undecided' attitudes since the instrument would be administered to pupils on entry to school before they have experienced any secondary school science.

4) The instrument should be constructed by well established and documented techniques, and the construction should not be too time consuming.

5) The instrument should be reliable (internal consistency and stability).

6) The instrument should be valid *i.e.* measures what it is supposed to measure.

7) It should be composed of unidimensional components relating to each of the five attitude objectives insofar as the interdependence among those objectives permits this.

8) The instrument should aim to provide an interval scale on which attitudes may be measured, or, if possible a ratio scale with a zero point related to a neutral attitude.

9) The scale should have high reproducibility *i.e.* two identical scores should not be obtainable in a number of different says (reproducibility of scores on a scale is closely related to the unidimensionality of the items of that scale, Shaw and Wright 1967, page 25).

No single instrument is completely satisfactory on all these criteria. The various methods have concentrated on different scale requirements. In comparison with, say, measures of cognitive achievement they are crude.

> 'Their measurements are relative, and the proper use of an attitude scale is to allow comparisons of the attitudes of different groups or of individuals compared with that of groups, or to assess the effect on attitudes of changes in, say, a school curriculum' (Nisbet and Entwistle, 1970, page 126).

and

> 'Such scales cannot, by themselves, be expected to provide us with subtle, insights in an individual case' (Oppenheim, 1966, page 121).

The decision was made to use a summated-rating or Likert-type scale (Likert 1932). This is probably the most widely used type of scale in behavioural research and particularly in educational research.

Likert scales are less laborious to construct than other commonly used scales such as the equal-appearing intervals of Thurstone (1929) scales, or the scalogram-analysis of Guttman (1950) scales.

Thurstone scales

The Thurstone scale construction directs itself to the establishment of equally-appearing interval items with assigned scale values that directly indicate the strength of attitude of a response that agrees with any particular item (see criterion 8 above). A large number of statements that express various feelings towards the issue are collected and presented to 'judges' for sorting into categories (usually about 11), ranging from 'most favourable' attitude towards the issue to 'least favourable'. The recommended number of judges varies from about 40 to 300 (though Edwards, 1957, page 95 does quote one study of scale values obtained independently by two groups of judges where the smaller group was only 15 in number and for which the inter-group correlation was 0.99). Who the judges should be is not always clear but Oppenheim (1966) states that

> 'By and large, it remains advisable to use as judges people similar to the respondents in our research sample' (page 133).

It is assumed that the judges' attitudes are irrelevant since their task is to interpret intelligently the implications of the attitude statement.

If there is marked disagreement among the judges about one particular statement, then it is normally discarded as being either irrelevant or ambiguous. Those statements retained are given a scale value according to their median position as assigned by the judges. A selection of 20 or 30 items is made, ensuring that the different intensities of attitude are covered without large gaps between scale values, and these are arranged in a random order. The subjects of the study are presented with the items and asked to indicate the statements with which they are in agreement, their score is then the median scale value of those chosen statements.

The judgment procedures are elaborate but crucial. The idea of carrying them out with 40+ early secondary school pupils as judges is daunting. The reliability of such scales is usually found to be adequate and the judgment phase frequently provides enough material for a parallel form of the test. Pupils being tested would be expected to agree with only one or two of the statements since the selection procedures tend to produce groups of items that are non-monotone *i.e.* a pupil whose attitude is either below or above the scale value of a particular item will disagree with that item. It is not clear that pupils on entry to secondary school with, as yet, inadequately formed attitudes to science, would be able to handle this type of response. Oppenheim (1966 page 132) comments that the ideal response where only one item is endorsed very seldom occurs, and several individuals can be placed at the same place on the scale though their endorsed items may not be identical *i.e.* the scale may be criticized in terms of its reproducibility (see criterion 9 above). It is probably necessary to carry out a pilot study using those items with small semi-interquartile ranges from the judges' ratings. Respondents would then tick those statements with which they agreed, and only if an individual endorsed items which were all of roughly similar scale values could the scale be said to approach adequate levels of reproducibility and unidimensionality.

Finally, Oppenheim (page 131) comments on the scale intervals:

> 'Although, as we have seen, Thurstone's procedure is primarily concerned with locating items at points on a hypothetical scale, we must be careful when treating these scale values as actual numbers, as units that are additive and interchangeable on a linear continuum. We are dealing not with equal but with equal - *appearing* intervals, with psychological rather than numerical units, which may or may not be equal'.

Guttman scales

The cumulative or Guttman scale concentrates its attention on unidimensionality and reproducibility. It consists of a small group of homogeneous items that measure *one variable only.* The items are selected and ranked so that an individual marking the statement that most nearly expresses his attitude, will also respond positively to all those items that have a lower rank than the one endorsed *i.e.* monotone items forming an ordinal and cumulative scale, each item expressing a slightly different attitude to the one next to it.

The procedure of scalogram-analysis is to select items and test to see if the set of attitude statements falls along one dimension from 'favourable' to 'unfavourable'. The basis for choice of a set of appropriate items is not clear cut.

> 'Just how these statements are selected remains something of a mystery....Guttman (1945) has expressed his belief that the selection of a small number of statements from the large number of possible statements representing the universe of content should be done upon the basis of intuition and experience.... We shall merely say that contrary to Guttman's early advice, it may be desirable to subject the statements relating to a given area of content to item analysis procedures prior to testing for scalability' (Edwards 1957, pages 177-178).

The set of items must be tested on a number (probably several hundred) of respondents to see whether it has a satisfactory coefficient of reproducibility R.[1] Unfortunately, as Oppenheim (1966 page 145) points out,

---

1

$$R = 1 - \frac{\text{no. of deviations from ideal scale pattern}}{\text{no. of respondents x no. of items}}$$

> 'not all areas of content will scale, especially not if they are rather wide and heterogeneous, and one cannot know beforehand whether attempts at scale construction will be successful. The procedure has been criticized for this and for a tendency to produce scales covering a very narrow universe of content'.

It has already been argued (Chapter 2) that some of the attitude objectives of Curriculum Paper 7 cover *broad* areas of affective behaviour and items related to these may well be impossible to scale in this way. In addition the computations involved in the construction of such scales are tedious and rather complex (Edwards 1957, Chapter 7).

Likert scales

The Likert or summated-ratings scales are comparable with other scales in some aspects and superior in others.

The Likert-type scale makes fewer statistical assumptions than the equal-appearing intervals of Thurstone (Nisbet and Entwistle 1970) yet measures on the two scales correlate highly. Edwards and Kenny (1946) found that the correlation between the two types of tests constructed using the *same* group of subjects but with *independent* selection of statements, reached levels of 0.79 and 0.92 (the Thurstone scale was developed as 'Form A' and 'Form B'). The same study suggested that the reliability of the 25 item Likert scale ($r = 0.94$) was above that for the 20 item Thurstone scale ($r = 0.88$). Edwards (1957 page 162) surveys a number of studies of reliabilities of the two types of scale and concludes

> 'According to the evidence at hand, there is no reason to doubt that scales constructed by the method of summated-ratings will yield reliability coefficients as high as or

higher than those obtained with scales constructed by the method of equal-appearing intervals'.

The procedure for construction of a summated rating scale (described below) is far less tedious and complex than that for the Guttman scale, and most writers suggest that it is simpler and less laborious than that for the Thurstone scale. However, on the latter point Edwards (1957 page 169) cites some 'devices' and studies with small (*i.e.* less than the usual minimum of 40) groups of judges where developments of the two types of scale are comparable in time and labour.

Likert scales require, for *every* item, responses along a continuum from strongly agree to strongly disagree, rather than asking the subject to check items that describe his attitudes. A preliminary session with S1 and S2 pupils from comprehensive schools suggested that they were more willing to follow this systematic pattern of response to each item indicating their degree of agreement or disagreement, than they were to choose particular items on a simple 'agree or not' basis as required by the other scales. The number of response categories for each item is variable. In this study five positions of 'strongly agree', 'agree', 'don't know', 'disagree' and 'strongly disagree' were provided in order that

a) pupils could to some extent, indicate the *intensity* of their attitudes,
b) areas where pupils' attitudes were, as yet, undeveloped (particularly for S1 pupils), could be accommodated by the 'don't know' category, and
c) the categories would not be so numerous that pupils would be unable to appreciate the distinctions between them.

The construction of a Likert scale starts with a pool of items designed to express a particular positive attitude and the corresponding negative attitude. Items which are neutral or very extreme are, in general, excluded (unlike the Thurstone scales). In this study, as described later, the original pool of items was reduced by a group of 'expert' judges who decided for each item whether it indicated a particular positive attitude, the corresponding negative attitude, neither of these, or both of these.

A number of respondents are then asked to indicate their levels of agreement or disagreement with each item. It is necessary that the response categories be weighted so that the respondents with the most favourable attitudes attain the highest scores. Edwards (1957, pages 149-151) describes a 'normal deviate system of weights' but concludes that since Likert (1932) found that the relatively simple assignment of integral weights (5, 4, 3, 2 and 1 to the five categories) correlated 0.99 with the more complex system, the simpler system should be used. In this study *positive* items were scored 5 for 'strongly agree' down to 1 for 'strongly disagree', and for *negative* items -5 for 'strongly agree' up to -1 for 'strongly disagree'.

Having obtained a set of scores the most suitable items for the scale must be selected. Oppenheim (1966, page 138) describes this procedure as

> 'something like an act of faith...Ideally, the item-analysis should take place by correlating each item with some reliable outside criterion of the attitude that it is supposed to measure and retaining only the items with the highest correlations. Such criteria are, however, almost never available...We must therefore say to ourselves that, for the moment at least, the best available measure of the attitude concerned is the total item pool that we have so carefully constructed. By purifying this, the items will at least be consistent and homogeneous - they will all be measuring the same thing - and the scale may possibly also be valid. It is rather like trying to pull ourselves up by our own bootstraps!'

The particular form of this 'internal-consistency' method of item analysis that was used in this study followed the procedure of Edwards (1957, page 152). The total scores of all respondents were used to identify a 'high scoring' criterion group and a 'low scoring' group. 't' tests were then carried out between the mean scores for the 'high' and 'low' groups on each item.

The 't' value indicates the extent to which the item differentiates between those respondents we have assumed to be displaying a very positive attitude and those displaying a very negative attitude. The 't' value should be at least 1.75 for us to say that there are significant differences between the 'high' and 'low' scoring groups, where the number in each group is 25 or more. Those items with the largest 't' values are then selected for the scale. In this study the minimum 't' value for items used in the scale was 4.51.

A logical outcome of this method of item analysis is that the scale will approach unidimensionality. However reproducibility is not a necessary characteristic of the Likert scale. Two individuals may achieve the same total attitude score with very different responses to individual items.

The allowance that is made for recording intensity of attitudes with the five response categories appears to be preferred by respondents, provides more information about the level of agreement or disagreement and leads to greater response variance. Unfortunately as Kerlinger (1969, page 484) observes

> 'The variance of summated rating scales...often seems to consist of response-set variance. Individuals have differential

tendencies to use certain types of responses: extreme responses, neutral responses, agree responses, disagree responses. This response variance confounds the attitude...variance.'

In general, scores on Likert scales cannot be interpreted in isolation. A pupil cannot be identified with a scale value on a *psychological continuum* as in the case of the Thurstone scale. We cannot assign a particular level of favourable or unfavourable attitude to an individual. The summated-ratings scale has no point that can be identified as a 'neutral' attitude point. If we are concerned only with comparison of mean attitude scores of different groups, or with correlations between attitude scores and other variables, or with comparing changes in attitude scores as a result of introducing different curricula, then there seems no necessity to identify the 'neutral' point and Likert scales will serve very well.

Other pen and paper attitude measures

There are a number of other attitude instruments. Some of these are impractical for very large numbers (*e.g.* essays, projective techniques); others are limited in the sorts of attitude that they assess (*e.g.* interest inventories); a few present considerable problems of interpretation and analysis (*e.g.* preference ranking, discussed in detail in Chapter 10).

One method that has been used with considerable success in the Harvard Project Physics work (Rothman 1967, Anderson and Walberg 1968, Welch 1969) is that of the semantic differential (Osgood, Suci and Tannenbaum 1957). The respondent is presented with a series of concepts (*e.g.* 'science lessons', 'technology', 'ecology') and a series of scales defined by adjective pairs ('useful/useless', 'enjoyable/unenjoyable', 'complex/simple'). These adjective pairs mark the extremes of 5 or 7-point rating scales, and the subject responds by indicating where he would place each concept on each scale. This rating is taken as a measure of the subject's attitude towards that concept and scores are usually explored using factor analysis which searches for basic dimensions of meaning. Use of such an instrument was considered for this study, but two objections led to the decision against this:

a) a preliminary exploration of pupils' reactions to this type of instrument revealed some lack of comprehension of what was required and considerable hostility towards it;
b) it was not clear that the instrument could be constructed to provide scales directly related to the attitude objectives.

Collection of items

Items for the scales were collected from audio tape recordings made during conversations with small groups of first and second year pupils from comprehensive schools. The items consisted of (i) statements made by these pupils; (ii) statements constructed using the vocabulary of the pupils. One hundred and fifty items together with the 'position statements' were submitted to a panel of 8 judges (1 educational psychologist, 4 HMI's, 2 science teachers, 1 university science education tutor) and each judge was asked to decide for each item whether it indicated a particular favourable (or positive) attitude, an unfavourable (or negative) attitude, both favourable and unfavourable attitudes, or neither favourable nor unfavourable. A final selection of 100 items, on which the judges were in agreement, was made. (Instruction for judgment of items will be found in Appendix A.)

These 100 items were made up of 10 positive and 10 negative items for each of the five subscales. The items were randomized within a single questionnaire.

This procedure follows that of Moore and Sutman (1970) and the assumption is made that:

> 'if a respondent agrees with the attitude statement [item], that agreement could be taken as evidence that he assumes the position statement for which the attitude statement was written'.

Pilot sample

Three groups of pupils were used in order to cover the complete age range over which this scale will be used. Seventy-five pupils from two schools formed the final year primary group, 125 pupils from two comprehensive schools formed the first year secondary group, 123 pupils from the same comprehensive schools formed the second year secondary group. One of the comprehensive schools followed the Integrated Science course for the first two years, the other followed separate biology, physics and chemistry courses. In each case mixed ability groups were used but children unable to read were excluded.

Procedure

All pupils were given the attitude scale once and those from one comprehensive school (64 first year, 63 second year) and one primary school (49 pupils) were given it again 4 days later. During this time they had no science lessons. In addition to this the AH4 group test of general intelligence was administered to the pupils of one of the comprehensive schools, together with three tests of 'divergency'. (The purposes of, and results from, the tests of divergency will be discussed in Chapter 8.)

Statistical Analysis

An item analysis of the 100 item scale was carried out using the following procedure; (i) identification of the top 25% and the lowest 25% of pupils' scores for each subscale total; (ii) computation of mean score for each of these groups on each item relating to this subscale; (iii) application of a 't' test to the difference between the high mean and low mean; (iv) selection of items with the highest 't' value (see Edwards 1957, page 152). The 't' values for the items accepted were all significant at a minimal 0.01 level. The final scale consisted of 60 items making up 5 subscales, each subscale containing 6 'positive' and 6 'negative' items relating to one of the attitude objectives. All attitude scores from this point on refer to this 60 item scale. (The 60 item questionnaire will be found in Appendix A.)

Pearson product-moment correlation coefficients for the test-retest reliabilities for each subscale and the full scale were determined for primary pupils, first year secondary, and second year secondary pupils. These coefficients were found to be satisfactory for all three groups of pupils provided the scale is to be used for comparing groups of pupils rather than assessing attitudes of individuals ($r = 0.93, 0.86, 0.92$, for total scores, lowest value of $r = 0.55$ for subscale 1 for primary pupils, see Table 5.1). 't' tests for correlated means were carried out on the mean scores for each of the groups of pupils on the test and retest. No

TABLE 5.1 TEST/RETEST RELIABILITY COEFFICIENTS : ATTITUDE SUBTOTALS AND TOTAL

		Subtotal 1	Subtotal 2	Subtotal 3	Subtotal 4	Subtotal 5	Total
Primary	n = 50	0.55	0.86	0.87	0.91	0.78	0.93
S1	n = 64	0.72	0.83	0.74	0.88	0.75	0.86
S2	n = 62	0.69	0.81	0.81	0.91	0.78	0.92

All r's significant .001 level

TOTAL ATTITUDE SCORE : DIFFERENCES BETWEEN TEST AND RETEST MEAN SCORES

	Primary	S1	S2
mean score test	20.98	30.01	29.53
s.d.	23.53	18.57	23.33
mean score re-test	24.38	32.35	33.14
s.d.	29.33	20.41	23.46
t value	0.64	0.67	0.85
	N.S.	N.S.	N.S.

significant differences were found (see Table 5.1). The internal consistency or homogeneity of items for each of the subscales was determined from the split half method giving values of 0.72, 0.83, 0.74, 0.93 and 0.62 for Spearman-Brown reliability coefficients.

A correlation matrix (Table 5.2) showing inter-relationships between subscales and the full scale was computed. Correlation coefficients for AH4 scores and attitude subscale scores for the smaller group of secondary school pupils are also shown in Table 5.2. Two-way analyses of variance on the scores of the secondary school pupils were carried out for each of the subscales to investigate the main and interaction effects of sex and type of science course, and sex and age. Those giving significant results are reported in Table 5.3. 't' tests were then carried out to investigate differences between boys' and girls' mean scores for the three age groups and for the 'integrated science' and 'separate science' secondary groups (Table 5.4). These differences between the mean scores of the various groups of pupils on each of the subscales were examined with a view to identification of hypotheses that seem to be worth exploring in the larger main study.

Discussion of results

An interesting feature of the test and retest scores that emerged was the consistent trend for all groups of pupils to show an increase in mean score (63% of the individual scores increased). Other evaluation projects (Wynn and Bledsoe 1967, Science 5/13 1971, and Nuffield Secondary Science - Evaluation, Alexander 1972) have found that scores on attitude measures have dropped for post-tests. This has been explained in terms of negative

TABLE 5.2 CORRELATION COEFFICIENTS BETWEEN ATTITUDES SUBSCALE SCORES AND A.H.4 SCORES
(Figures in brackets give level of significance)

		Subtotal 1 "inter-relationship between sciences"		Subtotal 2 "relationship of of science to other subjects"		Subtotal 3 "social and economic implications"		Subtotal 4 "interest"		Subtotal 5 "objectivity"	
All pupils n = 23	Subtotal 2	0.51	(0.001)								
	Subtotal 3	0.50	(0.001)	0.54	(0.001)						
	Subtotal 4	0.41	(0.001)	0.44	(0.001)	0.62	(0.001)				
	Subtotal 5	0.32	(0.001)	0.27	(0.001)	0.37	(0.001)	0.07	(N.S.)		
	Total:	0.71	(0.001)	0.75	(0.001)	0.85	(0.001)	0.80	(0.001)	0.46	(0.001)
First year pupils n = 64	A.H.4	0.44	(0.01)	0.24	(N.S.)	0.05	(N.S.)	0.05	(N.S.)	0.17	(N.S.)
Second year pupils n = 63	A.H.4	0.13	(N.S.)	0.22	(N.S.)	0.43	(0.01)	0.15	(N.S.)	0.30	(0.05)

TABLE 5.3 2-WAY ANALYSIS OF VARIANCE

	Source	Variance	df	Mean Sq.	F Ratio	Significance Level
Subtotal 2: "Relationship with other school disciplines"	Between Courses	460.7	1	460.7	12.15	0.001
	Between Years	105.8	1	105.8	2.79	N.S.
	Interaction Year/Course	10.9	1	10.9	0.29	N.S.
	Residual (within)	9250.8	244	37.91		
	Total:	9828.2	247			
Subtotal 3: "Social and economic implications"	Between Sex	290	1	290	8.19	0.01
	Between Course	50	1	50	1.41	N.S.
	Interaction Sex/Course	210	1	210	5.93	0.025
	Residual (within)	8640	244	35.4		
	Total:	9190	247			
Subtotal 4: "Interest"	Between Sex	1158.7	1	1158.7	11.58	0.001
	Between Year	271.3	1	271.3	2.71	N.S.
	Interaction Sex/Year	350	1	350	3.50	N.S.
	Residual (within)	24414	244	100.01		
	Total:	26194	247			
Subtotal 5: "Objectivity of observation and assessing observations"	Between Sex	50	1	50	2.19	N.S.
	Between Course	220	1	220	9.65	0.01
	Interaction Sex/Course	20	1	20	0.9	N.S.
	Residual (within)	5557	244	22.8		
	Total:	5860	247			

TABLE 5.4 'T' VALUES FOR DIFFERENCES BETWEEN MEAN SCORE OF PUPILS GROUPED BY SEX, YEAR AND TYPE OF SCIENCE COURSE
(Figures in brackets show significance level)

Group	Sub-groups for 't' test	Subtotal 1	Subtotal 2	Subtotal 3	Subtotal 4	Subtotal 5	Total
Primary	Boys & girls (b) (g)	2.45 (0.02) b > g	3.75 (0.001) b > g	1.63 (NS)	3.71 (0.001) b > g	0.64 (NS)	3.52 (0.01) b > g
1st year secondary	Boys and Girls	0.06 (NS)	1.49 (NS)	0.83 (NS)	1.32 (NS)	2.43 (0.02) b > g	1.78 (NS)
2nd year secondary	Boys & Girls	0.03 (NS)	1.36 (NS)	2.79 (0.01) b > g	3.25 (0.002) b > g	0.36 (NS)	2.46 (0.02) b > g
'Separate subject'	Boys & Girls	0.61 (NS)	1.52 (NS)	0.10 (NS)	1.20 (NS)	2.29 (0.05) b > g	1.43 (NS)
'Integrated Science'	Boys & Girls	0.58 (NS)	1.52 (NS)	3.64 (0.001) b > g	3.49 (0.001) b > g	0.12 (NS)	2.54 (0.02) b > g
1st year secondary n = 125	'Separate' (S) 'Integrated' (I)	1.52 (NS)	2.68 (0.01) I > S	0.02 (NS)	2.52 (0.02) S > I	2.21 (0.05) I > S	0.11 (NS)
2nd year secondary n = 123	'Separate' and 'Integrated'	1.57 (NS)	1.43 (NS)	1.49 (NS)	1.26 (NS)	1.97 (0.05) I > S	1.37 (NS)
Boys 1st & 2nd year n = 127	'Separate' and 'Integrated'	0.90 (NS)	2.09 (0.05) I > S	2.66 (0.01) I > S	0.82 (NS)	0.91 (NS)	1.58 (NS)
Girls 1st & 2nd year n = 119	'Separate' and 'Integrated'	2.35 (0.05) S > I	2.08 (0.05) I > S	1.04 (NS)	1.57 (NS)	3.54 (0.001) I > S	0.16 (NS)

reactions to the test itself (Science 5/13 1971), *e.g.* boredom at 'having to do the thing again', rather than a real unfavourable change in attitude towards science. It seems likely that this type of effect would be most apparent if the time between test and retest was short. Yet here 'no science' results in an apparent increase in score over a very short period of time, while other 'science courses' result in apparent decreases in score over longer periods of time.

The inter-subscale correlations (see Table 5.2) were all positive and significant at the 0.001 level with the exception of that between scores on subscale 4 (interest) and subscale 5 (objectivity). For this pair the correlation was not significant. It is suggested that these two may represent the affective and cognitive extremes of the attitudes being investigated. Insofar as the items of the scales reflect the position statements, subscale 5 demands a substantially higher level of cognitive activity than does the predominantly 'emotional' subscale 4. Several studies (*e.g.* Neidt and Hedlund 1969 and Rothman 1969) have pointed to cases in which gains in the cognitive and affective domains are not significantly correlated.

In studies of the relationship of *cogntive* achievement in science to intelligence for older (and often more able) pupils than our sample, positive relationships have been reported (Brandwein 1951, Anderson, Page and Smith 1958, Butcher, Ainsworth and Nesbitt 1963, Hudson 1966 and Meister 1966). On the other hand studies of younger pupils with a wider range of intelligence (Daniels 1956, Meyer 1961, Cline, Richard and Abe 1964, Tyler 1964, Winder 1966, Wynn and Bledsoe 1967) have indicated zero or negative relationships between intelligence and science bias this being particularly marked if the *affective* measure of 'interest in science' is used as the criterion for science bias. This is supported by the non-significant relationship between AH4 scores and the 'interest' scores of subscale 4. The suggestion that cognitive science achievement and general intelligence are related for *older* pupils is supported by the zero relationship between AH4 and subscales 3 and 5 for first year pupils, and the positive significant relationship for second year pupils.

The correlations between AH4 and the other attitude scores were low (see Table 5.2). There was a relationship ($p<0.01$) between 'awareness of the inter-relationship of the different disciplines of science' and AH4 for first year but not for second year pupils. It is suggested that in the short time that the first year pupils had been exposed to science it would be those of higher intelligence (as measured by AH4) who would have been able to sort out what the 'different disciplines of science' were. By the second year this may well have been achieved by most pupils so correlation with intelligence could be expected to disappear.

From Table 5.4 it is seen that for 'awareness of the inter-relationship of the different disciplines of science' (subscale 1) there was no significant difference between boys' mean scores and girls' mean scores except among the primary school pupils, the boys' mean score being higher. There was no significant difference between mean scores of the boys following the 'integrated science' course and those of boys following 'separate science' subjects. In the case of girls, however, the 'separate science' course group scored higher.

It has been suggested (Roe 1963) that a science sex-difference is culturally determined long before the pupils arrive at secondary school. The primary school and home may provide activities and reading material for boys that will reinforce a science orientation, and at the very least provide familiarity with words such as 'chemistry', 'physics' and 'biology'. The girls are less likely to have reinforcement in this area and so may arrive at secondary school quite unaware that there are 'different disciplines of science' let alone any inter-relationship. Once at secondary school it is possible that those directed to biology, chemistry and physics classes in a 'separate subjects' course will attain this awareness sooner than those in an 'integrated science' course.

Subscale 2 ('awareness of the relationship of science to other aspects of the curriculum') again showed differences between the sexes at the primary school level only, the boys having the higher mean score. Both boys and girls following the Integrated Science course had mean scores significantly ($p<0.001$) higher than those of pupils following separate subject courses, the effect being most pronounced in first year secondary pupils. Table 5.3 shows the significant mean effect of course type ($p<0.001$).

Tables 5.3 and 5.4 show that at an overall 1% level boys' mean scores were significantly higher than girls' on subscale 3 ('awareness of the social and economic implications of science'). The highest mean scores here were from boys following the 'integrated science' course. Their mean scores were significantly higher than boys following the 'separate subjects' course but there were no significant differences between the two groups of girls.

'Interest and enjoyment' (subscale 4) appeared as a male characteristic since primary boys, S2 boys and 'integrated science' boys had significantly higher mean scores than the corresponding groups of girls. The 'separate science' pupils were significantly higher on interest mean scores than the 'integrated science' pupils in S1 but there was no significant difference in S2. In general girls' mean scores fell as from first to second year. This sex difference 'which seems to be characteristic of our culture' (Wynn and Bledsoe 1967) is in agreement with other studies of similar groups in the United Kingdom and the United States (this will be discussed further in Chapter 7).

The significant *course* effect ($p<0.01$) for subscale 5 (see Table 5.3) suggests that Integrated Science may be having a favourable effect on pupils 'objectivity in observation and in assessing observations' (subscale 5) particularly for girls. There was a significant difference between sexes for the 'separate subjects' course (boys scoring higher) but not for the 'integrated science' course. In S1 and S2 the 'integrated science' course pupils' mean scores were significantly higher than the 'separate subject' pupils, this being due to the high scoring of the girls in the Integrated Science group. There was no significant difference between mean scores of boys in each of the two groups.

Field and Cropley (1969) in a study of 178 pupils in the 16 to 18 year age range found that for each sex 'a different pattern of cognitive functioning was...associated with high science achievement'. The major difference between the sexes appeared as a significant positive relationship between originality and science achievement for girls and a corresponding zero relationship for boys. They suggest that in the past school science with over-emphasis on content may have inhibited the development of scientific skills and interest among girls.

It has already been suggested that a relatively high level of cognitive science achievement may be necessary for the attainment of 'objectivity in observation and assessment of observations'. If this achievement is related to cognitive style then we might expect, according to the Field and Cropley hypothesis, that some teaching strategies would be more suitable for girls than others and they may achieve more in an atmosphere which is not primarily concerned with factual information and where alternative approaches to problems are not inhibited.

Curriculum Paper 7, page 22, stresses that 'In the integrated syllabus which has been prepared for the first two-year cycle, there is much reduced emphasis on the factual content of the syllabus' and that the subject should be 'taught in such a way that creativity...is encouraged and fostered at every possible opportunity'. Our results suggest that in the main study it might be worth exploring the hypothesis that girls achieve attitude objectives that are closely related to cognitive achievement to a greater degree in an Integrated Science course than in a traditional course.

Validity of the attitude subscales

Considerable effort has been made to produce a reliable and valid group of attitude measures. The reliability of each scale has been assessed, but validity is much more difficult to demonstrate and is not a necessary concomitant of reliability.

The process of *content validation* is basically judgmental. The question to be considered is: to what extent is this group of items relevant to, and representative of, the attitude objective under consideration? For this study the definition of the appropriate universe of content was attempted by formulation of 'position statements' by members of the Working Party.

These 'position statements' were used to provide directions to the 'expert' judges about what they were to make judgments on. Finally a pooling of all judgments led to selection of a group of items as the outcome of procedures aiming for a scale that adequately sampled and was representative of the attitude objective.

The *predictive or concurrent validity* of a scale is concerned with relating performance on some other measure (in the future or now) with performance on the scale: An outside criterion measure must be involved and the concern with such validity usually relates to some practical problem (*e.g.* the use of attitude scales to identify potential scientists, or to predict achievement in science examinations, or to act as a substitute for other more complex or cumbersome measures). The major difficulty is the absence of suitable criteria. If we wish to use the scale to predict some aspect of future behaviour we have the problem, discussed in Chapter 2, that the behaviour will have numerous determinants. It might be possible to use the criterion-group approach where, for example, the value of scores on 'Interest and enjoyment in science' might be examined as predictors of pupils who would choose to follow a scientific career. However, this appears an unsatisfactory measure for an objective which, it has been argued in Chapter 2, may be achieved by other behaviour than career choice. If a particular attitude objective has been included for the specific purpose of attaining a particular cognitive end, and if a measure of the appropriate cognitive achievement were available, then the level of predictive validity could be established. There is no evidence for such specific affective-cognitive links among these objectives. In the main study (Chapters 11, 12 and 13), the value of attitude measures taken on entry to secondary school as predictors of attitudes and of achievement in science in S2, and the relationships between attitude measures and teachers ratings of pupils, will be examined.

The concept of *construct validity* (Cronbach 1960, page 121) relates scale scores to some *theoretically* defined construct (or attribute) that can explain the variance in those scale scores. In other words the concern is with what it is that is being measured and which leads to individual differences in scale scores, rather than with the scale itself. To establish construct validity it would be necessary to describe constructs that might account for performances on the scale, to predict relationships between these constructs and other variables, and to test these hypotheses empirically. In earlier chapters we have already seen that there is no single generally accepted construct of 'attitude', that the attitude objectives relating to these scales have no substantial theoretical framework from which to derive suitable constructs, and that there is no information on how these objectives relate to any dimensions of attitudes to science that may be displayed by pupils (this question will, however, be considered in Chapter 6). At this time then, it is not possible to carry out a thorough construct validation for these scales.

However, the procedures have, to some extent, led to a description of some curriculum-based constructs (*i.e.* the position statements). These are probably better described as *pragmatic* rather than *theoretical* constructs, but they do permit relationships with other variables to be hypothesized. For example it has been suggested in this chapter and in Chapter 2 that the objectives corresponding to subscales 1, 2, 3 and 5 have strong cognitive components, while that of subscale 4 does not. We would expect that older pupils would have had more opportunities, and be at a more appropriate stage of development, for the necessary cognitive attainment than would younger pupils. We would therefore hypothesise that pupils would score higher on subscales 1, 2, 3 and 5 (but not on subscale 4) in the second year of secondary school than they would on entry to secondary school. This pilot study has explored the scores of pupils in S1 and S2 in two schools. While the evidence cannot be taken as conclusive (since *different* pupils were assessed in the S1 and S2 groups), Table 5.5 shows that for the two schools only one of the eight hypothesised increases in scores from S1 to S2 has failed to occur (although few of the increases are statistically significant).

The scores on subscale 4 ('Interest and enjoyment in science') appear to decrease from S1 to S2. It has already been suggested (see Chapter 2) that the construct corresponding to this objective is rather broad or global. However, whether the construct is interpreted as a predisposition to follow a career in science, or to opt for science courses in S3, or to tinker with science experiments at home, there is a clear relationship with sex and the hypothesis is that boys will score higher than girls on this scale. Table 5.3 and 5.4 provide firm confirmation of this hypothesis.

The weak relationship between cognitive measures (AH4) and the construct which is strongly affective rather than cognitive (subscale 4), and the stronger (and increasing from S1 to S2) relationship between AH4 and the more strongly cognitive constructs (subscales 3 and 5), have already been pointed to under 'Discussion of Results'.

These observed relationships by no means establish conclusive construct validity for the five subscales. Nevertheless, in terms of the pragmatic

TABLE 5.5

MEAN SCORES FOR S1 AND S2 PUPILS FROM TWO SCHOOLS ON EACH SUBSCALE
(figures in brackets show standard deviations)

School 1			School 2	
S1	S2		S1	S2
5.34	6.73	subscale 1	4.18	5.37
(4.53)	(4.59)		(3.91)	(4.93)
1.75	3.95	subscale 2	4.74	5.63
(6.24)	(6.23)		(6.11)	(6.65)
* 7.66	6.78	subscale 3	7.64	8.48
(5.68)	(5.80)		(5.75)	(6.76)
12.26	7.08	subscale 4	8.00	9.52
(8.06)	(10.41)		(10.57)	(10.89)
3.39	5.03	subscale 5	5.39	6.62
(5.11)	(4.69)		(4.91)	(4.10)

*Subscale 3 for school 1 was the only case for which a hypothesised larger mean score for S2 pupils over S1 pupils did not occur.

constructs we have, these preliminary results provide evidence that is consistent with construct validity.

Summary

A group of 5 attitude scales have been developed and shown to have satisfactory levels of internal consistency and stablility. There is reason to assume that the construction procedures have ensured content validity and some degree of construct validity.

From analysis of the attitude scores of pupils taking part in a pilot study the following hypotheses are proposed for the main study:

1. Pupils following the Integrated Science course will be expected at the end of S2 to achieve significantly higher scores than pupils following separate science subjects on scales assessing:
   (i) awareness of the relationship of science to other aspects of the curriculum,
   (ii) an objectivity in observation and in assessing observations.

2. Boys will be expected to achieve significantly higher scores than girls on scales assessing:
   (i) awareness of the contribution of science to the economic and social life of the community,
   (ii) Interest and enjoyment in science.

3. Pupils' scores on a test of general intelligence (AH4) will be expected to predict substantial variance at the end of S2 in their scores on scales assessing:

(i) awareness of the contribution of science to the economic and social life of the community,

(ii) an objectivity in observation and in assessing observations.

The hypotheses relating the type of science course to attitude measures must be tentative. Since this pilot study involved two schools only the significant effects observed may well be *school* rather than *course* effects.

CHAPTER 6

FACTOR ANALYSIS OF ITEM SCORES ON 'ATTITUDES TO SCIENCE' QUESTIONNAIRE

The questionnaire comprising the attitude scales was developed, as described in Chapter 5, from five dimensions of curriculum objectives for pupils' attitudes towards science. Whether or not the five attitude scales are satisfactory, in the sense of being composed of well functioning items, has been determined by the item analysis. However, whether or not such attitudes correspond to dimensions on which pupils can be seen to differ, has not been established.

No claims have been made by the curriculum planners that their objectives correspond to a series of independent attitude dimensions exhibited by pupils, and indeed it is not clear that such correspondence should necessarily be aimed for. It might, however, be fruitful to ask whether the scales are appropriate for assessing *pupil attitudes to science* as distinct from *achievement of attitude objectives*. This would be the case if the items within each scale were not only substantially correlated with each other, but were systematically less highly correlated with items on the other scales.

In order to sort out the apparent dimensions of pupils' attitudes to science, a factor analysis was carried out on the responses of 2815 pupils to each item. These responses were the post-test scores of pupils of the *main* study sample (the sampling procedures will be discussed in Chapter 8, all pupils were in the third term of their second year of secondary school at the time of the post-test).

It was considered more appropriate to carry out the analysis on the main sample rather than the pilot group since Raven, Ritchie and Baxter (1971) found that:

> '*all* the factors which emerged from the pilot, other than those which had been obtained in previous large studies, completely disappeared in the main study. It would appear that where a well established factor is fed into a small sample study it is likely to emerge in the analysis, but the reverse is not true: factors emerging from analyses of data obtained from small pilot samples are less likely to emerge in larger studies'.

This may be an overcautious approach and Guilford (1954 page 533) suggests that 'Factor loadings from samples near 200 have been fairly consistent with loadings in the same factors and tests from samples above 1,000'.

The need for a large sample arises from the problem of estimating errors. We have no satisfactory method of allowing for sampling fluctuations in rotated factor loadings and so must aim to make the errors in inter-item correlations as low as possible. It is also desirable to have normally distributed raw data. However, unless the data is excessively skewed, truncated or multi modal, most distributions can be subjected to factor analysis (Child 1970, page 17). As discussed later (Chapter 11 page 340) the data used here is apparently suitable in nature and distribution for the use of parametric statistics. Perhaps the most important assumption we are making, however, is that the correlations we are using are derived from item scores that are *linearly* related to each other.

Inter-item correlations were computed among scores on the 60 items comprising the questionnaire, and the correlation matrix was subjected to a Principal Components, followed by rotation to the Varimax criterion, and further transformation providing Promax oblique factors. The Computer program used was a modification of the Hallworth and Brebner (1965) STATPAK package.

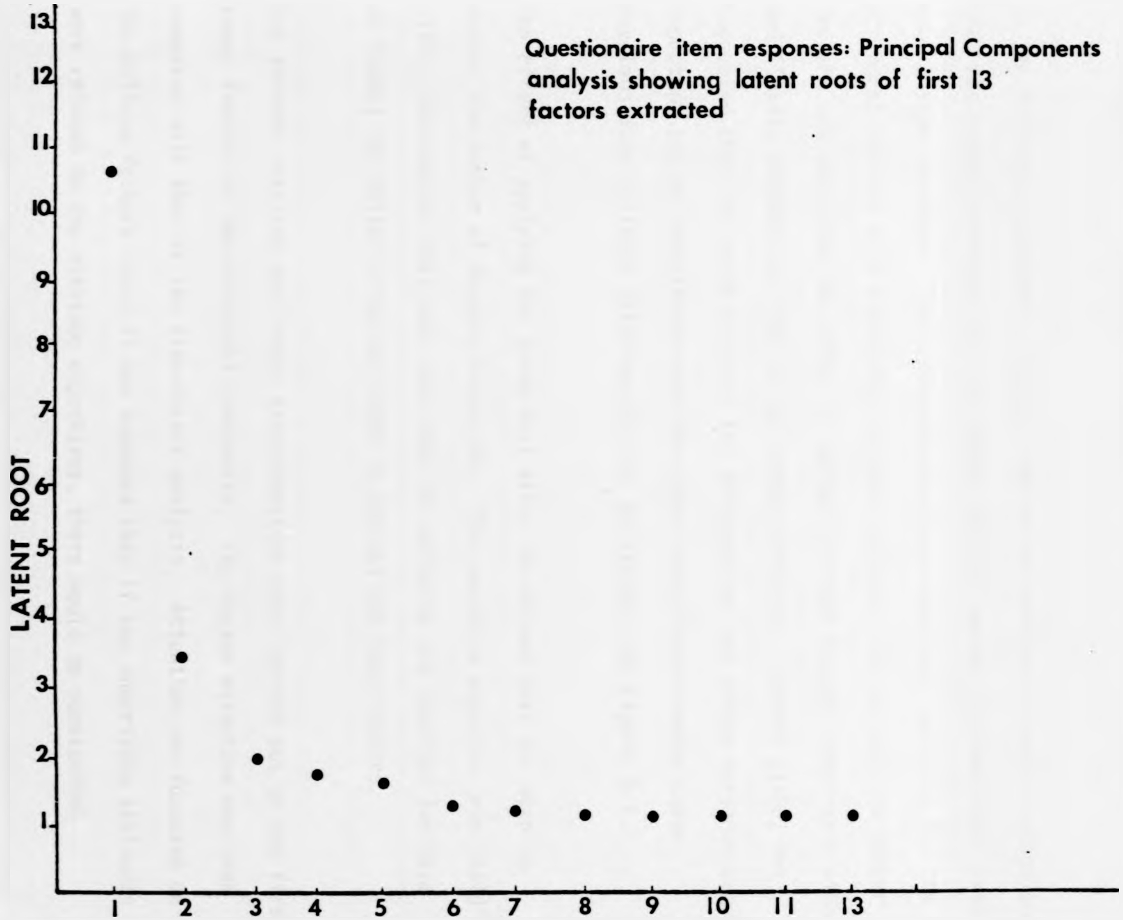
As always, there was the problem of how many factors to extract for interpretation in terms of underlying attitudes. It could be argued that we have hypothesised the existence of five attitude factors corresponding to the five attitude scales fed into the analysis. On the basis of this, one analysis was carried out which extracted *five* Principal Components (followed by Varimax rotation and Promax transformation).

However, the five attitude scales were based on what the planners saw as appropriate *curriculum objectives*, not on some psychological theory of the dimensions of *pupils' attitudes*. It therefore seemed necessary to make the decision on the number of factors to be extracted on the basis of criteria that are independent of the nature of the items. In addition, Guilford (1954, page 500) has warned that researchers are more likely to extract too *few* factors than too many.

With these considerations in mind, the Principal Components analysis was repeated and all factors having latent roots greater than one were extracted (Kaiser's Criterion). A graph of latent root against factor number was then plotted and a scree test (Cattell 1966) applied (see Figure 6.1). This indicated that seven of the thirteen factors should be used in the analysis.

The scree test is based on the growing influence of 'unique' variance as more and more 'common' factors are extracted. 'Unique' variance is that part of the 'total' variance of an item that a) arises from characteristics

Questionaire item responses: Principal Components analysis showing latent roots of first 13 factors extracted



(Fig.6.1)

FACTOR NUMBER	LATENT ROOT
1	10·41
2	3·49
3	2·04
4	1·71
5	1·61
6	1·40
7	1·20
8	1·16
9	1·09
10	1·09
11	1·05
12	1·03
13	1·01

of that item that are not shared with any other item, and b) error variance. The rest of the 'total' variance on that item is 'common' variance *i.e.* variance of the item in some 'common' factor for which other items also have significant loadings. So:

V_t	=	V_c	+	V_u	(see Child 1970, page 34)
total variance of item		common variance		unique variance	

In the Principal Components analysis the unique variance is not distinguished from the common variance, and the common factors become 'contaminated' with some unique variance. This contamination is unimportant initially for the strongest factors with substantial common variance, but as more and weaker factors are extracted the effect of unique variance becomes comparable with, and finally overwhelms, that of the common variance. Cattell (1966) has suggested that the point at which the influence of the unique variance becomes important can be identified where the latent root/factor number curve develops into a linear relationship *i.e.* at factor 7 on Figure 6.1.

The effect of applying the scree test after the Kaiser test has been to reduce the number of factors extracted. This would be expected from Child's (1970) observation that when more than 50 variables are involved (we have 60 items) the Kaiser criterion tends to extract too many factors.

The Varimax rotation and Promax transformation were carried out on the first seven factors of the Principal Components. The factor structure was then compared with that of the five-factor analysis. Attention was focussed on the oblique factors since it was expected that if the underlying attitudes were related to the attitude objectives, there would be substantial correlations between some of them (*cf.* inter-scale correlations observed in the pilot study).

Comparison between the 7- and 5-Factor Analyses

Tables 1B to 12B Appendix B list the items defining each factor emerging from both the 7-factor and 5-factor analyses, together with item loadings where these are greater than 0.3. The 'attitude scale' that the item contributes to is also indicated. Table 6.1summarizes the factor structure of the attitude questionnaire responses when 7 and 5 factors were rotated and translated. The *order* in which the factors emerged from the analysis does not correspond to the Factor Numbers in this table. The factors have been ordered here in accordance with the 'attitude scales' to which they appear to correspond. (Table 6.2 - variance accounted for - indicates both orderings.) Items making up each factor have been ranked in each case according to their factor loadings.

It can be seen from inspection of the table that the first four factors in the two analyses are closely related, and are made up predominantly of items from Attitude Scales 1, 2, 4 and 5 respectively. However, the 5-factor analysis has a number of items corresponding to Attitude Scale 3 included in the factor linked with Attitude Scale 2, and Factor 5 has a 'ragbag' appearance compared with Factors 5, 6 and 7 of the 7-factor analysis. Raven, Ritchie and Baxter (1971) have demonstrated that 'Rotation of fewer factors lead to items being assigned to larger factors in such a way as to obscure factor patterns', and found, as in this case, that the factor structure is blurred as fewer factors are extracted.

TABLE 6.1 Factor Analysis of Pupil Attitude Item Scores Promax Solutions identifying Loadings of Items on Factors when (a) seven factors extracted, (b) five factors extracted. Loadings less than 0.30 omitted. Decimal point omitted.

7 Factors extracted

Item Number	Attitude scale	Factor Number 1	2	3	4	5	6	7
17	-1	-66						
25	-1	-64						
54	-1	-61						
55	-1	-59						
50	+1	57						
59	+1	42						
22	+1	36						
32	+1	32					-32	
37	-2		-62					
60	-2		-55					
15	-2		-54					
42	-2		-51					
53	+2		50					
9	-2		-41					
	-2		-33					
12	+2		+32					
44	-3		-31					
7	+2		30					
57	+4			88				
36	+4			87				
30	+4			85				
39	-4			-83				
41	-4			-80				
24	-4			-80				
4	+4			-80				
31	-4			-64				
47	+4			51				
33	+4			49				
35	-4			-43			-34	
45	-3			-35			-34	
52	-5				-62			
13	-5				-59			
40	-5				-58			
6	-5				-50			
29	-5				-45			
10	-5				-44			
16	+5				42			
48	+5				30	39		
34	+5					57		
38	+3					46		
28	+5					45		
1	+5					44		
49	+3					38		
46	+5					36		
20	+3					33		
18	-4						-69	
26	-3						-60	
14	-3						-47	
21	-3						-44	
45	-3						-34	
5	+2							90
23	+2							84
12	+2							34

5 Factors extracted

Item Number	Attitude scale	Factor Number 1	2	3	4	5
17	-1	-55				
25	-1	-54				
54	-1	-51				
50	+1	47				
55	-1	-44				
32	+1	40				
19	+1	34				
59	+1	33				
27	+1	31				
15	-2		-62			
42	-2		-60			
37	-2		-59			
60	-2		-51			
9	-2		-50			
43	-2		-44			
-4	-3		-44			
44	-3		-42			
8	-1		-41			
7	+2		37			
12	+2		36			
53	+2		36			
45	-3		-34	-43		
21	-3		-33			
2	-1		-31			
51	+3		31			
30	+4			87		
39	-4			-87		
57	+4			86		
36	+4			85		
41	-4			-84		
24	-4			-83		
4	+4			76		
31	-4			-61		
47	+4			58		
35	-4			-53		
33	+4			50		
40	-5				-56	
6	-5				-54	
52	-5				-53	
29	-5				-53	
10	-5				-52	
13	-5				-52	
16	+5				38	
48	+5				36	44
23	+2					52
56	+3					49
49	+3					45
5	+2					45
28	+5					40
38	+3					40
20	+3					39
34	+5					35
22	+1					33
46	+5					32
3	+2					30

In view of this, discussion of the factor pattern will be directed towards the 7-factor analysis. Table 6.2 shows the variance accounted for by each of the 7 Varimax factors and Table 6.3 shows the intercorrelations of the factors after Promax translation.

The Factor Pattern

The first factor is defined by eight items, all of which come from Attitude Scale 1 ('awareness of the inter-relationship of the different disciplines of science'). These items (see Table 1B Appendix B) are concerned with the inter-relationships in a very broad, generalized sense, *e.g.*

25 'There are very clear boundaries separating physics, chemistry and biology'. (negative)

59 'Chemical energy is important to physics'. (positive)

Those four items corresponding to Attitude Scale 1 which do *not* have a high loading on Factor 1 all refer to specific areas of study in one science and the dependence of such areas on information or procedures from another science *i.e.*

19 'If you were interested in studying animals' eyes you would need to know some physics'. (positive)

27 'To study pond life you have to work like a physicist, chemist and biologist all combined'. (positive)

2 'Chemical reactions are of interest only to those who learn chemistry'. (negative)

8 'Biologists studying plants and animals do not need to know anything about electricity'. (negative)

TABLE 6.2

FACTOR ANALYSIS - 7 FACTORS UNDERGOING VARIMAX ROTATION

VARIANCE ACCOUNTED FOR BY VARIMAX FACTORS

(The order of factors is as described in the text and not in order of extraction. Order in which factors were extracted is given by figures in brackets)

	Factor Number	% Variance Accounted For
(4)	1	4.4
(5)	2	5.3
(1)	3	11.5
(2)	4	4.3
(3)	5	3.7
(6)	6	3.8
(7)	7	3.4

Total Variance accounted for = 36.4%

TABLE 6.3

FACTOR ANALYSIS OF PUPIL SCORES ON ATTITUDE ITEMS

CORRELATIONS BETWEEN PROMAX FACTORS -

7 FACTORS EXTRACTED (x 100)

Factor Number	1	2	3	4	5	6	7
1	100	40	28	23	33	-30	37
2		100	43	16	28	-38	38
3			100	01	29	-48	40
4				100	02	-21	14
5					100	-29	45
6						100	-53
7							100

None of these items have loadings as high as 0.3 on any of the seven factors.

Table 6.4 shows correlations between the eight items that define Factor 1. This indicates that, with the possible exception of item 32, they form a cluster of *moderately* correlated items. Raven, Ritchie and Baxter (1971) argue that 'good' attitude scales should ideally have *moderate* rather than *high* correlations between items:

> 'The theory here is that if an attitude is strong it will dominate a person's responses to a series of items which, in addition to tapping one common dimension, also tap a series of other dimensions. The theory assumes that *all* the items raise issues which are complex and that we should respond to *all* of them by saying "in some ways yes, in other ways no". However, if we feel strongly about something we will allow these considerations to overrule all other considerations. It is assumed that the more often we allow one set of considerations to overrule all sorts of alternative considerations the more strongly we feel about the basic issue and the more firmly it is anchored in our mental make-up.
>
> Thus, if we respond consistently to a series of items which have one attitudinal dimension in common but which also tap a *series* of other attitude dimensions then we must have a stronger, more all-pervasive, more organized, or more resistant to change, attitude on the common dimension....
>
> If the items are too highly correlated they will tap only one dimension and our inference that an individual who endorses all of them must have a stronger attitude because more divergent pulls were overcome will be erroneous'.

Factor 2 is defined by nine items from Attitude Scale 2 ('awareness of the relationship of science to other aspects of the curriculum') and one item from Attitude Scale 3. The Scale 2 items are concerned with whether or not a knowledge of science is necessary to, or useful for, the study of

TABLE 6.4

Factor 1 of 7-Factor Analysis. Correlations between Items with Loading 0.3 or Greater on this Factor. (Only correlations of 0.2 or greater are reported. Decimal points are omitted).

Item numbers

Item numbers	17	25	54	55	50	59	22	32
17								
25	4							
54	3	3						
55	2	3	4					
50	-2	-2	-3	-2				
59	-2	-2	-2	-3	2			
22		-2	-2	-3	2	3		
32			-2	-2	2	2	2	

other subjects *e.g.*

7 'Science is very useful to several of my other school subjects'. (positive)

15 'Science does not help someone to learn geography'. (negative)

(see Table 2B Appendix B)

Neither of the two items of Scale 2 concerned with the relationship between science and *mathematics* had high loadings on this factor. However, they had very high loadings on Factor 7:

5 'Mathematics is a great help to science'. (Positive, loading on Factor 7 = 0.90)

23 'Science would be very difficult if we had no mathematics'. (Positive, loading on Factor 7 = 0.84)

(see Table 7B Appendix B)

These two items are concerned with the usefulness of mathematics *to* science, rather than the usefulness *of* science to other subject areas (as are the items defining Factor 2).

One item:

12 'Geography provides examples of things we learn about in science'. (positive)

has loadings of 0.32 and 0.34 on Factors 2 and 7 respectively. It is not clear from this item which subject (Geography or Science) is considered as being 'useful' to the other.

If the inter-item correlations for the two factors are examined (Tables 6.5 and 6.6) it can be seen that item 12 achieves moderate correlations with only 5 of the other 9 items in the cluster defining Factor 2. The same item correlates moderately (r = 0.2) with the other items in the Factor 7 cluster, but this cluster is very small.

The other items defining Factor 2 all intercorrelate moderately with the exception of item 53. This item only achieves a correlation coefficient of 0.2 with 3 of the other 9 items in the cluster.

Eleven out of the 12 items of Attitude Scale 4 ('Interest and Enjoyment in Science') have loadings above 0.43 on Factor 3. The single scale 4 item that does not load on this factor is:

18 'Science is only for brainy folk'. (negative)

It seems reasonable to suggest that a pupil's response to this item may be independent of his interest and enjoyment in science.

If the inter-item correlations for this factor are examined it can be seen that there are a number of high correlations between items that have high loadings on this factor (Table 6.7). This indicates that items such as:

30 'I enjoy science'. (positive)

or

24 'I am not interested in science'. (negative)

tap *only* the attitude in which we are interested. Since there are no

TABLE 6.5

Factor 2 of 7-Factor Analysis. Correlations between Items with Loading 0.3 or Greater on this Factor. (Only correlations of 0.2 or greater are reported. Decimal points are omitted).

Item numbers

Item Numbers	37	60	15	42	53	9	43	12	44	7
37		2	2	3		2	2		2	-2
60			2	2	-2	2	2		2	-2
15				2		2	3	-4	2	-2
42					-2	3	3	-2	3	-2
53									-2	
9							2	-2	2	-2
43								-3	3	-2
12									-2	
44										-2
7										

TABLE 6.6

Factor 7 of 7-Factor Analysis. Correlations between Items with Loading 0.3 or Greater on this Factor. (Only correlations of 0.2 or greater are reported. Decimal points are omitted).

Item Numbers

Item Numbers	5	23	12
5		5	2
23			2
12			.

TABLE 6.7

Factor 3 of 7-Factor Analysis. Correlations between Items with Loading 0.3 or Greater on this Factor. (Only correlations of 0.2 or greater are reported. Decimal points are omitted).

Item Numbers

Item Numbers	57	36	30	39	41	24	4	31	47	33	35	45
57		7	7	-6	-6	-6	6	-4	5	4	-4	-4
36			6	-6	-6	-6	6	-4	5	4	-4	-4
30				-7	-7	-7	6	-4	5	4	-5	-5
39					7	7	-5	4	-5	-4	5	5
41						7	-5	4	-5	-4	5	5
24							-6	4	-4	-3	4	5
4								-5	4	4	-3	-4
31									-3	-3	3	3
47										3	-6	-4
33											-3	-2
35												4
45												

alternatives to be considered, in the form of diverse attitudes tapped by the item, the pupil's decision on his response will be easier and less likely to indicate the intensity of his attitude. Such items do not meet the requirements for 'good' attitude scale items, as proposed by Raven, Ritchie and Baxter (see page 180), in the way that a group of items which tapped a wide range of diverse attitudes, in addition to the particular one of interest, would be expected to.

Factor 4 is defined by eight of the twelve Attitude Scale 5 items, and six of the eight are 'unfavourable' items (see Table 4B Appendix B). These items appear to be those concerned with the source of scientific 'truths', or 'final answers' to scientific questions *e.g.*

52 'Good scientists know the true laws of science'. (negative)

16 'A good scientific theory does not supply the final answer to scientific questions'. (positive)

Table 6.8 shows the inter-item correlations for this factor. There is a fairly good pattern of moderate correlations, though items 13 and 16 are not so clearly members of the cluster as are the others.

Item 48, although included in Factor 4, also has a loading of 0.39 on Factor 5. This factor, unlike Factors 1 to 4, is not bipolar. It is defined by five positive Attitude Scale 5 items and three positive Attitude Scale 3 items (see Table 5B Appendix B). The general concern of these items is with the need of science to be supported by, and to have its strengths and limitations recognized by, everyone *e.g.*

TABLE 6.8

Factor 4 of 7-Factor Analysis. Correlations between Items with Loading 0.3 or Greater on this Factor. (Only correlations of 0.2 or greater are reported. Decimal points are omitted).

Item Numbers

Item Numbers	52	13	40	6	29	10	16	48
52		3	2	2	2	2		
13			2	2				
40				2	2	2	-2	-2
6					2	4		-2
29						3		-2
10								
16								2
48								

34 'Lots of information we get from science now will be changed in the future'.

38 'Science needs the understanding and support of ordinary people'.

If we look at the inter-item correlations on this factor (Table 6.9) we do not have a pattern that indicates a satisfactory cluster of items. Out of the 28 intercorrelations only 4 reach the moderate level of r = 0.2. Four items (1, 20, 34 and 46) do not achieve moderate correlations either with other items in this cluster or with any of the other items in the questionnaire. Three possible explanations for this absence of substantial correlations are:

a) most pupils gave the same response to the item, leading to an inadequate distribution of scores,

b) the item represented an attitude dimension untapped by any of the other items,

c) ambiguity of the item led to a number of different interpretations (and so responses) by different pupils.

If any of these explanations is applicable, then the items will be unsatisfactory contributors to the attitude factor.

Four items from Attitude Scale 3, two from Scale 4 and one from Scale 1 define Factor 6. This factor is not bipolar and the items reflect the perceived remoteness or alienation of science from the pupil and other members of the general population *e.g.*

18 'Science is only for brainy folk'.

14 'Only people who are going to do scientific work should have to learn science'.

TABLE 6.9

<u>Factor 5 of 7-Factor Analaysis. Correlations between Items with Loading 0.3 or Greater on this Factor. (Only correlations of 0.2 or greater are reported. Decimal points are omitted).</u>

Item Numbers

Item Numbers	34	38	28	1	48	49	46	20
34								
38			2		2	2		
28					2			
1								
48								
49								
46								
20								

Inspection of Table 6.10 which provides the inter-item correlations shows a satisfactory cluster pattern of moderate correlations for all but item 32.

From this discussion it seems that five of the factors (1, 2, 3, 4 and 6) are defined by items suitable for assessing pupils' expressed attitudes. Factor 5 is rejected because of the exceptionally low correlations between the items, and Factor 7 has too few items for serious consideration.

Table 6.3 shows the Promax factor intercorrelations. In general they are moderate (as expected) and two only will be commented on. Firstly, if we ignore Factors 5 and 7, the correlation coefficient with the greatest magnitude is that between Factor 3 (Interest in Science) and Factor 6 (Remoteness of Science) where $r = -0.48$, and, as anticipated, the pupil who views science as 'interesting and enjoyable' also views it as low on 'remoteness' from himself. Secondly, the lowest coefficient is that between Factors 3 and 4 ($r = 0.01$). One of these factors corresponds to a dimension which is predominantly affective ('interest') and the other to one which, it has been argued elsewhere, is predominantly cognitive ('scientific truths'). The scores on the Attitude Scales corresponding to the curriculum objectives (see Table 5.2 Chapter 5) also indicated that there was no relationship between the achievement of an objective that was largely affective (Scale 4) and achievement of an objective that was strongly cognitive (Scale 5). Although Factor 7 is defined by very few items, the substantial negative correlation with Factor 6 is interesting. It appears that pupils with the greatest alienation from science are those with the least appreciation of the contribution of mathematics (and possibly other disciplines) to science.

TABLE 6.10

Factor 6 of 7-Factor Analysis. Correlations between Items with Loading 0.3 or Greater on this Factor. (Only correlations of 0.2 are reported. Decimal points are omitted).

Item Numbers

Item Numbers	18	26	14	21	35	45	32
18		3	3	2	2	3	
26			3	2	2	3	
14				2	3	4	
21					2	2	
35						4	
32						-	

Finally, although it appears that four of the five attitude objectives relate quite closely to dimensions on which pupils appear to differ, the factors extracted in this analysis were only able to account for 36% of the variance in scores (see Table 6.2). This suggests *either* that the pupils' responses to the questionnaire items were predominantly idiosyncratic, *or* that there are a number of further dimensions that are untapped by substantial *groups* of items (they may, however, be tapped by *individual* items). The latter explanation implies that a comprehensive study of pupils' 'attitudes to science' should reveal more dimensions (in addition to the 'Alienation from Science' factor identified in this analysis) than those related to the scales used in this study.

CHAPTER 7

ASSESSMENT OF ATTITUDES - RESEARCH QUESTIONS

The publication in 1969 of Curriculum Paper 7 provided encouragement from the Scottish Education Department for those schools (assumed to be the majority) teaching three separate science subjects to pupils in S1 and S2 to change to a pattern where all the science was taught by one teacher. By 1971 the view of the inspectorate was that about 50% of schools were adopting the one-teacher pattern. The concern of this curriculum with affective objectives has already been discussed in Chapters 1, 2, 3 and 4, and Curriculum Paper 7 finishes its section on S1 and S2 science by saying 'There is here a wide field for further investigations of a similar kind [evaluation of the syllabus] including assessment of attitudes' (page 31).

The planning of this empirical research started with an attempt to identify the *research questions* with which Curriculum Paper 7 was concerned in asking for 'assessment of attitudes'. These questions would be expected to lead to formulation of *hypotheses* which would determine the *research design*. This chapter provides a discussion of the research questions and attempts to formulate some substantive hypotheses.

Research Questions

The scope for basing such questions on a *theoretical* framework provided by Curriculum Paper 7 has been limited, as discussed earlier, and has provided only those hypotheses identified at the end of Chapter 4. The empirical work of the pilot study suggested further hypotheses to be investigated (Chapter 5), but these appeared inadequate in coverage for the sorts of questions that teachers and HMI's were asking.

Discussions were held with individual HMI's and teachers involved in S1 and S2 science teaching, to identify a) what they considered to be the *important* questions related to the assessment of attitudes and b) any *information* that would aid the development of research hypotheses. The selection of *important* issues appeared idiosyncratic, and the *information* supplied was often conflicting and based only on personal experience.

However, there did seem to be agreement that there was a need to assay the 'attitudes to science' among pupils in all schools, *i.e.* to assess the attitude characteristics of the whole population. The advantages of the wide scope of survey research were seen as outweighing any disadvantages of lack of depth in exploration of relationships. The teachers, in particular, suggested on the one hand that much curriculum development in the past had been in the hands of teachers and HMI's whose teaching experience had been in schools where the abilities of the *pupils* had been atypical, and on the other hand that the evaluation of curricula had itself been carried out in carefully selected pilot schools which might well be atypical in the quality of their *staff*.

Survey research is time consuming and expensive. Apart from the major operations of sampling and development of instruments, the collection and processing of data is a massive procedure compared with that in case study research. Nevertheless, it was decided that in order to get as accurate a picture as possible (within sampling error ranges) of the 'attitudes to science' of pupils during the first two years of secondary school, a longitudinal study over two years of a minimum of 3000 pupils would be undertaken (this represents about 4% of the total S1 school population).

From the discussions with the interested parties the following main research questions were identified.

(i) *How does the achievement of affective objectives over S1 and S2 relate to the different types of science course experienced by the pupils?*

Chapters 4 and 5 have already provided the discussion and hypotheses related to this question.

(ii) *How is the achievement of affective objectives related to characteristics of the school such as its size, its location, its denominational status, the socio-economic level of its catchment area, and the average level of ability of its intake?*

Inclusion of these sociological variables of the school resulted from various expressed opinions such as the view that large schools are impersonal and inhibit the development of favourable attitudes towards school and school subjects; or that small schools, particularly those that have developed from junior secondaries, have too restricted an outlook and area of competence to develop satisfactory attitudes to science among their pupils; or that schools in the cities were so preoccupied with problems of discipline that there would be no chance to spend time and effort on the fostering of favourable attitudes to science; or that pupils from industrial areas would be expected to develop more interest in science because of the ease of demonstrating the direct relevance of science to their future occupations; or that the attitudes displayed by the pupils depended on the general background and level of ability of the school's intake; or that attitudes and ability were unrelated and some low achieving non-certificate pupils showed the most favourable attitudes to science.

Much of such opinion was based on limited personal experiences and was frequently contradicted by the opinions of other individuals. No theoretical framework for the relationship between attitudes to science and school variables such as size, denomination and location has been proposed. Empirical data on such relationships is sparse. The Science 5/13 (University of Bristol, School of Education 1970) found that improved attitudes towards science among this younger group of children clustered with a number of other variables including school size and location. This 'positive' group indicated that primary schools with eight classes or less, with suburban catchment areas, following an integrated day, with pupils working at their own science tasks individually or in groups, and using active discovery methods, could be expected to show an increase in Attitude to Science Activities test score. In other words 'findings which simply confirm one's feeling that small, flexible schools are more likely to foster positive attitudes than large, less personal ones' (Wynne Harlen private communication).

For the 5/13 study then the variables of location and size belong to a cluster of factors related to the 'climate' of the school. Anderson and Walberg (1968) also found that attitudes were related to pupils' perceptions of their classroom climate. However, we have no reason to believe that the size and location of secondary schools in Scotland is systematically related to the climate in their science classrooms, and so can hypothes:.c no relationship between these variables and pupils attitudes to science.

Clarke (1972) attempted to identify the differences among urban, suburban and rural childrens' *particular* interests in science. His sample consisted of 776 pupils from three classes in each of 6 schools in Massachusetts, U.S.A.. He found no significant differences, for the sample as a whole, among numbers of pupils from the 3 communities with regard to their preferred area of science interest (biological, earth or physical science). There was some indication of a community x sex interaction effect: urban and suburban boys preferred physical science, urban girls preferred biological and earth sciences, suburban and rural girls preferred biological sciences, and rural boys showed no particular interest in science.

There has been some speculation that pupils in non-denominational schools would be expected to achieve the affective objectives to a greater extent than would pupils in Roman Catholic schools. This view has been based partly on a stereotype of a supposed style of teaching in Catholic schools (authoritarian, didactic) that is assumed to inhibit the formation of favourable attitudes, and partly on the statistics which show the unfavourable position of Roman Catholic schools on such measures as numbers of early school leavers, numbers of pupils achieving 'O' grades, and teacher shortages. In the January 1971 Scottish Education Department staff survey the shortage of qualified science staff in Roman Catholic schools was 16% (the 'shortage' is the difference between the actual number and the 'requirements' as laid down by Circular 714, Scottish Education Department), and 4.5% in non-denominational schools. This may have been due, in part, to the Roman Catholic schools being predominantly located in the west of Scotland where the shortage was more acute than in the east. However, in

12 of the 17 counties having both non-denominational and Roman Catholic schools, the percentage shortage of science teachers was greater for the Roman Catholic.

These sorts of constraints may or may not influence the development of attitudes to science. To formulate a positive hypothesis it would be necessary to relate the various characteristics of the schools to a theory of attitude development.

The only study known to this author that attempted to compare such attitudes of pupils attending Roman Catholic schools with those of pupils attending non-denominational schools was carried out in Sydney, Australia by Meyer (1963). He used Thurstone-Likert tests to assess attitudes towards science as a social concept. He categorized pupils, firstly, into 'Dislike Science', 'Indifferent to Science' and 'Like Science' groups, and secondly into groups for and against science as a subject in the school curriculum. He found no significant differences in the distribution of numbers of pupils so categorized among Roman Catholic, state selective, state non-selective or independent schools.

There have been some empirical studies on the relationship of socio-economic status to attitudes to science among *individuals*. There is, however, a paucity of work on the relationship of socio-economic status of the catchment area of the school to attitudes to science among *groups* of pupils. Lowery (1967) sampled 11-year old pupils from 12 classes in three socio-economic areas of California, U.S.A., and assessed their attitudes to science before and after the 'treatment', based on new curriculum material, had been administered to half of these classes. He did not use means of *group* mean measures for his analysis, but compared means of *individual* pupil measures. He found that on pretest and post-test pupils from the upper socio-economic area held attitudes to science that were significantly higher (as measured by t-ratio) than those of pupils from middle or lower socio-economic areas, but there were no significant differences between pupils from middle and lower areas. However, *gains* in attitude scores after 'treatment' were highest for pupils from schools in the lowest socio-economic area. Gain scores are difficult to interpret (this problem will be considered in detail in Chapter 11). It is frequently the case that gains are negatively correlated with initial scores since there is more room for improvement among those who score low initially. The pupils from the higher socio-economic areas may have started off with very favourable attitudes with little potential for gain and in this event, the results would not be reflecting any relationship between attitude and socio-economic status.

It is difficult to relate Lowery's findings directly to this study since it is not at all clear what correspondence there is between the attitudes that he investigated and the ones with which we are concerned. He used a Projective Test of Attitudes which

> 'is an open-ended instrument designed to uncover and measure attitudes which are overlooked by the usual techniques. It is designed to uncover attitudes of students toward the field of science, the scientific process and the scientist as an individual'.

Nevertheless, his findings are in agreement with the general findings (*e.g.* Neale and Proshek 1967, Glick 1970) that pupils from schools with high socio-economic catchment areas have more positive attitudes towards school and teachers than do pupils from schools in low socio-economic areas.

A substantial number of studies have investigated the relationship between an *individual's* intelligence and attitudes to science but information on the relationship of *group* mean attitudes and mean levels of intelligence is difficult to find. Most of the work in the area is concerned with 'interest in science' and not with attitudes that have strong cognitive components. Meyer (1961a) cites eight studies which suggest that there is no significant relationship between a pupil's science interest and intelligence, and, in his own study of 150 London 11 to 15 year olds, confirmed this view. Clarke (1972) found no significant effect of IQ on pupils' ratings of science and other school subjects on a Semantic Differential scale characterizing the subjects on factors of 'Vigor and Certainty'. However, Aiken and Aiken (1969) report two unpublished studies which identified

> 'positive relationships between intelligence and favourable attitudes towards science and scientists...and found that in a sample of over 1000 elementary-school children, intelligence was related to science attitude and science information'.

It may be the case that a positive relationship exists between intelligence and attitude to science where the particular attitude under consideration has an emphatic cognitive component.

There has been considerable argument since the early 1950's that general intelligence tests have been very limited in the abilities that they assess and the sorts of item that they produce. The result of this has been a plethora of studies using tests of divergent thinking which purport to tap different abilities from those tapped by the intelligence test. The distinction between the different types of ability is most clearly made in Guilford's model of the intellect (1956). Divergent thinking is open-ended and used to tackle problems where there are numerous 'right' answers or no 'right' answers. This is contrasted with convergent thinking (assessed by an intelligence test) which is used to solve a problem which has one correct answer.

If divergent and convergent abilities are sufficiently independent (*i.e.* a substantial proportion of the variance in divergency measures is *not* shared as common variance with convergency measures) then we might expect that measures of divergent thinking would be able to predict additional variance in pupils' learning *after* the prediction by measures of convergent thinking had been allowed for. Unfortunately, and remarkably as Butcher (1972) observes, very little work on such predictive validity of divergency measures, over and above that of convergency measures, has been carried out.

Much of the interest in the value of divergency in the prediction of attitudes to *science* has been generated by Hudson's (1966) widely read study of English Public School and Grammar School boys in which science specialists are identified as 'convergers' and arts specialists as 'divergers'. If 'convergency' and 'divergency' are distinguishable abilities it would be expected that scores on tests of divergency would not correlate highly with intelligence scores. Getzels and Jackson (1962) working, like Hudson, with children of high ability found intercorrelations of 0.131 to 0.378 between IQ and divergency test measures. On the other hand Hasan and Butcher (1966) working with 175 Scottish pupils from the whole ability range (*i.e.* a sample corresponding closely to the pupils of this study) found much greater overlap between IQ and divergency (intercorrelations 0.317 to 0.772)

Butcher (1968 page 103) reports six articles which provide an explanation in the form of a

> 'plausible "threshold" suggestion about the relation between general intelligence and creative ability, which suggests that up to a level of about IQ 120 general intelligence is the most important factor, particularly in predicting school achievement, but at levels about this creative abilities begin to assume more importance'.

In a later paper (Butcher 1972) he reports an unpublished study by Nuttall (1971a) in which 600+ fifteen and sixteen year-old pupils had their scores on a battery of convergent and divergent measures factor analysed. Three analyses were carried out for a) all the pupils, b) those pupils from grammar schools, and c) non-grammar school pupils. Butcher states

> 'In the total group it was clear that no factor of divergent thinking emerged; each test of divergent thinking loaded on a different factor. The analysis of the grammar school data produced a very different result. In an orthogonal solution, the second factor was clearly interpretable as one of divergent thinking, the three "established" tests of this type had loadings between 0.64 and 0.78 but of the other ten variables only one loaded over 0.12. An oblique solution produced a closely similar pattern. A third analysis (of the non-grammar school group) produced a pattern of factors and loadings similar to that for the total group....This greater differentiation of divergent and convergent ability in the more able group (whose mean score was higher on divergent as well as convergent measures)...seems very relevant to the threshold thesis whereby divergent and convergent abilities are more readily distinguishable in high ability groups'.

In view of the fact that our study is concerned with pupils of the whole ability range in S1 and S2, and not just with the more able groups, there is no reason to believe that divergency measures will be of value in predicting learning over and above that predicted by measures of general intelligence. However, in view of the interest shown in this area by the HMI's and teachers it was decided to include variables related to measures of divergency.

(iii) How is the achievement of affective objectives related to characteristics of class organization such as class size, ability grouping, and number of periods of science?

Curriculum Paper 7 states that 'Since this is essentially a practical course a class of 20, or less is frequently a reasonable number' (page 46). A number of schools have not been able to keep class sizes down to this level. It has been suggested that the Curriculum Paper's recommended pattern of 'guided discovery' is impossible to follow with large classes. Some teachers assert that 16 is the maximum number of pupils that can be coped with if this course is to be adequately followed. The assumption is being made that large classes are less likely to develop favourable attitudes to science than will smaller classes. This relates closely to Getzel's (1969) view that the small group will generate a more favourable social climate through greater opportunities for personal interactions and individual participation than will a large group.

McKeachie (1963) has reviewed studies relating learning to sizes of discussion and lecture classes. He concludes that group size is a much more relevant variable in the discussion classes than in the lecture, and cites studies which found higher levels of satisfaction and less teacher dominance in small discussion groups (page 1143). He also concluded that small lecture classes were superior in affective outcomes. He cites a series of 4 studies using a battery of criterion measures that found 'significant differences favouring small classes...on measures of student attitudes toward all the courses' (page 1131). He sums up by saying

> 'large lecture classes are not generally inferior to smaller lecture classes if one uses traditional achievement tests as a criterion...[but]...The weight of the evidence seems to favour small classes if one uses student or faculty satisfaction as a criterion' (pages 1131-1132).

These conclusions are based on studies of much older pupils in very different institutions from those of this study. However, one point that McKeachie makes may be directly relevant to this case. He suggests that

> 'One unplanned consequence of increasing class size may be a restriction upon the teacher's freedom to vary his methods to fit his objectives' (page 1132).

Teaching towards attitude objectives probably requires different, and a greater variety of, methods than the traditional cognitive-oriented programme. Consequently such constraints on the teacher may assume some importance.

Walberg (1969) has investigated the influence of class size on pupils' 'Science Interest', 'Physics Activities' and 'Physics Interest'. The correlations between class size and each of the criteria were not significant.

Generalizations are difficult to make in this area since the comparisons of some studies are between groups of 4 and 10 as the extremes, while in others they are between groups of 12 and 150. Hypotheses on class sizes of 15 to 30 pupils cannot confidently be formulated.

The Integrated Science course has been developed specifically for mixed ability grouping. It does not *necessarily* follow that achievement of its objectives among mixed ability groups should be greater than among streamed or banded groups, but the intention that this should be the case is implied.

In general Khan and Weiss (1973) observe that 'Research on affective aspects of ability grouping is practically non-existant'. For those committed to mixed-ability grouping there appears to be a fundamental belief in the affective benefits acruing that is based on personal experience rather than systematic research. Kelly, A.V. (1974) states that

> 'It has been noted that pupils' attitudes to school become more positive...and this improvement permeates all aspects of their work. In schools where they have been given a limited opportunity to work in mixed-ability situations... improvements have been detected in their attitudes to work' (pages 84-85).

However, he warns that

> 'this is not an inevitable concomitant of the change; it is the result of the way the teachers use the advantages the change gives them to develop more productive relationships with their pupils' (page 85).

It is unlikely that the information on whether or not a class has a mixed ability composition will, in itself, substantially predict class attitudes to science. Only if there is some reason to identify mixed ability grouping with particular teaching methods, will hypothesis formation in this area become feasible.

The suggestion is made that 'For science, one eighth of the teaching time is recommended, to be distributed by periods, 5-5, 4-6, or 6-4 over the first two years' (Curriculum Paper 7, page 17). Some schools have allocated less time than this to science, and teachers have claimed that a 4-4 programme does not provide enough time for attention to be given to development of attitudes. Schools following separate subject courses are more likely to provide 6 (2, 2, 2) periods for the three sciences.

If, as suggested in Chapter 2, the development or modification of attitudes is dependent on acquisition of cognitive learnings, then the more time devoted to science the more opportunities there will be to develop such cognitions. However, whether or not the difference between 4 and 6 periods per week is crucial has not been established.

(iv) How is the achievement of affective objectives related to pupils' sex, social class, general intelligence measures, divergency measures, perceived achievement in science, and teachers' ratings?

A wealth of studies (reviewed by Gardner 1974(a), 1975, Kelly, A. 1974) have demonstrated the higher level of interest in science among boys than girls. The pilot study described in Chapter 5 provided similar evidence among S1 and S2 pupils in Scotland.

It seems to make little difference whether 'interest' is operationalized as career preference (Butcher and Pont 1968), or as most preferred school subject (Sumner and Wilson 1972), or as interest in reading or talking about science (Hilton and Berglund 1974), or as interest in practical experimenting at home or repairing electrical appliances (Walberg 1967), the same differential effect is observed. However, the differences are reduced among older groups of science pupils. The girls among these groups are normally in the minority and are a very select sample of the female population, most girls having elected to study something else. Gardner (1974a) has used Welch's (1969) data to demonstrate that there is some evidence suggesting that girls who elect to study science 'are an extremely select group of the population, whereas boys are less extreme'.

The literature on attitudes to science other than 'interest' is much more limited and the results are not nearly so clear cut. The evidence that is available is largely from rather older pupils than those of this study. Alexander (1972) and Nuttall (1971b) used the NFER Science Attitude Questionnaire (Skurnick and Jeffs 1971) with non-certificate pupils (Nuffield Secondary Science) and GCE pupils respectively. Neither identified any significant sex differences on the 'Social Implications of Science' scale. Omerod (1971) in a study of English 14 year-olds found that the relationship between choice of science option and score on a 'Social Implications of Science' scale was stronger for girls than boys. These girls were of the 'extremely select group' discussed above and unlikely to display attitudes of unselected groups of girls. Walberg (1967), also assessing a group of pupils who had opted for physics, found that girls scored significantly higher than boys on an 'Academic' science activity scale ('scholastic' rather than 'direct involvement'), on 'Nature Study' and 'Applied Life'. In contrast, Cooley and Reed (1961) found, in a study of a younger (14 year-old) unselected group, that boys scored significantly higher than girls on a composite score from the same three scales together with two other scales. Another composite measure was used in the IEA study (Keeves 1973). Weighted measures on scales assessing 'View of Science in the World', 'Mathematics as important to Science' and 'Interests and activities in Science', were computed, and the scores of a minimum of 1000 fourteen year-old pupils in ten countries were analysed. On these composite scores boys displayed more favourable attitudes to science than girls (for Scotland the correlation between sex and attitude was 0.23, about the median point for the ten countries).

The pilot study described in Chapter 5 identified significantly higher scores for boys on the 'Economic and Social Implications of Science' scale as well as on the 'Interest and Enjoyment in Science' scale for the fourteen year-old (second year) pupils.

The evidence points clearly to the hypothesis that scores for boys on scales of Interest in Science will be significantly higher than those for girls. For other, more strongly cognitive, attitude scales the evidence is inadequate for any hypotheses to be formulated with confidence.

The general relationships between socio-economic status and attitudes have been discussed under question (ii). Clarke (1972) concluded from his study of 776 urban, suburban and rural children that the social class status of the individual pupil had no influence on that pupil's preference for science in relation to other school subjects. Meyer (1961b) was unable to find any significant relationship between social status of father's occupation and pupil's interest in science in either his own empirical study of 150 London schoolchildren or in the relevant literature. However, there is some evidence that social status or parental occupation is related positively to choice of a career in science (Brown 1953, Terman 1954, James and Pafford 1973).

Schwirian and Thomson (1972) found that 'typically the higher the occupational status of the student's father the more positive their attitude toward science'. The particular dimensions of attitude with which they were concerned were developed from a theoretical position on those

> 'particular cultural values which are more conducive to growth and development of the scientific institution in society than are their opposites' (Schwirian 1968).

These values were identified (after Barber 1962) as

> '*Rationality:* "The modern world thinks the rule of reason more important than the rule of custom and ritual"....
> *Utilitarianism:* "By the value of utilitarianism, we mean the predominant interest modern man has in affairs of this world, this natural world"....
> *Universalism:* Approval of the cultural value of universalism is clearly manifest in...assuring a man his just position in society based on his abilities and performance.
> *Individualism:* The approval placed on the value of individualism serves...as a force in determining the extent and focus of the *process* of rationality...When individualism is an agreed-upon value, members of the society are committed to the moral preference of individual conscience rather than for those of organized authority.
> *Progress and Meliorism*...The final cultural value, then, which is necessary for the optimum growth and development of science is that active rationality can and should improve man's lot in this world, coupled with belief in and approval of progress'.

Each of these values were operationalized in the context of the scientific institution for construction of the attitude scales. Their findings were of undergraduate's attitudes and the scales do not relate directly to any of the attitude scales of this study. Nevertheless, the results are of interest since they relate to attitudes with strong cognitive elements unlike most other studies which have concerned themselves with pupils *liking* for science.

Insofar as this study is concerned with some attitudes that are also strongly cognitive it might be suggested that there would be some relationship between attitude social class. However, in view of the lack of evidence from comparable groups of pupils, such a hypothesis is considered too tenuous.

The evidence relating attitudes to various measures of pupils' intelligence has been reviewed already under research question (ii). Attitude/

divergency relationships have been largely unexplored, and attitude/IQ relationships have been found to be non-significant or weak.

The relationship between attitudes and achievement in science is by no means established. Meyer (1961a) refers to seven studies, in addition to his own, which have failed to reveal such a relationship, but he also refers to six others which have identified positive but weak relationships.

Wynn and Bledsoe (1967) investigated the relationship between achievement (science, biology and overall academic) and *changing* interest in science. Their conclusion was that

> 'interest change is a phenomenon in itself which is unique, and important in itself, independent of such factors as intelligence, academic achievement and home background'.

Wick and Yager (1966) used the *change in grades* the pupil receives as his perceived achievement 'success' in a science course, and related this to *change in attitude*. They hypothesized that

> 'If the grades do have a considerable influence on the students' attitudes, we would find that...more positive attitude toward the course was caused by students' increased success, as measured by their grades'.

However, when they examined the changes in grade of 20 sixteen to seventeen year-old pupils, who had all shown a substantial net *gain* on an Attitude towards School Science scale over one year, they found that:

8 showed increasing grades,

6 showed no change in grades,

6 showed decreasing grades.

Examination of the grades of 20 pupils showing corresponding *losses* in attitude scores showed that:

9 showed increasing grades,

4 showed no change in grade,

7 showed decreasing grades.

They conclude

> 'We cannot discount the attitude-grade relationship, but we are tempted to drop it to a lower level of importance'.

Other writers have, however, found some positive relationships. The NFER scales of the *Science Attitude Questionnaire* (Skurnick and Jeffs 1971) are actually *validated* in terms of their correlations with science examination grades.

Welch (1969) found that the level of 'Course Satisfaction' with high school physics among a group of 68 boys was related ($r = 0.35$, $p<0.01$) to the *change* in their scores on the 'Physics Achievement Test' (administered before and after the course), and to the *change* in their course grades ($r = 0.28$, $p<0.01$). However, 'Course Satisfaction' was unrelated to *initial* achievement scores ($r = 0.02$, $n = 125$, N.S.). For 82 girls, 'Course Satisfaction' was related to *initial* achievement scores ($r = 0.33$, $p<0.01$) but not to *changes* in scores ($r = 0.22$, $n = 41$, N.S.). Like the boys their changes in course grades were positively related ($r = 0.35$, $n = 41$, $p<0.01$) to their 'Course Satisfaction'.

The IEA study in seventeen countries (Comber and Keeves, 1973) identified positive relationships around 0.2 to 0.3 between science interest and achievement among 14 year-old pupils (and rather higher for older pupils who had, for the most part, opted either in, or out, of science by that time).

Few of these studies indicate that there is as much as 10% common variance between attitudes to, and achievement in, science, unless a rather older age group than that of this study is used.

(v) How are teachers' behaviour and attitudes related to pupils' achievement of affective objectives?

As Khan and Weiss (1973) observe 'it would appear that the relationship between teachers' and students' attitudes has been regarded as axiomatic with no need for empirical research' (page 774). The empirical work on science-related attitudes has certainly been sparse and in general has revealed either weak relationships or none at all.

In an attempt to examine factors that influence pupils preference for, and enrolments in, high school physics courses, Welch and Walberg (1967) administered a questionnaire to obtain teachers' opinions on a number of topics regarding physics teaching. They concluded that there was nothing in the attitudes and beliefs of the teachers that appeared to be a direct cause of low physics enrolments.

Rothman, Welch and Walberg (1969) found no relationship between pupils' attitudes to physics and teachers' scores on the Minnesota Teacher Attitude Inventory. Aiken and Aiken (1969) report an unpublished study (Taylor 1965) in which teachers' attitudes were found to be unrelated to 10 year-old pupils' attitudes towards science materials.

Aiken and Aiken also refer to another unpublished report (Bixler 1958) of a study of 1000 elementary school children from 62 classes which provides some support for the view that teachers' attitudes towards science influence pupils' attitudes towards the subject. Greenblatt (1962), again working with elementary school pupils, observed a similar effect. Rothman (1969) in a 'pseudo-replicate' of the Rothman, Welch and Walberg study, found that teacher attitudes were significantly, but in general weakly, related to students' attitudes towards physics. The relationship was complex:

> 'Students find physics easier, and they gain understanding of the universe but lose interest in physics when taught by a teacher convinced of the importance of learning physics. If in addition the teacher indicates that he has an esthetic appreciation of science (Universe-Beautiful), his students tend to find physics and the learning of physics less orderly. However, students of teachers who feel that although physics is complex, learning about it is orderly, exhibited the greatest gain on the Understandable scale'.

The Science 5/13 project (University of Bristol, School of Education 1970) found that pupils' increasing scores on an attitude to science activities test clustered with teachers' familiarity with and favourable attitude towards Nuffield Junior Science, with teachers' satisfaction in their own way of dealing with science activities before the trial, and with teachers' approval and appreciation of Science 5/13 ideas and objectives. This cluster also contained items relating to teachers' classroom behaviour: use of discovery approach, teacher reacts to flagging interest by suggesting fresh approach, individual or group working, children given some choice of science activities and allowed to work outside the classroom.

There may be a stronger relationship between pupils' attitudes and teachers' classroom behaviour rather than teachers' attitudes. Gardner (1974c) found that pupils' enjoyment of physics was related to pupils' perceptions of their teachers as intellectually stimulating. However, the teachers' behaviours appeared to have little influence on other attitudes towards physics (*i.e.* 'physics learning discovery based', 'physics openness', and 'scientists normal'). Other studies have found that teachers perceived as warm (Reed 1961), well prepared (Elliott 1971) and encouraging of pupil participation (Tisher and Power 1973) tend to foster more favourable attitudes to science among pupils than do other teachers.

Gardner (1974b) has suggested that process-product studies which investigate the effects of teacher behaviour on pupil learning have failed to take account of the differential effects that such behaviour may have on pupils of various abilities, or levels of motivation, or personalities. He was able to demonstrate that when *individual pupil* personality variables were included in the analysis:

> 'Achievement-pressing teachers tend to have a beneficial effect on the enjoyment of physics of highly achievement-motivated pupils, but a relatively deleterious effect upon pupils with low achievement motivation. Very serious teachers tend to have a beneficial effect upon already more favourable attitudes of serious pupils, but a deleterious effect upon the attitudes of playful pupils'.

He suggests that techniques employing *class mean* attitude scores will fail to uncover such relationships.

No clear picture of the relationships between teacher characteristics and pupils' attitudes has emerged. If the teachers' attitudes do influence the pupil, then it is likely that this is done through the medium of teachers' classroom behaviour. The problem is still either the identification and assessment of this behaviour, or the identification and assessment of those attitudes or ideologies which might allow prediction of this behaviour.

(vi) Do pupils' attitudes to science change over the first two years in secondary school and, if so, in what direction? To what extent are attitudes assessed on entry to secondary school valuable as predictors of attitudes held at the end of two years?

There appear to be few longitudinal studies of attitudes to science among pupils. There are some cross sectional studies but often these, like the IEA work cited earlier, assess populations that are different in nature (*e.g.* older pupils who have *chosen* to continue with science and younger pupils who are all *required* to take science). Perrodin (1966) does seem to have looked at comparable groups of 10, 12 and 14 year-old pupils, and found that attitudes to science became more positive between 10 and 12, and less positive between 12 and 14.

Mallison and Van Dragt (1952) assessed the scientific interests of 240 fifteen year-old pupils (using the Kuder Preference Record) and administered the same instrument again three years later. They concluded that the initial measures were dubious predictors of the final measures since scientific interests of 112 pupils increased, 13 pupils were unchanged, 115 pupils decreased, and only 13 of the 29 pupils who ranked science as their first interest in the first administration still ranked it first three years later. Stoops (1953), however, also using the Kuder Preference Record with a similar age group found correlations of 0.70 and 0.85 (girls and boys respectively) between scores on the Scientific Scale administered at the start and end of a two year period.

Wynn and Bledsoe (1967) in a more recent study have used the same instrument with 325 comprehensive high school students. Using t-tests they could find no significant differences between mean scores of pupils on entry to high school and scores on the same test administered two and three years later. They found that 68 pupils showed significant changes (at least one standard deviation) between pretest and post-test, and of these 30 showed significant gains and 38 significant losses. More than 75% of each of the four groups (junior boys, junior girls, senior boys, senior girls) showed no significant change in scientific interest. They conclude 'that the interests of high school students are more stable than has been generally believed'.

Evaluation of science curricula in the United Kingdom provides some evidence (Science 5/13, 1971, and Nuffield Secondary Science, 1972) that pupils' scores on attitude measures have fallen from pretest (before experience of the science course) to post-test (after undergoing the course). This has been explained in terms of negative reactions to the *test* rather than to the course, and has not been viewed as evidence of real change or instability of attitudes. Such arguments are tenuous.

Most of these studies are concerned with pupils' interest in science, but even in this area there is no substantial evidence to suggest an hypothesis regarding the stability, or otherwise, of pupils' scientific interests over long periods of time. Further, the value of attitude measures taken early in the pupil's school career as predictors of attitudes held later on has not been established.

Summary

The empirical evidence relating to the various research questions, pertaining to the assessments of the attitudes to science of this group of pupils, has been reviewed. Only a few of the relationships between attitudes to science and other social and personal variables can be taken as 'established' with any degree of confidence. For the purposes of this study only one positive relationship can be hypothesized *i.e.* that boys' level of *Interest and enjoyment in science* will be above that of girls'.

For the rest, it is hypothesized that none of the following variables will provide substantial prediction of pupils' attitudes to science along any of the five dimensions under consideration: sex (except for hypothesis above), measures of general intelligence, measures of divergency, science achievement measures, social class, size of school, location of school, denomination of school, teachers' attitudes, size of science class, ability grouping of class, number of periods allocated to science, teachers' ratings of pupils, pupils' pretest attitude measures.

In addition there is no evidence to suggest that pupils' attitudes to science will become more or less favourable during the first two years of secondary school.

CHAPTER 8

RESEARCH STRUCTURE AND STRATEGY

Chapters 4, 5 and 7 provided a number of substantive hypotheses relating pupils' achievement of affective objectives to other variables describing characteristics of pupils, schools, science courses and teachers. This chapter is concerned with the development of a scheme for operation of these variables, through which statistical hypotheses, directly related to the substantive hypotheses, may be tested.

In general, research design attempts to control variance, and, in particular, to maximize the variance of the variables of the substantive hypotheses. In this case the independent variables (sex, IQ, size of school *etc*) which the hypotheses relate to pupils' affective achievements, are not manipulable and so cannot be controlled. The study therefore has the inherent weakness of what Kerlinger (1969) calls

> 'Ex post facto research [which] may be defined as *that research in which the independent variable or variables have already occurred and in which the researcher starts with the observation of a dependent variable or variables. He then studies the independent variables in retrospect for their possible relations to, and effects on, the dependent variable or variables'* (page 360).

As well as the problem of lack of control of independent variables, a second weakness of ex post facto research is the researcher's inability to assign 'treatments' to groups at random (*e.g.* a randomly selected group of pupils cannot be sent a large, Roman Catholic school in a city following Integrated Science in S1, we have to make do with the pupils who are already there). This means that, strictly speaking, it will not be possible to make statements about affective achievement and its relationships with other variables for the *general population* on the basis of the research sample. However, it will be possible to select pupils at random from within the groups to which they have already 'assigned themselves'.

The third weakness of this type of research lies in the increased possibilities of invalid interpretation of results. Without random assignment of subjects to 'treatments' it is possible that there may be some other variable that correlates with the 'treatment' and which is the 'true' source of some observed 'treatment' effect. For example suppose that we find that 'medium' sized schools apparently have pupils with higher scores on attitude scales than other sized schools. The school size variable may be crucial, or it may be irrelevant. It may be the case that 'medium' schools have middle class catchment areas and pupils of above average ability. It is possible then that social class or intelligence may be the real determinent of attitudes. Only by randomization of pupils to treatments can unique interpretation of results be made with confidence.

The following four sections of this chapter are concerned with:

(i) A census of all secondary departments in Scotland to ascertain the types of science course followed in S1 and S2, the ability grouping of pupils, the number of periods allocated to S1 and S2 science, and the size and nature of the S1 intake.

(ii) The selection of criterion and independent variables for testing of hypotheses, and the sampling procedures adopted.

(iii) The procedures used to gather the data.

(iv) The preliminary studies of pupils' ability measures.

S1 and S2 Science Course Census

Before the final decisions on the research design could be taken, it was necessary to gather information on the S1 and S2 science courses being followed in Scottish secondary schools.

In March 1971 a questionnaire (see Appendix A-5A) was sent to all secondary departments in Education Authority and Grant Aided schools in Scotland requesting information on the type and organization of science courses for pupils in S1 and S2.

Each school was asked for details of the courses they expected to offer to their first year pupils in 1971-72 and to their second year pupils in 1972-73. They were asked to make a choice from the following alternative types of course:

1) Fully Integrated Science - each class following science in an integrated form taught by one teacher.
2) Integrated Science - each class following science in an integrated form under more than one teacher; each teacher responsible for part of the course only.
3) Physics, Chemistry and Biology taught as three separate subjects.
4) Physics/Chemistry taught as a combined subject. Biology taught separately.
5) Physics/Chemistry - no Biology taught at this stage.
6) Other type.

In addition to this schools were asked:

i) to estimate the number of pupils in S1 and S2,
ii) to estimate the number of periods of science to be allocated to S1

and S2,

iii) to state the average length of a lesson period,

iv) to indicate whether their S1 intake would be unselected, selected certificate, or selected non-certificate,

v) to indicate whether S1 and S2 science classes would be of mixed ability pupils or ability banded groups.

Five hundred and fifty-one replies to the questionnaire were received (100% response). Of these, 56 were from secondary departments which i) had closed, ii) expected to close within the next two years, or iii) were to be amalgamated with other secondary departments. Fifteen replies stated that their school offered no science at the S1/S2 stage and 480 secondary departments claimed to offer science to S1 and S2. The 15 schools offering no science in S1 and S2 included:

a) 'top-tier' schools offering classes S3 - S6,

b) schools with acute staff shortages,

c) schools lacking science accommodation,

d) special schools.

Type of science course followed

Table 8.1 shows the number of schools in each county or burgh claiming to follow fully Integrated Science, Integrated Science/Team Teaching and Separate Science Subjects courses. It can be seen that there was considerable variation between areas. The percentage of schools following the 'separate subjects' course throughout the two years varied from 60% (Clackmannan) to 0%. Separate figures for Education Authority schools, Grant Aided schools, Roman Catholic schools and Non-Denominational schools are also given.

TABLE 8.1b THE NUMBER OF SECONDARY DEPARTMENTS (L.E.A. AND G.A.) AND THE VARIOUS SCIENCE COURSES THEY EXPECTED TO FOLLOW WITH S1 AND S2 IN THE PERIOD 1971 - 1973 ARE SHOWN

Science Course / Burgh	Fully integrated science course	Integrated science Team Teaching	Separate Science Subjects	Other	Total
Aberdeen	8	0	3	4	15
Dundee	6	1	3	3	13
Edinburgh	20	1	6	4	31
Glasgow	32	7	9	17	65
Total	66	9	21	28	124
Total (counties & burghs)	297	33	57	93	480

TABLE 8.1c

Science Course / School Type	Fully integrated science course	Integrated science Team Teaching	Separate Science Subjects	Other	Total
Educ. Auth.	291	32	44	85	452
Grant Aided	6	1	13	8	28
All Schools	297	33	57	93	480
R.C. Schools	50	4	9	15	78
Non. Den. Schools	247	29	48	78	402
All Schools	297	33	57	93	480

TABLE 8.1a S1 AND S2 SCIENCE COURSES 1971 - 1973

County \ Science Course	Fully integrated science course	Integrated science Team Teaching	Separate Science Subjects	Other	Total
Aberdeenshire	18	3	0	4	25
Angus	5	0	1	3	9
Argyll	8	0	0	1	9
Ayrshire	16	3	0	9	28
Banff	4	0	2	0	6
Berwickshire	2	0	1	0	3
Bute	3	0	0	0	3
Caithness	1	0	1	0	2
Clackmannan	2	0	3	0	5
Dumfriesshire	7	2	1	2	12
Dunbartonshire	11	1	2	5	19
East Lothian	1	1	2	1	5
Fife	12	1	2	7	22
Inverness	16	2	0	3	21
Kincardine	3	0	0	0	3
Kirkcudbright	5	1	0	0	6
Lanarkshire	33	3	3	10	49
Midlothian	3	0	5	2	10
Moray & Nairn	3	2	0	1	6
Orkney	3	1	0	3	7
Peebles	1	0	0	0	1
Perthshire	10	0	4	3	17
Renfrewshire	25	0	1	0	26
Ross & Cromarty	12	1	1	0	14
Roxburghshire	2	0	1	0	3
Selkirkshire	1	0	0	1	2
Stirlingshire	4	2	5	5	16
Sutherland	3	0	1	2	6
West Lothian	7	1	0	2	10
Wigtownshire	2	0	0	1	3
Zetland	8	0	0	0	8
Total	231	24	36	65	356

Table 8.2 shows the percentages of the different types of schools following the various types of course. A striking difference between the situation in Grant Aided schools and that in Education Authority schools is shown. In the former only 21% claimed to follow a fully integrated course for the first two years while nearly half followed a completely separate subject course. In all the other groupings of schools considered here nearly two thirds followed a fully integrated course while 12% or less followed a separate subjects course.

Table 8.3 provides some information on schools following 'other' courses. The number of schools following courses labelled 1, 2 and 3, in the table was considerably higher than those labelled 5 and 6. One, 2 and 3 could be identified as having a strong 'integrated' factor and 5 and 6 a strong 'separate subject' factor. Some schools were included in more than one of the categories 1 to 8. For example, there were 6 schools that followed a fully integrated course in the first year and a separate subject course in the second year; 4 schools split their S1 class into two groups, one group following an integrated course for two years while the other followed a separate subject course; 8 schools followed a common Integrated Science course in the first year and split their class into an Integrated Science group and a separate subject science group for the second year.

Table 8.4 shows the estimated number of pupils who were expected to follow the various courses starting in 1971.

Time allocation for science

Table 8.5 shows the number of periods (equivalent 35-40 minutes) of science

TABLE 8.2 % OF SCHOOLS OF DIFFERENT TYPES FOLLOWING THE VARIOUS SCIENCE COURSES

Science Course / Types of Schools	Fully Integrated Science Course	Integrated Science Team Teaching	Separate Science Subjects	Other	Total
Educ. Auth.	64%	7%	10%	19%	100%
Grant Aided	21%	4%	46%	29%	100%
R.C. Schools	64%	5%	12%	19%	100%
Non. Den. Schools	62%	7%	12%	19%	100%
All Schools	62%	7%	12%	19%	100%

TABLE 8.3 NUMBERS OF SCHOOLS FOLLOWING "OTHER" COURSES

	"Other" Course	Number of Schools
1.	Fully Integrated Science for S1 only	37
2.	Part of class follows fully Integrated Science for S1 and S2	27
3.	Whole class follows "other" type of Integrated Science Course	5
4.	Whole class has Physics/Chem. integrated, Biology separate	8
5.	Part of the class follows separate subject science	7
6.	Separate subject science for S2 only	14
7.	Physics/Chem only, no Biology	14
8.	Biology or Rural Science only	1

TABLE 8.4 NUMBER OF PUPILS EXPECTED TO FOLLOW VARIOUS TYPES OF TWO YEAR SCIENCE COURSES FROM 1971 TO 1973

	Number of pupils	%
Number of pupils following fully integrated science	48,000	(58%)
Number of pupils following team teaching integrated science	6,000	(7%)
Number of pupils following separate subject science	9,000	(11%)
Number of pupils following other science courses	20,000	(24%)
Total number of pupils in science course intake	83,000	

TABLE 8.5 NUMBER OF PERIODS DEPARTMENTS EXPECT TO BE ALLOCATED TO SCIENCE IN S1 - S2 1971 - 73

	Schools with less than 4 periods	Schools with 4 periods	Schools with 5 periods	Schools with 6 or more periods	No information	Total schools
Educ. Auth.	11	146	189	79	27	452
Grant Aided	0	5	7	14	2	28
Roman Catholic	4	36	25	7	6	78
Non-Denom	7	115	171	86	23	402
All Schools	11	151	196	93	29	480

that schools expected to be allocated to S1 and S2. The figures given are an average over the two years. The Roman Catholic schools had the largest percentage of schools with less than 4 periods per week and the smallest percentage of schools with 6 or more periods per week. On these counts, the Grant Aided schools were the most favourably placed. Some ambiguity in the wording of this question resulted in 29 schools providing inadequate information for the purposes of this table.

Type of intake

The type of S1 intake in the different groups of schools is shown in Table 8.6. This table did, perhaps, oversimplify the real situation. Several secondary schools took an unselected intake from some of their feeder primary schools, but a selected intake from others. This was often the case for schools serving rural areas. A number of schools had an unselected intake at S1 but selection took place at the end of the first or second year when pupils might transfer to another school. Some schools had a mainly unselected intake but the very high ability pupils might be creamed off and sent, at the Education Authority's expense, to fee paying schools. There are some areas in which a significant percentage of the high ability pupils were able, and chose, to attend Grant Aided schools. This table should not be used as an indication of the number of comprehensive, senior secondary, and junior secondary schools at that time.

Ability grouping

The ability grouping of science classes is shown in Table 8.7. Seventy-six per cent of schools had mixed ability classes in S1. This number falls to

TABLE 8.6

TYPE OF SI INTAKE SHOWING NUMBER OF SCHOOLS IN EACH CATEGORY

	Unselected	Selected Certificate	Selected Non-Certificate	Unclassified	Total
Educ. Auth.	338	32	62	20	452
Grant Aided	10	18	0	0	28
Roman Catholic	56	9	11	2	78
Non-Denom.	292	41	51	18	402
All Schools	348	50	62	20	480

TABLE 8.7

ABILITY GROUPING IN SI AND SII SCIENCE

	No. of schools mixed ability classes SI and SII	No. of schools mixed ability classes SI only	No. of schools with no mixed ability classes	Total schools
Educ. Auth.	223	125	104	452
Grant Aided	11	7	10	28
Roman Catholic	32	26	20	78
Non-Denom.	202	106	94	402
All Schools	234	132	114	480

just under 50% in S2. Only 24% of the schools used ability banding or streaming throughout the two years, but among Grant Aided schools this percentage was higher - 39%.

Selection of Variables

Criterion variables

The criterion measures used for the major part of this study were scores on the five attitude scales developed in the pilot study as described in Chapter 5 for the purpose of assessing the five attitude objectives of Curriculum Paper 7. However, these five scales are concerned with various attitudes *towards* science rather than the 'scientific attitudes' supposedly characteristic of scientists. The fifth attitude objective 'an objectivity in observation and in assessing observations' probably reflects one such 'scientific attitude' and a Likert scale may be an inadequate way of assessing the extent to which a pupil is adopting such behaviour.

Observation of 'objectivity' in this study (3000 pupils) was considered to be impractical, but it was felt that some effort should be made to view pupils' behaviour from a second perspective. It has been suggested that one of the distinctive characteristics of the 'new' science curricula is the influence that they exert on the ways in which pupils attend to cognitive material (Heath 1964). These differences in modes of attention have been called 'cognitive preferences' and several tests purporting to assess such preferences have been developed. It has been suggested that such tests enable characteristics of the pupils' *cognitive style* to be identified. Cognitive style is related to the ways in which the pupil performs cognitive tasks, his typical patterns of thinking and problem solving, his preferred and characteristic behaviours.[1] It was decided to investigate the cognitive preference tests, and if possible use them to identify differences in scientific attitudes between pupils following a 'new' (Integrated Science) curriculum and those following a 'traditional (separate sciences) curriculum.

Since the literature on cognitive preferences is limited, detailed study of the nature and analysis of cognitive preferences was carried out. This will be found in Chapter 10.

Pupil variables

The social class variables were related to parental occupation classified according to the Registrar General's 'Classification of Occupations 1970'.

Each pupil's social class was translated into two dichotomous variables:

a) parental occupation in social classes 1 and 2, or not

b) parental occupation in social classes 4 and 5, or not.

For measures of pupils' ability it was decided to provide a battery of tests that would cover the various aspects (convergent and divergent) of 'intelligence' as adequately as possible. In this way it was hoped to identify those aspects which best predict pupils' attitudes.

---

[1] Making and assessing observations are cognitive tasks, objectivity is a characteristic of the way in which some pupils perform such tasks. It appeared that there might be some relationship between cognitive style (and so presumably cognitive preferences) and attitudes such as objectivity.

tasks, his typical patterns of thinking and problem solving, his preferred and characteristic behaviours.[1] It was decided to investigate the cognitive preference tests, and if possible use them to identify differences in scientific attitudes between pupils following a 'new' (Integrated Science) curriculum and those following a 'traditional (separate sciences) curriculum.

Since the literature on cognitive preferences is limited, detailed study of the nature and analysis of cognitive preferences was carried out. This will be found in Chapter 10.

Pupil variables

The social class variables were related to parental occupation classified according to the Registrar General's 'Classification of Occupations 1970'.

Each pupil's social class was translated into two dichotomous variables:

a) parental occupation in social classes 1 and 2, or not

b) parental occupation in social classes 4 and 5, or not.

For measures of pupils' ability it was decided to provide a battery of tests that would cover the various aspects (convergent and divergent) of 'intelligence' as adequately as possible. In this way it was hoped to identify those aspects which best predict pupils' attitudes.

---

[1]Making and assessing observations are cognitive tasks, objectivity is a characteristic of the way in which some pupils perform such tasks. It appeared that there might be some relationship between cognitive style (and so presumably cognitive preferences) and attitudes such as objectivity.

It might have been possible to collect VRQ measures of pupils from primary school records. However, there was some variation among counties in a) the administration of the tests (some counties no longer tested pupils), and b) the precise form of test used. It was decided to use AH4, a well tried test appropriate for this age group but little used in Scotland. It would be administered to all pupils in the study on the assumption that none would have encountered it before.

AH4 is a group test of general intelligence and while

> 'Many current intelligence tests consist of problems involving only one 'bias' and only one type of principle, the aim in AH4 is to incorporate as many different biases and principles as is consistent with a reasonably short test' . (Heim 1970a).

Part I of the test consists of 65 questions which have a *verbal* or *numerical* bias. Part II consists of 65 questions that have a *diagrammatic* bias. The emphasis is on deductive reasoning, the answer to each question being a single digit, unique number. In other words this is a test covering several aspects of ability by assessing convergent behaviour.

Butcher (1968) suggests that it may be

> 'instructive to treat the distinction between divergent and convergent thinking not as a dichotomy but as a continuum and to construct tests accordingly' (page 109).

He also warns against using only tests of divergency that relate closely to measures of verbal ability.

With such advice in mind three measures of divergency were selected for the study (see Appendix A for instruction sheets):

1) 'Circles', a test developed by Torrance (1962), for which the pupil is presented with a large number of small circles printed on a sheet of paper together with instructions to sketch as many objects as possible for which the circle is the main element. The pupil may add lines inside and/or outside the circle, the task being open-ended apart from the required inclusion of the circle in the sketch. This test uses non-verbal stimuli and responses, and Butcher (1968) reports an unpublished factor analysis of data from the study by Hasan and Butcher (1966) which revealed that the 'Circles' test

> 'has a uniquely high loading on a factor orthogonal to that of verbal fluency and interpretable as divergent thinking applied to graphic or diagrammatic material' (page 109).

2) 'Uses for Things', is a test first developed by Guilford as 'Unusual Uses for a Brick'. Here the pupil is instructed to think of as many uses as he can for a brick, an elastic band, a blanket *etc.* This test is also open-ended, except for the focus object. It is, however, verbal in both stimulus and response.

3) 'Meanings of Words', a test modified by Hudson (1966) from Guildford's original idea, requires the pupils to write down as many meanings as he can think of for various common words with multiple meanings (*e.g.* 'post', 'state', 'bar'). This is not an open-ended test insofar as there are a finite number of meanings that are acceptable, and it is clearly verbal in nature. There is evidence (Christie, 1969) from its loadings in factor analytic studies that it is more closely related to tests of verbal intelligence than to other tests of divergent abilities.

As well as subject-independent measures of ability, a measure of science achievement was required. It was decided to investigate the relationships between attitudes and pupils' *school science achievement marks* rather than marks on a special test administered for the study. The argument for this procedure was that the pupil's attitude, insofar as it is influenced by achievement at all, will be related to the pupil's *own perception* of his achievement. This perception will be influenced by the grades he is given in school. In addition, on the assumption that the pupil's attitudes may be related to his perceptions of his teacher's expectations of him, it was decided to collect ratings made by teachers on a) pupils' levels of interest, and b) pupils' academic ability. This implies the further assumption that teachers transmit their expectations of a pupil to that pupil.

Teacher variables

The attempt to isolate relevant and salient characteristics of teachers' attitudes or behaviour, which might influence pupils' attitudes to science, involved the modification of a 'Perceptions of Effective Science Teaching' scale for science teachers (Taylor, Christie and Platts 1970). The rationale for choice of this instrument, a description of its theoretical basis, and a discussion of the results obtained from its administration to science teachers in this study is given in Chapter 9.

School variables and sampling procedure

Information on some school variables, identified as being of interest in the research questions, was obtained from the secondary school census (*i.e.* size of school, location, denomination and type of science course followed). Other measures, such as socio-economic status of catchment area and general ability level of pupil intake, were not available. It was decided to determine measures of the latter group of variables after the sample had been drawn. The socio-economic status of the school's catchment area would be assessed on two variables:

a) the proportion of pupils from that school taking part in the study, whose parent's occupation was in class 1 or 2 of the Registrar General's Classification of Occupations (1970),
b) the proportion of pupils from that school taking part in the study, whose parent's occupation was in class 4 or 5 of the Registrar General's Classification of Occupations (1970).

The mean level of ability of the school's intake would be measured on two sets of variables:

a) the mean scores for boys and girls separately in each school on the AH4 test of general intelligence;
b) the mean scores for boys and girls separately on measures of divergency (fluency only) on 'Circles', 'Uses' and 'Meanings of Words'.

For the variables of size, denomination, location and type of science course followed, it was decided to adopt a method of sampling which would provide, directly, the categories by which the results might be analysed. A random stratified sample of schools was therefore taken, the basis of the stratification being existing classes of schools. Butcher(1965a) suggests that

> 'it is rarely worthwhile to use more than two, or at most three, bases of stratification for both theoretical and practical reasons. The theoretical reason,...is that the

> bases of classification will rarely themselves be independent, with the result that the degree of correlation or association between the basis of stratification and the variable under study will increase relatively slightly as new bases of classification are added. The practical reason is that, as the design becomes progressively more complicated, it will be increasingly difficult to ensure the right proportion of the sample in each cell of the cross-classification' (page 9).

Table 8.8 provides details of the number of schools of various types, the number of pupils entering S1 in these schools, and the different science courses offered to S1 and S2, as indicated by responses to the census forms. The horizontal classification conflates the denomination and location variables. The distinction between Roman Catholic (n = 78) and non-denominational schools (n = 402) was clear. It was also suggested by teachers and HMI's that there were general distinctions to be made between schools within the city burghs and those outwith the cities. However, there are some schools outwith the city boundaries but so close to the cities (particularly Glasgow) that they are, for all practical purposes, city schools. If these boundary schools are a *small* proportion of the non-city schools they are unlikely to influence the results substantially. This is the case for the non-denominational schools but not for the Roman Catholic. Roman Catholic schools are concentrated in the west of Scotland and are predominantly situated either within the Glasgow city boundaries or in the counties of Dunbartonshire, Lanarkshire and Renfrewshire but within the Glasgow conurbation.

It was therefore decided to sample schools for analysis of variance on a denominational/location dimension employing three categories of population

a) all Roman Catholic schools;

b) all non-denominational schools within the four city burghs of Aberdeen, Dundee, Edinburgh and Glasgow;

c) all non-denominational schools outwith the city burghs.

TABLE 8.8

NUMBER OF SCHOOLS IN VARIOUS SIZE, DENOMINATION, AND TYPE OF S1 SCIENCE COURSE CATEGORIES (DATA COLLECTED MARCH 1971)

		Roman Catholic	Non-Denominational (Cities)	Non-Denominational (Outwith the Cities)	Totals by size of school
SMALL	No. of schools	27	31	157	215
	No. of schools following I.S.	21	15	120	156
	No. of schools following S.S.	2	8*	8	18
	No. of schools following 'other' courses	4	8	29	41
	No. of S1 pupils (nearest 1000)	2	3	10	15
MEDIUM	No. of schools	35	52	112	199
	No. of schools following I.S.	21	27	53	101
	No. of schools following S.S.	5	7	20	32
	No. of schools following 'other' courses	9	18	39	66
	No. of S1 pupils (nearest 1000)	8	11	26	45
LARGE	No. of schools	16	16	34	66
	No. of schools following I.S.	8	13	19	40
	No. of schools following S.S.	2	1	4	7
	No. of schools following 'other' courses	6	2	11	19
	No. of S1 pupils (nearest 1000)	7	6	13	26
ALL SIZES					GRAND TOTAL
	No. of schools	78	99	303	480
	No. of schools following I.S.	50	55	192	297
	No. of schools following S.S.	9	16	32	57
	No. of schools following 'other' courses	19	28	79	126
	No. of S1 pupils (nearest 1000)	17	20	49	86

*including 7 grant aided

The vertical classification by size of school indicates the *number of pupils in the S1 intake:*

a) small - S1 intake less than 150,

b) medium - S1 intake between 150 and 300,

c) large - S1 intake greater than 300.

The basis on which to classify size was problematic. Is it the total number of pupils in the school, or the total number of pupils in his own year, that is likely to influence the pupil's attitudes? If the former, then does it matter that the school roll covers primary and secondary departments, or just six years of secondary, or four years of secondary, or only the first two years of secondary? Can the influence of 800 other pupils of various ages from 5 to 18 be considered equivalent to 800 in the pupil's own and one adjacent year?

In the absence of any theory, a relatively arbitrary choice had to be made on the basis by which size of school would be measured. It was decided to use the size of S1 intake since this was the only aspect for which all the schools were strictly comparable.

A third classification was the type of science course followed - Integrated Science or separate science subjects. Inspection of Table 8.8 showed that 'medium sized schools' was the only size stratum for which all categories of schools had a minimum of 10% following separate science subjects. It was decided that the comparison of attitudes between groups of pupils following the two types of science course would only be feasible for medium sized schools.

To summarize, a sample of 36 schools was made, 3 schools being selected at random from each of the following groups:

i) Roman Catholic, small sized schools following an Integrated Science course in S1 and S2.

ii) Roman Catholic, medium sized schools, following an Integrated Science course in S1 and S2.

iii) Roman Catholic, medium sized schools, following three separate science subject courses in S1 and S2.

iv) Roman Catholic, large sized schools, following an Integrated Science course in S1 and S2.

v) Non-denominational schools in the cities, small sized, following an Integrated Science course in S1 and S2.

vi) Non-denominational schools in the cities, medium sized, following an Integrated Science course in S1 and S2.

vii) Non-denominational schools in the cities, medium sized, following three separate science subject courses in S1 and S2.

viii) Non-denominational schools in the cities, large sized, following an Integrated Science course in S1 and S2.

ix) Non-denominational schools, outwith the cities, small sized, following an Integrated Science course in S1 and S2.

x) Non-denominational schools, outwith the cities, medium sized, following an Integrated Science course in S1 and S2.

xi) Non-denominational schools, outwith the cities, medium sized, following three separate science subject courses in S1 and S2.

xii) Non-denominational schools, outwith the cities, large sized following an Integrated Science course in S1 and S2.

Within these schools individual pupils were sampled on a random or systematic (unrelated to ability) basis from among all pupils entering S1 in August 1971 with the exception of those pupils whose reading level was too low for them to understand the test items. The decisions on non-readers to be excluded were taken by the teachers in the schools. Approximately 90 pupils were sampled from each school. In some cases the schools formed mixed ability classes on a similar sampling basis. Since these groups had had no secondary school science, the pupils forming them were accepted as an appropriate sample. For several of the small sized schools the total S1 intake was less than 90. Since *mean scores for groups* were to be used in a major part of the study (analysis of variance, see Chapter 11) the reduced numbers for small schools were immaterial. However, for the factor analytic study of the dimensions of pupils' attitudes (see Chapter 6) it was considered that roughly equal numbers of pupils from the various types of school should be included. To this end, four extra schools were sampled: one small Roman Catholic school, one small non-denominational city school, and two small non-denominational schools outwith the cities. The full 40 schools were used in some of the multiple regression analyses to be described in Chapter 11.

The population of schools from which sub-populations (i) to (xii) are taken, is all Scottish secondary schools offering S1 and S2 science with the exception of the Independent schools. In 1971 an estimated 98.5% of all S1 pupils in Scotland attended either LEA or Grant Aided schools.

Of the original 'without replacement' stratified random sample of schools, four schools later had to be replaced by further random sampling within appropriate strata. Two of these were small, Roman Catholic schools and two were small, non-denominational, outwith the cities schools. The reasons for removal from the sample were that two of the schools were due to close during the period of the study, one school was due to move *twice* during the study, and the fourth school had completed the census questionnaire incorrectly and in fact taught no biology at all in S1 and S2.

Table 14A Appendix A provides the list of schools that participated in the study.

Collection of Data

Each of the 40 schools in the study was asked to appoint a co-ordinator who would make all arrangements within the school. A visit of at least half a day was made to each of the schools and co-ordinators in order to explain the purpose and plan of the research. (Appendix A-14A gives a list of school co-ordinators.)

The schools undertook to administer the 'Attitudes to Science Questionnaire' to the sample of pupils within the first few days of their entry to secondary school in August 1971, *before they had experienced any secondary school science.* The instructions for administration of the questionnaire are given in Appendix A-14A. Pupils' scores on the five scales as measured in 1971 will be referred to as the 'pretest attitude scores'. The same questionnaire was administered to the same sample of pupils in the last term of their second year in secondary school *i.e.* June 1973. Scores obtained at the second administration are referred to as 'post-test attitude scores'.

During the first half of the first term of the 1971-72 academic year a test of general intelligence (AH4), three tests of divergency ('Circles', 'Uses' and 'Meanings of Words') and a 'pupil questionnaire' were administered to the sample of pupils. The 'pupil questionnaire' was primarily concerned with identification of social class. The various tests, questionnaires and instructions for administration are given in Appendix A-3A to 13A.

The testing in two schools was carried out by the researcher in order to identify any problems that might have been overlooked in the pilot study.

The co-ordinator was asked to submit information about the science teaching groups for the sample of pupils throughout the two years. This information consisted of:

(i) size of the teaching groups,

(ii) ability grouping,

(iii) sex grouping,

(iv) number of periods of science per week.

At the end of S1 and S2 the teachers were asked to rate those pupils from the sample who belonged to their teaching groups for

a) interest in science,

b) academic ability.

The ratings were to be on a five point scale, they were to relate to the full academic year and the teachers were free to use their own criteria for their assessment of pupils. In addition to ratings, teachers were asked to provide any achievement marks that were given to pupils for performance on science tests or examinations. (Instructions to co-ordinators and teachers, and details of these measures will be found in Appendix A-16A to 17A.

The study of cognitive preferences was carried out in six schools, one from each of the 'medium sized' school categories. All the testing was carried out by the researcher in June 1973 and will be described in detail in Chapter 10.

Responses on the pretest measure of attitudes were obtained from 3612 pupils. Table 1C Appendix C shows how pupils were distributed among the various types of school and social classes (these figures were taken after pupils had completed the AH4 test, divergency tests, and pupil questionnaire, and since some pupils missed one or more of these the total number is reduced to 3027). The final analysis was carried out on the scores of 2815 pupils who had completed all the tests in the study.

Preliminary Studies of Pupils' Ability Measures

The distribution of intelligence test scores among the sample

An exploratory study of the AH4 scores of the pupils in this sample was carried out in order:

a) to establish norms for this population of pupils (no previous data were available on AH4 scores of pupils of this age group in Scotland),

b) to check that there was nothing obviously distorted or atypical about the levels and distribution of intelligence among this sample of pupils from secondary schools which, although mostly comprehensive, contained examples of schools with selected intakes.

Table 8.9 provides a comparison of the distribution and mean AH4 score of the sample with the norms set up for the same age group in England (Heim 1970a). Figure 8.1 shows the frequency distribution curve for the Scottish pupils for which there is no evidence of skewness or other distortion from the normal curve.

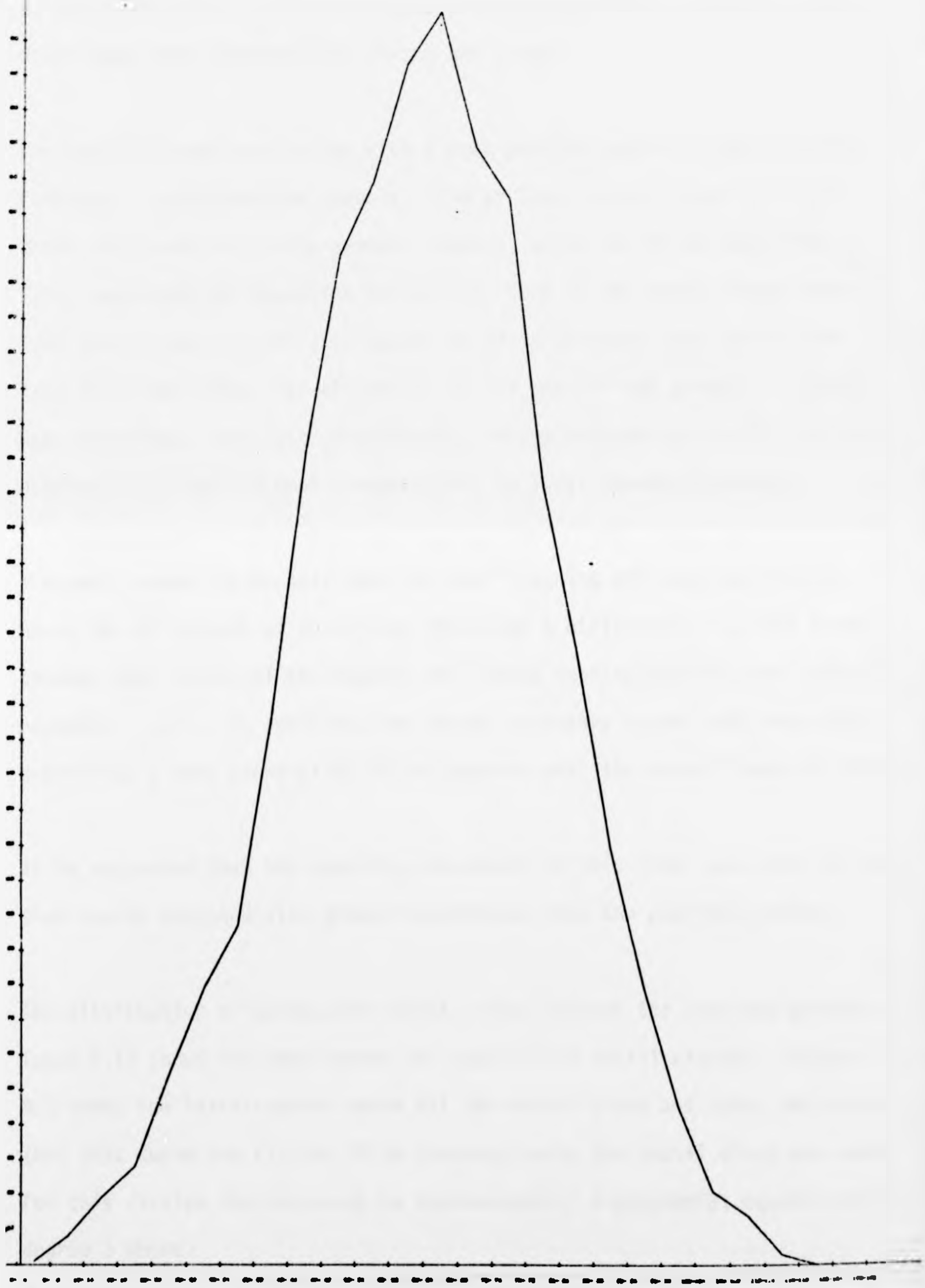
The Scots'mean score is 5 marks higher on AH4 total score than that of the English comprehensive school 12-year olds (a difference that is significant

TABLE 8.9

A COMPARISON OF THE TEST SCORES OF SCOTTISH SECONDARY SCHOOL CHILDREN (TWELVE YEARS) WITH A.H. 4 TEST NORMS OF ENGLISH COMPREHENSIVE SCHOOL CHILDREN (TWELVE YEARS)

	Total Score	
	Scottish Sample (N = 3273)	English Sample (N = 790)
Grade A (top 10%)	84 to 130	76 to 130
Grade B (next 20%)	70 to 83	65 to 75
Grade C (next 40%)	50 to 69	46 to 64
Grade D (next 20%)	36 to 49	31 to 45
Grade E (bottom 10%)	0 to 35	o to 30
	Mean Score = 59.40	Mean Score = 54.40
	S.D. = 18.29	S.D. = 17.11

FIGURE 8.1



Frequency distribution curve for 3273 twelve-year-old Scottish school children on AH4 total score.
Mean score = 59·397; standard deviation = 18·290.

at the 0.001 level). This is a substantial difference, but there is some doubt about the comparability of the two groups.

The English norms were set up with a much smaller number of pupils (790) from only 3 comprehensive schools. Two of these schools were in areas where there were also some grammar schools, while the third came from a fully comprehensive education authority. Some of the pupils tested were from second year classes (equivalent to S1 in Scotland) but others came from first and third year classes (*i.e.* 11+ and 13+ age groups).[1] There was, therefore, some lack of uniformity in age and some possibility of high scoring pupils having been 'creamed off' to local grammar schools.

The small number of schools used and the 'creaming off' may be crucial. Among the 40 schools of this study there was a difference of 37 AH4 marks between mean scores of the highest and lowest scoring schools (see Table 2C Appendix C). In addition, one senior secondary school with very able pupils had a mean score of 82.52 as compared with the overall mean of 59.40.

It is suggested that the sampling procedures of this study have led to norms that can be accepted with greater confidence than the published norms.

The distribution of scores with social class follows the expected pattern. Table 8.10 shows the mean scores for pupils in 5 social classes. Figure 8.2 shows the least-squares curve fit for social class and total AH4 score. (For this curve the fiction of an *interval scale* for social class was used. For this fiction the curve may be represented by a polynomial equation of degree 3 where:

---

[1] I am very grateful to Dr. Alice Heim for providing this detailed information in a private communication.

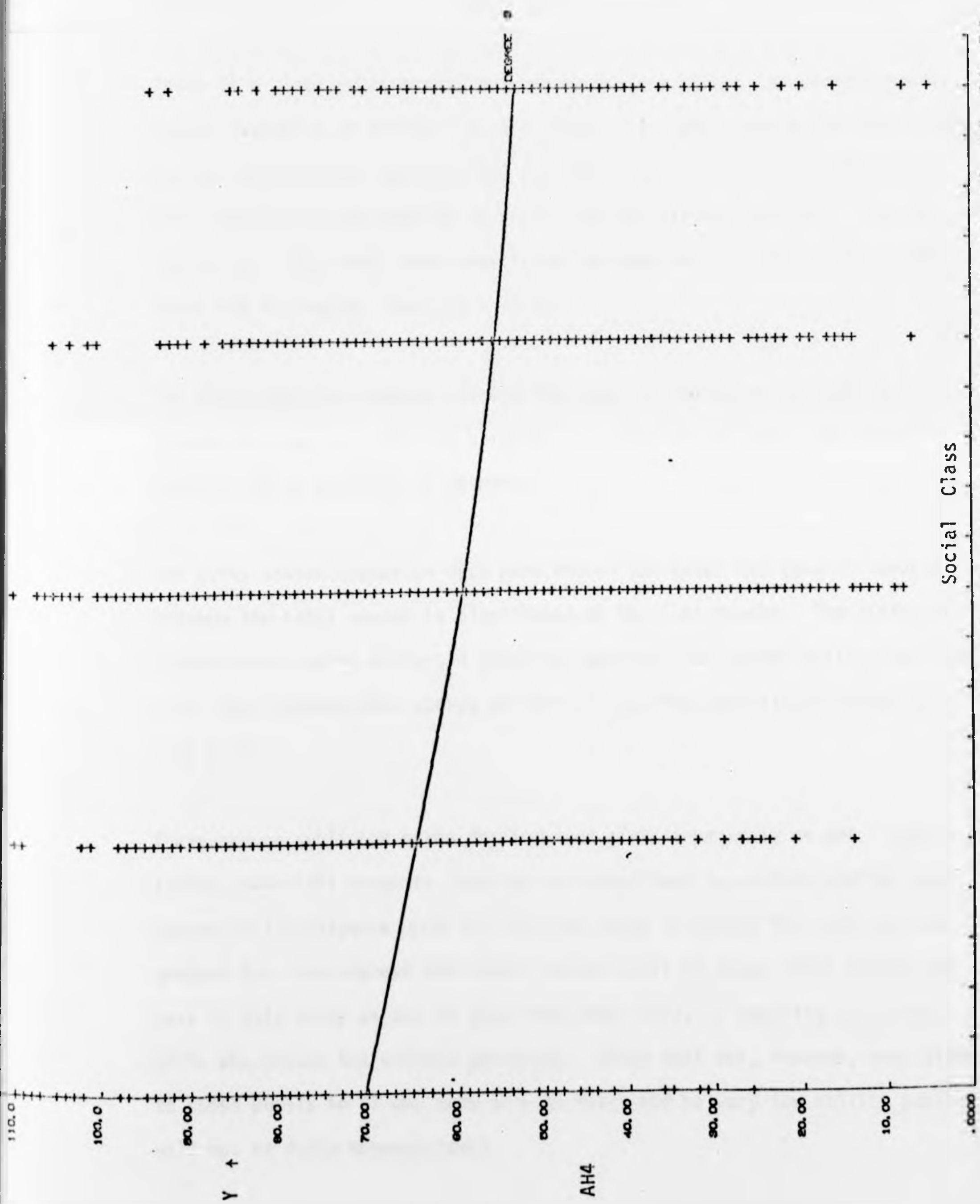
TABLE 8.10

AH4 MEAN SCORES AND STANDARD DEVIATIONS FOR SOCIAL CLASSES

Test Scores	Social Classes 1 (N=136)	2 (N=471)	3 (N=1860)	4 (N=454)	5 (N=106)
Part I					
Mean Score	31.30	28.06	25.27	23.87	21.81
S.D.	8.65	9.22	9.07	8.59	8.97
Part II					
Mean Score	38.90	36.99	34.11	32.36	30.84
S.D.	10.95	9.49	10.61	10.29	10.90
Total					
Mean Score	70.20	65.05	59.38	56.38	52.65
S.D.	17.68	16.80	17.96	17.18	17.51

FIGURE 8.2

LEAST SQUARES FIT FOR SOCIAL CLASS AND TOTAL AH4



AH4 score = $f(x) = 77.29 - 6.99x + 0.25x^2 + 0.04x^3$

where x = 1, 2, 3, 4 or 5 *i.e.* social class.)

There is a clear relationship between *mean class scores* and socio-economic status (assuming an interval social class scale, the correlation coefficient between social class and mean AH4 for that class = -0.99), but Figure 8.2 shows the very great overlap in scores between all the classes. This is reflected in the much lower coefficient between an *individual pupil's* AH4 score and his social class (r = -0.19).

The distribution of scores between the sexes is indicated by Table 8.11. Figures 8.3 and 8.4 show the unimodal distributions of boys' and girls' scores with no evidence of skewness.

The girls scored higher on each part and on the total AH4 (the difference between the total scores is significant at the 0.01 level). The difference between mean scores on Part I (testing numerical and verbal skills) was 1.49 while that between mean scores on Part II (testing spatial abilities) was only 0.55.

There are no published norms for boys and girls separately on AH4. Heim (1970b, page 136) suggests that the two sexes tend to achieve similar mean scores on intelligence tests but that the range of scores for boys will be greater *i.e.* the highest and lowest scorers will be boys. This is not the case in this study as can be seen from Table 8.12, if anything it is the girls who occupy the extreme positions. (This test was, however, only given to those pupils in S1 who were able to read, and so very low ability pupils will not be fully represented.)

TABLE 8.11

A COMPARISON OF THE DISTRIBUTION OF THE AH4 SCORES OF BOYS AND GIRLS

Total Scores

	Boys (N = 1605)	Girls (N = 1668)
Grade A	83 to 130	84 to 130
Grade B	69 to 82	71 to 83
Grade C	49 to 68	51 to 70
Grade D	35 to 48	36 to 50
Grade E	0 to 34	0 to 35
	Mean Score = 58.36	Mean Score = 60.40
	S.D. = 18.20	S.D. = 18.32

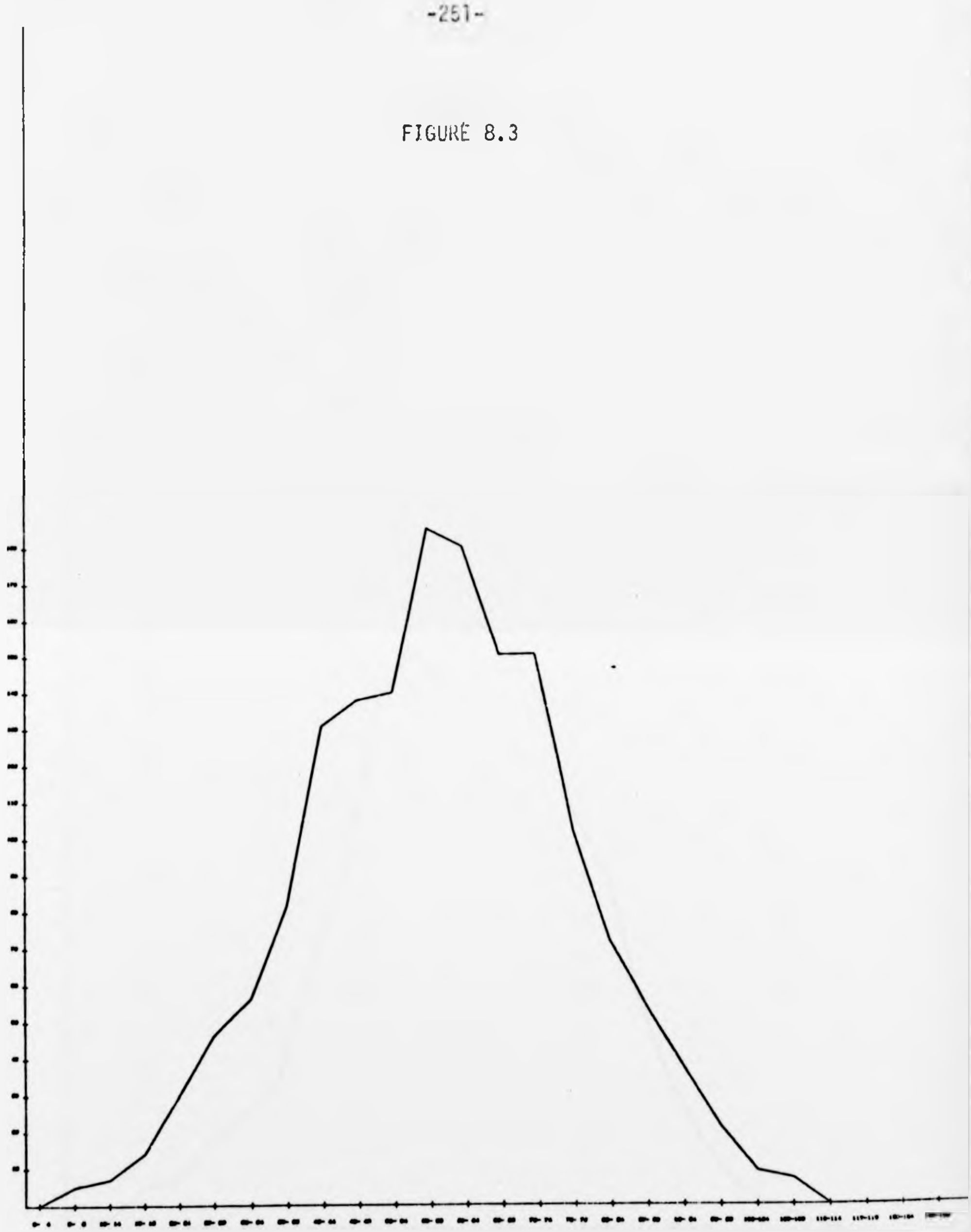
Part I

	Boys	Girls
	Mean Score = 24.57	Mean Score = 26.07
	S.D. = 9.31	S.D. = 9.09

Part II

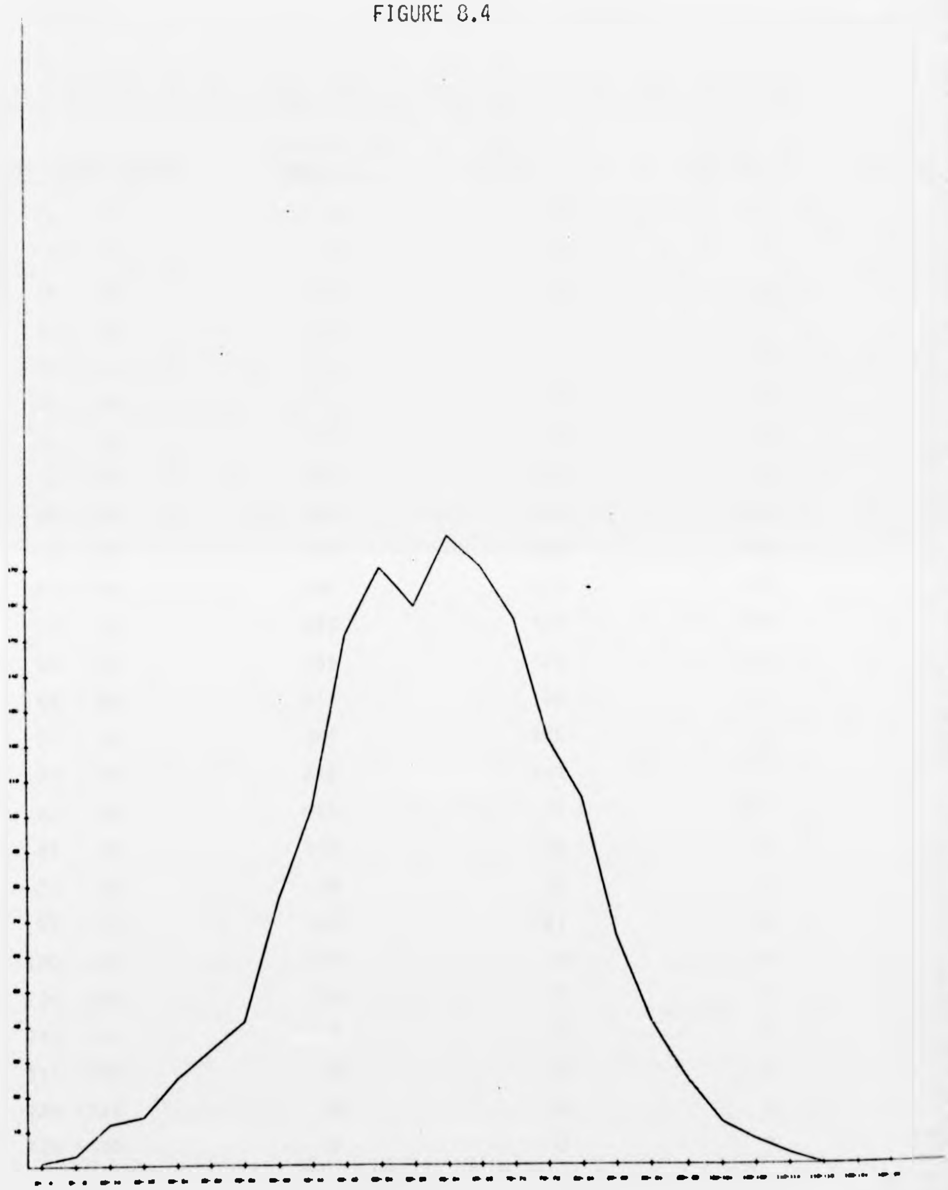
	Boys	Girls
	Mean Score = 33.78	Mean Score = 34.33
	S.D. = 10.84	S.D. = 10.63

FIGURE 8.3



Frequency distribution curve for 1605 twelve-year-old Scottish school boys on A H4 total score.
Mean score = 58·358; standard deviation = 18·202.

FIGURE 8.4



Frequency distribution curve for 1668 twelve-year-old Scottish school girls on AH4 total score.
Mean score = 60·397; standard deviation = 18·320

TABLE 8.12

DISTRIBUTION OF AH4 SCORES FOR ALL PUPILS, BOYS ONLY AND GIRLS ONLY

AH4 Total Score	Number of Pupils	Number of Boys	Number of Girls
0 - 4	1	0	1
5 - 9	8	5	3
10 - 14	19	7	12
15 - 19	28	14	14
20 - 24	55	30	25
25 - 29	79	46	33
30 - 34	97	56	41
35 - 39	157	81	76
40 - 44	232	130	102
45 - 49	288	137	151
50 - 54	309	139	170
55 - 59	343	184	159
60 - 64	358	179	179
65 - 69	319	149	170
70 - 74	304	149	155
75 - 79	222	101	121
80 - 84	175	71	104
85 - 89	118	53	65
90 - 94	78	37	41
95 - 99	45	21	24
100 -104	21	9	12
105 -109	14	7	7
110 -114	3	0	3
115 -119	0	0	0
120 -124	0	0	0
125 -130	0	0	0

The higher scores for girls on Part I and Part II support the general findings (reviewed by Gardner 1974a) that girls are superior to boys in verbal skills, but do not support the consistently found superiority of boys over girls in tests of spatial ability (in her review of work in this area Kelly, A., 1974 claims that sex differences in spatial ability are even evident in rats).

Since much of the analysis of results was to be of single sex groups in schools of various types an analysis of variance of AH4 mean scores was carried out. This was a 4-factor design (denomination of school, size of school or type of science course, school and sex), with *schools* nested within *denomination* x *size (or science course)* cells and crossed with *sex*. This analysis is described in detail in Chapter 11, pages 334 to 338.

The results of this analysis will be found in Tables 1D to 4D Appendix D. Tables 1D and 3D show that there are significant *school* effects and a significant *denomination* effect. In other words there are significant differences ($p<0.05$) among schools following Integrated Science in S1 in the mean levels of intelligence of their intakes, and furthermore there are significant differences ($p<0.05$) between the intakes to schools of differing denominations (the significant difference here is the low level for the Roman Catholic schools in comparison with the other two denominational categories). These differences are reflected in the differences between the mean scores of all individual pupils attending non-denominational schools and all those attending Roman Catholic schools (see Table 8.13). Figures 3C and 4C Appendix C show the frequency distributions of scores of these two groups. The lower AH4 scores of pupils attending Roman Catholic schools mirrors the social class distributions. Table 1C Appendix C shows that the proportion of the pupils from social-classes 1 and 2 in

TABLE 8.13

A COMPARISON OF TEST NORMS OF ROMAN CATHOLIC AND NON-DENOMINATIONAL POPULATIONS

	Non-Denominational (N 1 2269)	Roman Catholic (N 1 1004)
Grade A	85 to 130	78 to 130
Grade B	72 to 84	65 to 77
Grade C	53 to 71	47 to 64
Grade D	38 to 51	33 to 46
Grade E	0 to 37	0 to 32
	Mean Score = 61.05	Mean Score = 55.67
	S.D. = 18.42	S.D. = 17.44

Roman Catholic schools (approximately 10%) is less than half that for the pupils from non-denominational schools (approximately 25% and 23%).

There were no significant differences between group mean AH4 scores from schools of different *sizes*.

From this analysis it was concluded that the AH4 scores were suitable for use as predictors in various regression analyses to be carried out (see Chapter 11), but that the significant differences between *denominations* and *schools* identified in the analysis of variance made them unsuitable for use as covariates in an analysis of covariance to be carried out with attitude scores as the dependent variables (this will be argued in detail in Chapter 11).

Divergency measures: reliability study and exploration of inter-test relationships[1]

The three tests of divergency chosen for this study have been used in a large number of other research studies, but several different scoring methods have been employed and inter-marker reliabilities have shown considerable variation. It was considered necessary to carry out a pilot study in order:

a) to investigate different scoring methods and the relationships between scores so obtained;
b) to ascertain the inter-marker reliability for each scoring method;
c) to examine the general pattern of intercorrelations between scores of pupils on the various tests.

---

[1]I am most grateful to Andrew Comrie, an Honours student in the Education Department, University of Stirling, for the substantial help he provided in the preliminary study of divergency measures. As well as being one of the two main markers, he calculated all the inter-marker and inter-scale correlation coefficients for the pilot study.

This pilot work was carried out with 65 S1 pupils from one of the schools used in the pilot study (described in Chapter 5). The pupils completed the tests of 'Circles', 'Uses' and 'Meanings' in that order and before administration of AH4. The intention here was to avoid establishing a convergent mental set just before the administration of open-ended tasks, and it was assumed that in this sequence the pupil moved along a continuum from 'open' to 'closed' activity.

The various scoring methods that have been used for divergency tests - fluency, flexibility, originality, uniqueness, elaboration - tend to be time consuming, tedious and complex. The very large number of pupils in the main study necessitated a limitation on the number of ways in which each script was marked. It was decided to explore three methods only:

1) Fluency: the pupil scores one point for each separate idea indicated by his responses to the test.

2) Originality: each response is weighted in relation to the number of pupils who present the same idea, according to the following scale (Cropley and Clapson 1971)

Weighting	Proportion of scripts on which response appears	Corresponding no. of scripts in Pilot School
0	15% or more	10 or more
1	7% to 14%	5 to 9
2	3% to 6%	2 to 4
3	1% to 2%	1
4	below 1%	-

3) Uniqueness: the pupil scores one point for those responses for which he is the *only* pupil responding in that way; all other responses are non-scoring.

(Since 'Meanings of Words' is not an open-ended test it could only be marked for fluency.)

In the pilot study no 'rules' for scoring were laid down other than these general outlines. There was, for example, no criterion provided to identify a 'separate idea'. Each scorer developed personal criteria for judgment of different responses. Even this rather loose procedure led to satisfactory levels of inter-marker reliability (two markers) with correlation coefficients between 0.88 and 0.99 (see Table 8.14) and no significant differences between the mean values of marks awarded by different markers (see Table 8.15). These reliability measures compare not unfavourably with a very elaborate and ingenious system commented on by Goldman (1964), where coefficients from 0.84 to 1.00 were recorded for fluency, flexibility and originality scores.

On the basis of the experience of pilot study some guide-lines were laid down for the main study scoring. Two more markers assisted with the fluency scoring in the main study. Table 8.16 presents the inter-marker reliability coefficients between each of these markers and one of the original markers. For this reliability study 3 pupil scripts were picked at random for each of the three tests and from each of 39 out of the 40 sample schools. These 3 x 117 scripts were marked independently by the three markers and Pearson product moment inter-marker correlations were calculated for each test. The guide-lines provided for each marker consisted of

a) provisional scoring rules for 'Circles' (see Appendix A-18A);

TABLE 8.14 INTER-MARKER CORRELATIONS ON DIVERGENCY TEST SCORES (PILOT)

n = 65

	Tests of Divergency		
	Circles	Uses for Things	Meanings of Words
Fluency	0.96	0.99	0.99
Originality	0.95	0.93	-
Uniqueness	0.88	0.92	-

TABLE 8.15 t-VALUES INDICATING SIGNIFICANCE LEVELS OF DIFFERENCES BETWEEN THE MARKS AWARDED BY THE TWO MARKERS ON THE PILOT TESTS OF DIVERGENCY

	Degrees of freedom = 64		
	Uses for Things	Circles	Meanings of Words
Fluency	0.10	0.78	0.02
Originality	0.16	0.90	-
Uniqueness	0.00	0.43	-

None of these values reach the 5% significance level

TABLE 8.16

INTER-MARKER CORRELATIONS - DIVERGENCY TESTS RELIABILITY MAIN STUDY

'Circles' - fluency

Marker	P	H	S
P		0.96	0.94
H			0.97
S			

n = 117

'Uses' - fluency

Marker	P	H	S
P		0.88	0.95
H			0.91
S			

n = 117

'Meanings' - fluency

Marker	P	H	S
P		0.94	0.95
H			0.96
S			

n = 117

b) a list of about 250 responses to the 'Circles' test collected from the pilot study;

c) provisional scoring rules for 'Uses' (see Appendix A-19A);

d) lists for each of the focus objects in the 'Uses' test (*i.e.* a blanket, a brick, an elastic band, a barrel) of responses collected in the pilot study (between 40 and 80 for each object);

e) a list of acceptable and distinguishable dictionary meanings for each of the words used in the 'Meanings' test (*i.e.* bar, post, set, terms, form, box, lead, state).

The decision about which of the scoring methods should be retained for the main study rested on two requirements:

(i) that a measure should be retained if there was evidence to suggest that it had substantial variance that was *not* common with the other measures used,

(ii) that the amount of scoring be kept to a minimum.

Intercorrelations between pupils' scores on the different procedures within the pilot tests are shown in Table 8.17. Originality correlates highly with both fluency and uniqueness (minimum r = 0.80). Fluency and uniqueness, however, only share 28% and 38% variance on the 'Circles' and 'Uses' tests respectively. It was decided, therefore, that:

(i) originality, which involved the most time-consuming marking procedures and which correlated highly with the other two measures, would not be used for the main study;

(ii) fluency would be used, and measures of fluency could be compared between schools;

(iii) uniqueness would be used, but in view of the very different environments and experiences of pupils from schools in different

TABLE 8.17 CORRELATIONS BETWEEN PUPILS SCORES ON DIFFERENT SCORING PROCEDURES WITHIN TESTS (PILOT)

	Scoring Procedures under Comparison		
Tests	Fluency and Uniqueness	Fluency and Originality	Uniqueness and Originality
Circles	0.53	0.82	0.80
Uses for Things	0.61	0.81	0.81

TABLE 8.18 INTERCORRELATIONS BETWEEN SCORES ON DIFFERENT DIVERGENCY TESTS (PILOT STUDY)

n = 65

	Scoring Procedures		
Tests under Comparison	Fluency	Originality	Uniqueness
Circles v. Uses	0.28	0.21	0.24
Circles v Meanings	0.31	-	-
Uses v Meanings	0.46	-	-

parts of the country, comparisons on uniqueness scores could only be made within schools.

The relatively low proportion of shared variance between fluency and uniqueness measures was reflected in the 'within-schools' part of the main study (see Chapter 11). A combined r matrix showing intercorrelations between test scores in 12 schools is shown in Table 8.19. The common variance between 'Circles' fluency and uniqueness scores is 17%, and for 'Uses' it is 24%.

Tables 8.18 and 8.19 indicate the patterns of between-test correlations of pupils' scores on the pilot and the main studies. In each case the greater shared variance is between 'Uses' and 'Meanings' (21% and 20%), probably reflecting the common verbal ability required for the responses. The least shared variance on fluency measures is between 'Circles' and 'Meanings' (10% and 8%), reflecting the 'open/diagrammatic' nature of one and the 'closed/verbal' of the other. The 'Circles'/'Uses', fluency relationship increases substantially from pilot to main (8% to 16%), but this is not the case for the uniqueness relationship (6% to 5%). The main study results for this pair of tests suggest that while there may be some evidence to suggest that pupils who can produce plenty of ideas in an open-ended verbal response situation may be expected to do the same in a diagrammatic context, the ability to produce *unique or highly creative* ideas in one situation is unrelated to a similar ability in the other.

Table 8.19 also shows the divergency test correlations with AH4 for the main study. Relationships between AH4 and the uniqueness measures (r_c = 0.19, r_u = 0.18) are weak suggesting that performance on general intelligence tests is not indicative of the ability to produce unique ideas.

TABLE 8.19

INTER-CORRELATIONS BETWEEN TEST SCORES ABSTRACTED FROM 'COMBINED' r MATRIX OF WITHIN-SCHOOL ANALYSIS OF MAIN STUDY

Test Score	(1)	(2)	(3)	(4)	(5)	(6)
1. Circles fluency		0.40	0.29	0.41	0.28	0.30
2. Uses fluency			0.45	0.24	0.49	0.37
3. Meanings fluency				0.21	0.27	0.51
4. Circles uniqueness					0.23	0.19
5. Uses uniqueness						0.16
6. AH4						

Correlations of AH4 with 'Circles' and 'Uses' fluency are moderate (r_c = 0.30, r_u = 0.37), reflecting the diagrammatic and verbal components of AH4 respectively. The highest correlation coefficient in the matrix is that for the relationship between AH4 and 'Meanings' (r = 0.51), suggesting that 'Meanings' is more closely related to 'traditional' tests of general intelligence (and in particular to verbal components of such tests) than it is to other tests of divergency that have a more open-ended nature than its own.

The pilot study revealed a few problems (mainly administrative) that led to minor changes in two of the tests. 'Meanings of Words' and 'Uses for Things' were reduced in length and the time allowed for each was changed from 15 to 10 minutes. Since the time set for 'Circles' was already 10 minutes, teachers administering the three tests in the main study found that they were able to fit them, together with AH4, into one double period.

CHAPTER 9

THE ASSESSMENT OF THE TEACHERS' ATTITUDES

This chapter describes the attempt that has been made to differentiate among science teachers in this study on the basis of salient attitude factors that might be expected to influence the attitudes to science of their pupils. In Chapter 7 it was suggested that teacher attitudes suitable for this purpose would be related to the teacher's classroom behaviour.

Assessments of teachers' attitudes towards teaching or teachers' own conceptions of good teaching have been widely used in the study of teaching competence and in efforts to identify 'effective teachers'. The most popular instrument in this area has been the Minnesota Teacher Attitude Inventory (MTAI: Cook, Leeds and Callis, 1951). The construction and validation of this scale did not derive from a theoretical model but was empirically developed using the researchers' preconceived criteria of what constitutes 'good teaching'.

Other instruments, more relevant to education in Britain, such as the Manchester Scales of Opinions about Education (Oliver and Butcher, 1962), and the Role Definition Instrument (Finlayson and Cohen, 1967), are useful in the comparison of the attitudes of different groups of teachers or student teachers (Butcher, 1965b; Finlayson and Cohen, 1967; McIntyre and Morrison, 1967; Oliver and Butcher, 1968) but do not appear to carry clear implications for the teacher's classroom behaviour.

There is evidence (for example, McLean, Gowan and Gowan, 1955) using the Study of Values (Allport, Vernon and Lindzey, 1951) that teachers' ideas about teaching are strongly influenced by the subject which they are teaching. Getzels and Jackson (1963) point out that the use of other instruments such as the Kuder Preference Record (Kuder, 1953) and the Strong Vocational Interest Blank (Strong, 1943) would not be very profitable in studying 'good' teachers *versus* 'poor' teachers unless the teachers were subdivided into subject matter areas. They point to the dangers of assuming a unitary nature for teachers' attitudes to teaching across subjects and teaching levels.

In order to assess attitudes or ideologies which might allow prediction of the teacher's classroom behaviour it would seem desirable to construct instruments that fulfil the following conditions:

(a) they should be based on *teachers'* constructs of 'good teaching' (unlike MTAI);

(b) they should be directly related to classroom teaching (unlike the Manchester Scales);

(c) they should take account of subject area.

Taylor, Christie and Platts (1970) attempted to construct such an instrument in the area of science teaching. Initially they maintained two concurrent activities: the development of a conceptual framework or sub-theory, and discussions with experienced science teachers of what constituted teaching effectiveness. These teachers produced some 300 statements relating to their perceptions of effective science teaching. The theoretical model on which the study was based has three basic dimensions:

A. the teacher's classroom behaviour and relationships;

B. the teacher's preparatory behaviour including his relationships with his colleagues and with pupils in general outside the classroom and laboratory;

C. the standing conditions or requirements of teaching including the teacher's professional qualifications in his subject and in the study of education, and his desire to pursue these further.

These areas are further sub-divided. 'A' covers classroom and laboratory organization; the act ot teaching; discipline and control; and evaluating. 'B' covers lesson planning and preparation; activities and relationships within the school and outwith the school; and co-operation with other teachers. 'C' covers qualifications and training; attitudes, values and interests; and competence and professionalism.

The authors found that they were able to allocate 106 of the 300 statements to the 13 sub-divisions of the model without apparent ambiguity, duplication or undue complexity. These 106 items were placed in random order and administered as a 'Questionnaire Rating Effective Science Teaching' to 58 science teachers in England.

Factor analysis of the responses to the questionnaire suggested eight second-order factors isolated as perceived characteristics of styles of effective science teaching, about 45 of the original 106 items having sufficiently high loadings on the factors to be used in defining the eight styles. The authors suggest that these, together with additional items defining the eight scales could be used to construct a 'Perceptions of Effective Science Teaching' scale. The interpretations put by the authors on each of their eight factors are not immediately obvious on inspection of items loading highly on these factors. This is perhaps reflected in the rather vague (though plausible) names given to some factors:

(1) 'Face-to-face science teaching'.

(2) 'Face-to-science science teaching'.

(3) 'Impersonal, evaluative (pens out, facts down) science teaching'.

(4) 'Pupil autonomy science teaching'.

(5) 'Classroom management science teaching'.

(6) 'Competent science teaching'.

(7) 'Laboratory-assistant or production-line science teaching'.

(8) 'Science subject matter teaching'.

Such an instrument appeared to be appropriate for our study but before use could be made of the scales developed by Taylor *et al.*, there was a need to replicate their work in order (a) to test the items in the Scottish context; (b) to test a larger sample; (c) to test the generality of the factor structure. Inspection of the items loading heavily on the factors isolated in the English study suggested that they would all be relevant to science teaching in Scotland. Accordingly, these items were extracted from the original 106-item questionnaire and used to investigate perceptions of effective science teaching among Scottish science teachers.

Procedure

A questionnaire of 44 items, distributed at random, was drawn up. This consisted of those items of the 106-item scale described above with the highest loadings on each of the eight factors. A copy of this questionnaire will be found in Appendix A-20A. The factor number from the Taylor *et al.* study on which each item had its highest loading is indicated. Examples of items in each of the categories A, B and C of the model are also to be found in Appendix A-21A.

Each factor was described by at least four items and only one of the factors was bipolar. The distribution of items among the three basic dimensions of the model for those items selected for the 44-item scale was not significantly different ($\chi^2 = 1.8911$; df = 2) from that of those items NOT selected for the scale. The distribution of items among the three categories of the model for the 106-item scale and the 44-item scale are shown in Table 9.1.

The questionnaire was completed by 157 science teachers all of whom were involved in first and second year secondary school science teaching. The sample had a mean teaching experience of eight years and 42 per cent had less than five years' teaching experience. They came from 39 of the 40 schools that formed the stratified random sample for this study. Each of the items was rated on a five-point scale, 5 signifying that the item was considered 'extremely important' and 1 signifying that the item was 'not relevant in science teaching'.

Analysis

Since this is an exploratory study of science teachers in Scotland factor analysis appeared as an appropriate technique, and in order to make

TABLE 9.1 THE DISTRIBUTIONS OF ITEMS BETWEEN THE THREE CATEGORIES OF THE MODEL FOR THE 106 - ITEM SCALE AND THE 44 - ITEM SCALE

	106 - item scale (English)	44 - item scale (Scottish)
A: Teacher's Classroom Behaviour	53.8% (n = 56)	61.4% (n = 27)
B: Teacher's Preparatory Behaviour	12.5% (n = 13)	9.1% (n = 4)
C: Standing Requirements of Teaching	33.7% (n = 35)	29.5% (n = 13)

comparisons with the earlier study the procedure of Taylor *et al.* was followed as closely as possible. The mean value and standard deviation for the Scottish teachers' scores for each of the 44 items was calculated. A Pearson product-moment correlation was used to investigate the relationship between these 44 mean scores with the mean scores of the English teachers on the same items. 't' tests were applied to the differences between the Scottish and English means on each item.

Correlations between item scores for Scottish teachers were computed and factor analyses at the first and second-order level were carried out using STATPACK, a program developed at the Computer Unit, University of Warwick, from the Hallworth and Brebner package for multivariate analysis (Hallworth and Brebner, 1965). This program provides principal components from the inter-variable correlation matrix, a rotation to the varimax criterion, transformation of the varimax factors to promax oblique factors, and second-order factors (principal components followed by varimax). The factor analysis was carried out on raw scores for the teachers' responses to the 44 items. The mean score and standard deviation for each teacher over the 44 items was then calculated, the standardised z-score for each teacher on each item was computed, and the factor analysis was repeated using z-scores.

Results

Figure 9.1 shows the distribution of mean scores on the 44 items. There was a high correlation ($r = 0.88$) between mean scores on each item for the Scottish teachers and the English teachers. 't'-tests for the difference between mean scores of the two samples from the same population (*i.e.*,

FIG. 9.1 DISTRIBUTION OF MEAN SCORES ON ITEMS FROM 44 - ITEM SCALE (SCOTTISH)

1.0–1.5	1.5–2.0	2.0–2.5	2.5–3.0	3.0–3.5	3.5–4.0	4.0–4.5	4.5–5.0
					42	44,	
					39, 41	40, 43	
					32, 35	36, 37	
				30, 34	28, 31	25, 29	
				23, 24	26, 27	16, 19	
				20, 22	17, 21	12, 13	
			38	9, 14	10, 15	4, 5	
		18,	8, 33	6, 7	3, 11	1, 2	

1.0 Not relevant, 2.0 Unimportant, 3.0 Important, 4.0 Very important, 5.0 Extremely important

Rating scale

Numbers refer to item numbers on 44 - item questionnaire (Appendix A) secondary school science teachers) showed significant (above 5 per cent level) differences on 10 items. Items with significant 't' values are listed below:

(a) Rated higher by teachers in England

5. Relates new learning to natural phenomena within experience of pupil in order to develop meaningful associations. (t 4.6; $P<0.001$).
7. Encourages pupils to set themselves goals according to their abilities. (t 4.6; $P<0.001$).
9. Affects his pupils so that they wish to take more advanced courses in science. (t 2.0; $P<0.05$).
17. Can help pupils differentiate between hypotheses, facts, superstition and theory as well as encourage pupils to suspend judgment when faced with inadequate scientific evidence. (t 2.1; $P<0.05$).
19. Has patience in dealing with pupils. (t 2.1; $P<0.05$).
24. Can locate sources for free and inexpensive science teaching material. (t 2.7; $P<0.01$).
30. Encourages pupils to bring appropriate materials and specimens to class. (t 3.1; $P<0.002$).
33. Helps pupils to prepare for a career in science or technology. (t 2.3; $P<0.05$).

(b) Rated higher by teachers in Scotland

36. Is confident and at ease when teaching. (t 3.4; $P<0.001$).
37. Knows how to proceed if there is a serious problem of discipline. (t 3.8; $P<0.001$).

The 44 mean item scores were divided into three rating groups: 'high', 'medium' and 'low'. These groups were arbitrarily defined to allow the numbers of items in each group to be roughly equal. 'High' corresponds to ratings above 4.0 ('very important'), 'medium' to ratings between 3.5 and 4.0, and 'low' to ratings below 3.5. There were only four items in the *Teacher's Preparatory Behaviour* category, one with 'medium rating' and three with 'low' rating. There was no significant difference between the distributions of items in the *Teacher's Classroom Behaviour* and *Standing Requirement of Teaching* categories over the 'high', 'medium' and 'low' ratings (χ^2 = 0.13; df = 2). The corresponding value of χ^2 = 3.10 (df = 2) for the 106-item scale in the English study, though considerably higher, is still not significant. If the English responses on the 44 items only are considered, an even lower value of χ^2 = 0.06 (df = 2) is obtained (see Tables 9.2 and 9.3).

Items with 'high' mean rating for Scottish teachers (*i.e.*, above 'very important', are listed below. With the exception of two items (1 and 25) they appeared in the 'high' group of the English study.

1. Is consistently fair and emotionally calm when enforcing rules.
2. Teaches for understanding rather than reproduction of learned material.
4. Has a genuine interest in science and believes in the academic and practical use of the subject.
5. Relates new learning to natural phenomena within the experience of the pupil in order to develop meaningful associations.
12. Can devise experiments which involve pupil participation in learning.
13. Willingly consults colleagues in case of professional difficulties.

TABLE 9.2 DISTRIBUTION OF ITEMS WITH 'HIGH', 'MEDIUM' AND 'LOW' RATINGS AMONG THE THREE BASIC DIMENSIONS OF THE MODEL

SCOTTISH TEACHERS : 44 - item scale

	No. of items rated 'high' (score > 4.0)	No. of items rated 'medium' (4.0 > score > 3.5)	No. of items rated 'low' (3.5 > score)	Total No. of items in category
A. Teacher's Classroom Behaviour	10	10	7	27 (61%)
B. Teacher's Preparatory Behaviour	0	1	3	4 (9%)
C. Standing Requirements of Teaching	5	4	4	13 (30%)
ENGLISH TEACHERS : 106 - item scale				
A. Teacher's Classroom Behaviour	21	26	10	57 (54%)
B. Teacher's Preparatory Behaviour	1	5	7	13 (12%)
C. Standing Requirements of Teaching	10	14	12	36 (34%)
ENGLISH TEACHERS : 44 - item scale				
A. Teacher's Classroom Behaviour	10	12	5	27 (61%)
B. Teacher's Preparatory Behaviour	0	2	2	4 (9%)
C. Standing Requirements of Teaching	5	6	2	13

TABLE 9.3

VALUES OF χ^2 FOR DISTRIBUTIONS OF 'HIGH', 'MEDIUM' AND 'LOW' ITEMS AMONG TWO CATEGORIES OF THE MODEL

	χ^2 (df = 2)
Distribution of "Scottish" items in categories A and C among "high", "medium" and "low" rating groups. (44-item scale)	0.18 (N.S)
Distribution of "English" items in categories A and C among "high, "medium" and "low" rating groups. (106-item scale)	3.10 (N.S)
Distribution of "English" items in categories A and C among "high", "medium" and "low" rating groups. (44-item scale)	0.06 (N.S)

16. Tries to stimulate pupils to think for themselves about science.
19. Has patience in his dealing with pupils.
25. Is clear and unequivocal in his personal relationship with pupils.
29. Changes curriculum and methods to keep up to date with developments in his subject and methods for teaching it.
36. Is confident and at ease when teaching.
37. Knows how to proceed if there is a serious problem of discipline.
40. Develops interests in science in his pupils.
43. Is willing to change an opinion or conclusion because of later evidence.
44. Is constructive and helpful in his criticism of pupils.

Of the items with mean scores of less than 3.0 (*i.e.*, below 'important') listed below, two (8 and 18) were also rated below 'important' by the teachers in England.

8. Can interpret the results of diagnostic instruments used in schools (IQ, aptitude and achievement tests).
18. Uses pupils to carry out routine duties such as giving out books, cleaning the blackboard, *etc.*
33. Helps pupils to prepare for a career in science or technology.
38. Has studied the philosophy and psychology of education.

Since the items forming the scale were based on those whose loadings most completely defined the eight factors extracted previously, it was hoped that the first-order factor analysis might produce an appropriate group of factors. Using the Kaiser criterion, that all factors with roots greater than unity be extracted in the principal components, 12 factors emerged. It will be seen, however (Table 9.4) that the total variance accounted for by these twelve is 64.2 per cent of which 29.1 per cent is due to Factor 1. The latent root values of the other factors suggest they are all scree and that there is little justification for accepting more than one general factor (see Figure 9.2). The inclusion of three more factors with roots just less than unity in the table supports the view that there is little reason for a cut-off at 12 factors (see Table 9.4).

However, for interest the first-order varimax and promax solutions were computed and followed by a second-order principal components and varimax. The second-order analysis identified three factors accounting for 54.8 per cent of the variance. When the loadings for each item were examined no relationship between the factors emerging from Taylor's study and either our first- or second-order factors could be seen.

The dominant general factor in the first-order principal components analysis had a positive loading above 0.45 for 40 out of the 44 variables making item discrimination difficult. This suggested that the factor might be accounted for by the response sets of individual teachers (*i.e.*, one teacher might take 3-'important' as a mean position while another might take 4-'very important' as a mean position; similarly the range of ratings will vary between individual teachers). This would lead to positive correlations between scores on any pair of items. The data showed that out of 946 such correlations 940 were, indeed, positive.

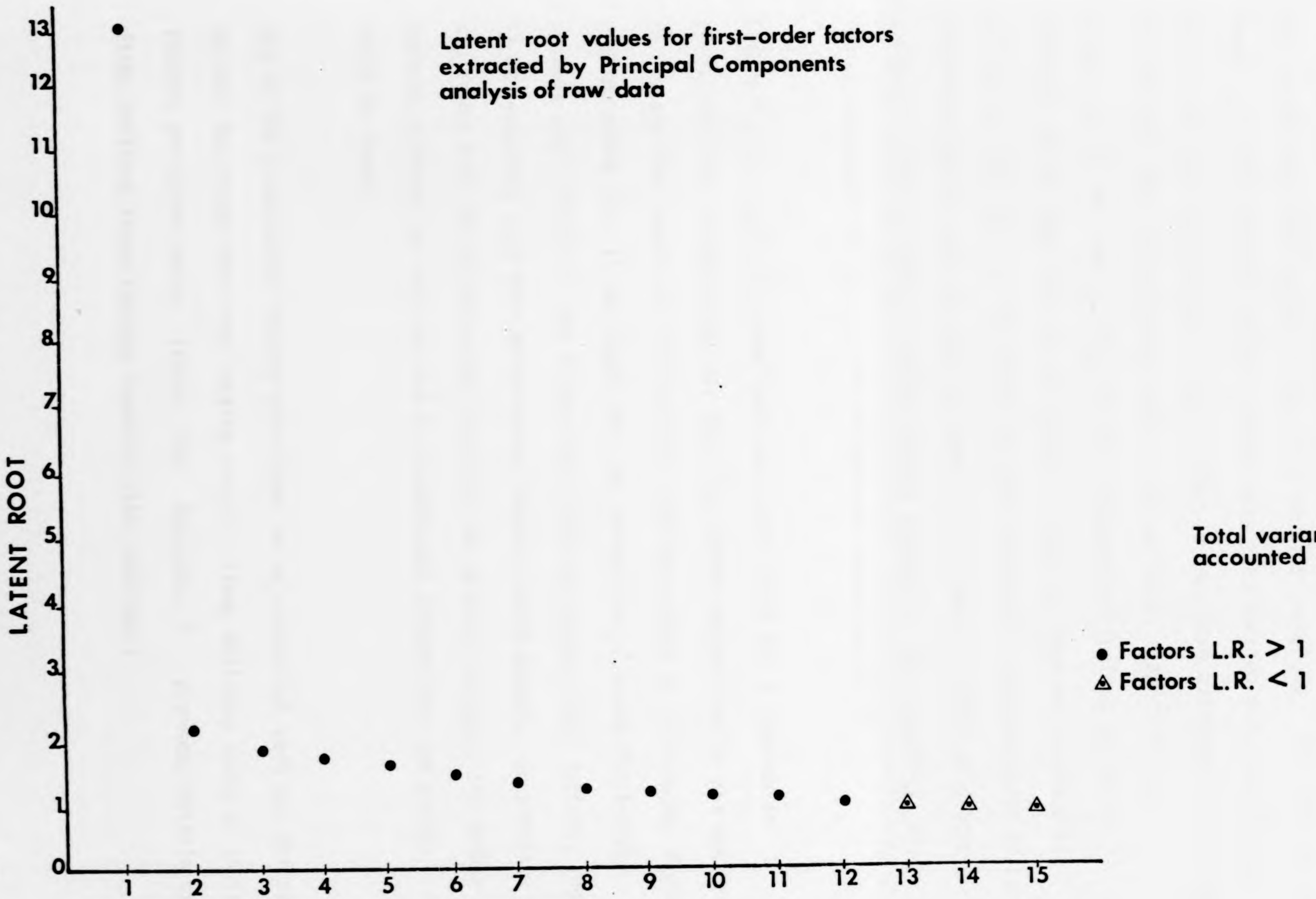
TABLE 9.4

LATENT ROOT VALUES AND PERCENTAGE VARIANCE ACCOUNTED FOR BY FIRST-ORDER FACTORS EXTRACTED IN PRINCIPAL COMPONENTS ANALYSIS OF RAW DATA

Total variance accounted for 64.2 per cent

Factor Number	Latent Root	% variance accounted for
1	12.81	29.1
2	2.13	4.8
3	1.76	3.0
4	1.68	3.8
5	1.62	3.7
6	1.43	3.2
7	1.37	3.1
8	1.18	2.7
9	1.15	2.6
10	1.10	2.5
11	1.04	2.4
12	1.01	2.3
13	0.96	2.2
14	0.90	2.1
15	0.89	2.0

Latent root values for first-order factors extracted by Principal Components analysis of raw data



(Fig. 9.2)

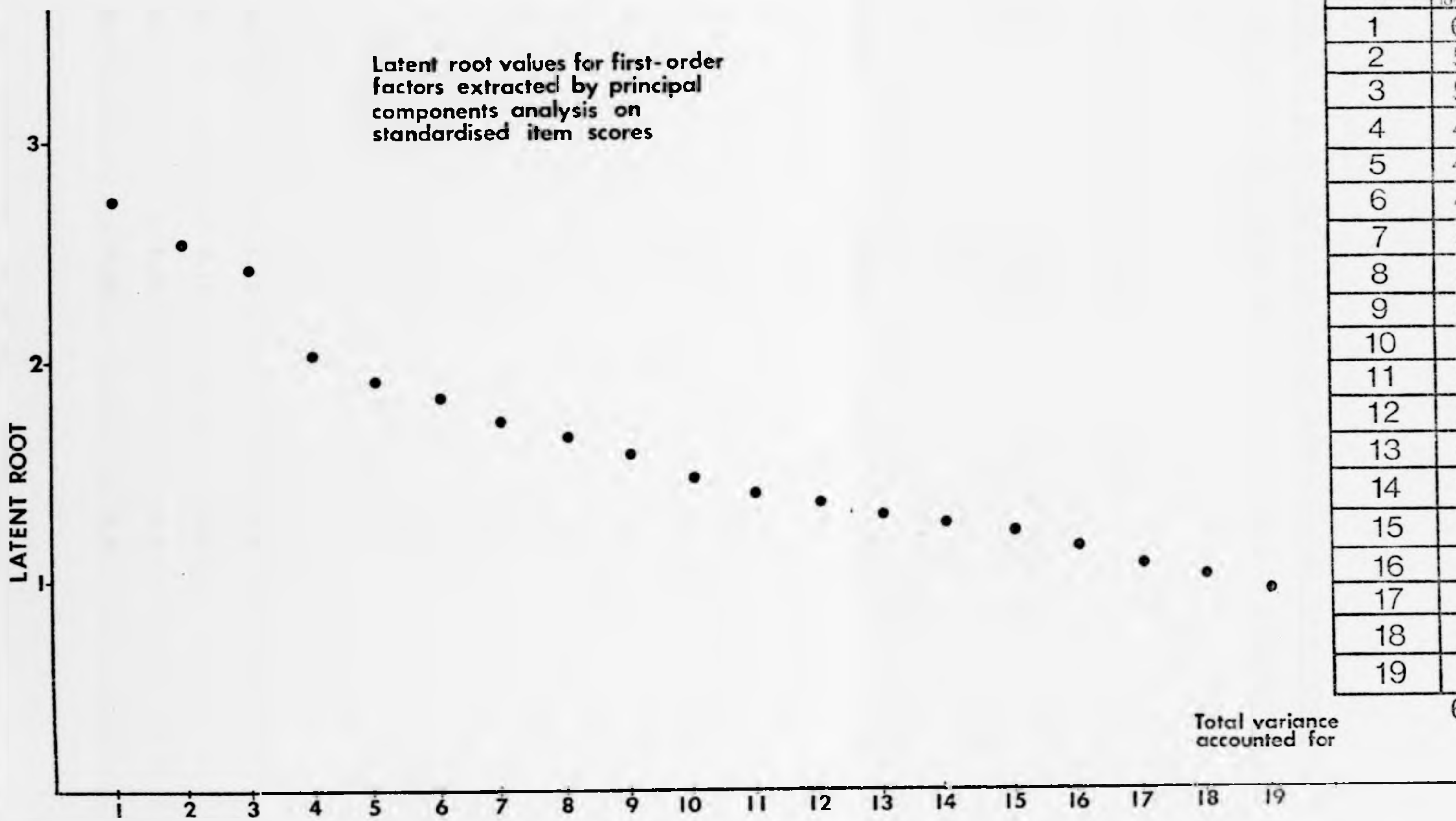
FACTOR NUMBER	% variance accounted for
1	29·1
2	4·8
3	4·0
4	3·8
5	3·7
6	3·2
7	3·1
8	2·7
9	2·6
10	2·5
11	2·4
12	2·3
Total variance accounted for	64·2%

To avoid this problem standardized z-scores were computed for each individual teacher so that each individual's distribution of ratings had the same mean and standard deviation. This made no difference to the rank order of items' mean scores and that aspect of the study remains unchanged. Correlations and factor analysis based on these z-scores were then carried out, and nineteen first-order and seven second-order bipolar factors were extracted accounting for 69.3 per cent and 61.8 per cent of the variance, respectively. Inspection of the item loadings on the second-order factors failed to identify any of them with any of Taylor's factors. However, a plausible (though as with the English study far from 'obvious') interpretation of our factor structure could be made in terms of the identification of discrete styles of teaching among *Scottish* science teachers. Unfortunately, this factor structure was found to be unstable as described below.

Figure 9.3 and Table 9.5 show that once again there was a problem in justifying the cut-off point for the first-order factors and it was decided to compare the results of two analyses, one extracting 15 first-order factors and the other 19. It was hoped that the extraction of extra first-order factors would result in the firmer definition of second-order factors, and it was expected that more second-order factors would appear. In practice it was found that the second-order structure was greatly changed; the number of factors differed by only one but no resemblance between the two groups could be found.

Six of the *first-order* factors were found to be stable but each was defined by only two items with high loading (>0.5). Items defining three of these factors are given below. (Table 22A Appendix A provides details of items defining these factors together with loadings.)

Latent root values for first-order factors extracted by principal components analysis on standardised item scores



FACTOR NUMBER	% variance accounted for
1	6·1
2	5·6
3	5·4
4	4·6
5	4·4
6	4·0
7	3·9
8	3·8
9	3·6
10	3·3
11	3·1
12	3·0
13	3·0
14	2·9
15	2·8
16	2·6
17	2·5
18	2·4
19	2·3
Total variance accounted for	69·2%

(Fig. 9.3)

TABLE 9.5

LATENT ROOT VALUES AND PERCENTAGE VARIANCE ACCOUNTED FOR BY FIRST-ORDER FACTORS EXTRACTED IN PRINCIPAL COMPONENTS ANALYSIS ON STANDARDISED ITEM RATINGS

Total variance accounted for: 69.3 per cent

Factor Number	Latent Root	% variance accounted for
1	2.66	6.1
2	2.46	5.6
3	2.38	5.4
4	2.04	4.6
5	1.94	4.4
6	1.76	4.0
7	1.71	3.9
8	1.65	3.8
9	1.56	3.6
10	1.44	3.3
11	1.38	3.1
12	1.33	3.0
13	1.30	3.0
14	1.26	2.9
15	1.22	2.8
16	1.15	2.6
17	1.11	2.5
18	1.06	2.4
19	1.01	2.3

Factor 1.

36. Is confident and at ease when teaching.

37. Knows how to proceed if there is a serious problem of discipline.

Factor 2.

34. Can apply his knowledge of the psychology of learning to the teaching of his subject.

38. Has studied the philosophy and psychology of education.

Factor 3.

33. Helps pupils to prepare for a career in science or technology.

9. Affects his pupils so that they wish to take more advanced courses in science.

Most pairs provided a fairly straightforward factor interpretation but no pair was to be found together in one of the groups of items defining Taylor's factors.

Discussion

While there was a high correlation between the English and Scottish teachers' ratings of individual items, no clear picture of emphasis on a particular category of the model of teaching effectiveness emerged from the Scottish responses to the shortened version of the scale. It would appear that items in the category of *Teacher's Preparatory Behaviour* were rated low, but there was no evidence from either Scottish or English ratings on the 44 items to support the claim that 'science teachers place the greatest emphasis on the dimension of effective science teaching concerned with the teacher's classroom behaviour, and the least on the dimension concerned with the standing requirements of teaching' (Taylor *et al.*, 1970).

The two items rated higher by the Scots than the English were concerned with smooth running of the classroom. It is, perhaps, not surprising that the Scottish teachers, the majority of whom were concerned with teaching Integrated Science to lower secondary age-groups from the whole range of abilities in mixed ability classes, should rate such items more highly than the English teachers who were teaching separate science subjects to pupils of average or above average ability. Explicit reference to pupils in seven out of eight items scored higher by teachers in England points to a far more pupil-centred strategy than that of the Scots.

From the items rated 'high' by the Scots there emerges an 'ideal' picture of science teaching in Scotland where the pupils are kept under control but in an atmosphere without undue restraint; with a teacher who presents the material in a clear and unbiased way placing emphasis on pupils' participation, understanding and interest in the work; and where the teacher's interests in science are supported by a willingness to seek outside help and modify his opinions and teaching programme. With seven of the 15 'high' items referring explicitly to the pupil we might conclude that a pupil-centred approach was favoured, but this proportion is not significantly different from that of the scale as a whole, for which 21 out of 44 items refer explicitly to pupils.

Of the items rated below 'important', two (8 and 38) perhaps indicate a lack of emphasis on the theoretical and evaluative aspects of education. Analysis has not revealed any salient factors by which Scottish science teachers could be classified in terms of variation in their perceptions of effective science teaching. Variations from the common pattern of priorities found in both England and Scotland appear to be idiosyncratic rather than to reflect differing ideologies.

It is possible that the failure to replicate the English findings is related to the more centralized pattern of education in Scotland. The fact that the examination syllabuses, secondary science curriculum development, and in-service training courses have been largely in the hands of the same group of people, may have led to a relative absence of ideological debate among Scottish science teachers and so inhibited the development of clear-cut dimensions of ideology.

However, speculation on the differences in structure of ideologies between science teachers in Scotland and England would be unwise until replication of Taylor's study has shown that the instrument can be assumed valid for application to English teachers. Regrettably, it is clearly invalid for discriminating among Scottish science teachers.

The items of Taylor's questionnaire appear to provide a comprehensive coverage of those views of teachers that might be expected to influence their classroom teaching. The failure of these items to identify any salient attitude factors among the teachers suggests that there are no distinct ideological dimensions, closely related to their behaviour in the science classroom, by which teachers may be characterized. Characterization may only be possible on very specific scales relating to views on narrow areas (such as those of the first-order factors on page 285). Such scales would not be sufficiently general to account for much of the variance in teacher behaviour and even less of the variance in pupil attitudes. It was, therefore, concluded that this method for assessing the 'teacher' influence on pupil attitudes should be abandoned.

CHAPTER 10

COGNITIVE PREFERENCES IN SCIENCE

It was suggested in Chapter 8 that it would be desirable to assess the scientific attitudes of pupils from some other perspective in addition to that of the Likert attitude scale. Furthermore, it appeared that the test of Cognitive Preferences in Science might provide an appropriate means for such assessment.

There is a substantial body of literature describing empirical studies of cognitive preferences, but there is little available regarding the nature of such preferences, their construct validity, or the statistical analysis of data collected.

This chapter then examines, in some detail, the structure of the Cognitive Preference test instrument, the results of empirical studies, the nature and validity of the cognitive preference constructs, the various methods of analysing the data, and, finally, a study of the cognitive preferences of 218 pupils following the Integrated Science course and 227 pupils following the separate science subjects course.

Review of Studies in Cognitive Preference

It is a decade since Heath (1964) first introduced a cognitive preference test in physics. He proposed this as an instrument to supplement traditional testing procedures in assessing the achievement of the distinctive goals of the new science curricula. He suggested that while many of the goals of the new courses were identical with those of the traditional courses, emphasis was now put on a 'firm understanding of basic ideas that permeate the discipline rather than technical vocabulary and practical applications' and that they have 'tried to teach the student how to acquire, evaluate and retrieve knowledge in a field and have devoted less energy to presenting him with a survey of this knowledge'. This, he argued, could lead to a difference in the mode of attending to material exhibited by pupils following different courses. Such modes he termed 'cognitive preferences', and his suggestion was that our interest should be in what a pupil does with information (*i.e.* his intellectual style) rather than whether he can identify the information as correct or incorrect. He proposed the Cognitive Preference Test as an appropriate measure of the achievement goals that are designed to change a pupil's intellectual style. His purpose was to permit the pupils to 'demonstrate differing styles' using test items that 'allow the student to exhibit some preference in cognition'.

The test instrument developed by Heath aimed at the identification of four cognitive preferences:

(a) MEMORY (recall) of specific facts or terms,

(b) practical APPLICATION,

(c) critical QUESTIONING of information,

(d) fundamental PRINCIPLES.

Each item contained an introductory (or stem) statement, sometimes accompanied by an illustration. This was followed by four 'alternative' statements and the pupil was told to select from the four choices the one that most appealed to him. He was told that all the 'answers' were correct and that choice was a matter of preference only. Each of the four alternatives was closely related to the introductory statement and was designed to represent a particular cognitive preference:

> 'The acceleration of a body is directly proportional to the force acting on it.
>
> (A) When gases from a rocket produce a constant "thrust", acceleration results.
>
> (B) f = ma (where f = force; m = mass; a = acceleration).
>
> (C) If, according to the "Law of Reaction" forces always occur in pairs (equal and opposite) how can acceleration result?
>
> (D) This implies that the final velocity attained by a massive body is limited only by the duration of the accelerating force' (Educational Testing Service, undated).

'B' is a factual statement closely related to the introduction and endorsement of this choice contributes to the pupil's score on a preference for facts or MEMORY (M). 'A' describes a practical APPLICATION (A) of the information provided in the stem. 'C' is designed to indicate a pupil's preference for expanding or critically QUESTIONING (Q) the introduction. 'D' is a statement of fundamental PRINCIPLE (P), generalization, or abstraction underlying the introductory statement.

Taking the Physical Science Study Committee (PSSC) goals as an exemplification of the goals of the new curricula, Heath hypothesised that pupils following a PSSC course would demonstrate a stronger preference for fundamental Principles and Questioning than would pupils following a traditional course, and that pupils following a traditional course would demonstrate a stronger preference for Memory of facts and terms and for practical Application. His sample consisted of 30 PSSC classes (1027 pupils) and 49 Control classes (2100 pupils). His results showed significant differences between the mean of the PSSC class means and the mean of the Control class means on preferences for Memory, critical Questioning, and Principles (in the direction confirming his hypotheses) but not on Application. In addition he was able to support a further hypothesis that scores on Principles and critical Questioning related more positively to achievement among the PSSC pupils than for the Control group (this was the case for both 'traditional' and 'PSSC' physics achievement tests). Scholastic aptitude (as measured by the School and College Ability Test) showed its highest correlations with Application for the PSSC group (unexpected) and with Memory for the Control group (as expected). A test purporting to measure 'the ability to change the function or significance of structural elements of an object to use them in a new way' (Concealed Figures Test), showed its highest correlation with Principles for the PSSC group (r = 0.42) and with Memory for the Control group (r = 0.41).

In his discussion of results Heath is concerned with:

a) achievement of course objectives - 'the data reported here indicate the PSSC students show a greater inclination to attend to fundamental concepts and to ask questions in pursuit of basic understanding. These differences are consistent with the reported objectives of the PSSC course'.

and b) identification of cognitive style - 'the generally positive results obtained suggest that this type of instrument can identify in a meaningful context curriculum-related differences in cognitive style....When such objectives are operationally expressed in a curriculum it appears that a cognitive preference test can assess curricular effects on the students' styles of cognition in that discipline. The contrasts between groups in Cognitive Preference Test scores and Concealed Figures Test scores are consistent with the notion of different cognitive preferences being formed by differing curriculums'.

Mackay (1972) uses a Cognitive Preference Test, similar to that of Heath, on PSSC pupils in schools in Victoria, Australia. After a survey of PSSC literature he concludes that a) the course places increased emphasis on fundamental principles and decreased emphasis on factual recall compared with traditional courses; and b) the course tends to concentrate on the theoretical aspects while giving little emphasis to the technological applications of physics. He derives two objectives for the course from these findings:

'1. Students should develop a greater preference for critical questioning and theoretical extension of information over practical applications of physical information.

2. Students should develop a greater cognitive preference for identification of fundamental principles over rote remembering of specific facts or terms.'

Two cognitive preference scores (Q - A, and P - M) are used on this test since Wish (1964) concluded from a reanalysis of Heath's data using unfolding analysis (Coombs, 1964) that two dimensions were sufficient to account for the P, M, Q and A scores.

Mackay was able to demonstrate changes in the cognitive preferences of pupils over two years (Grades 11 and 12) in the direction specified by the objectives (he does not, however, make clear the significance levels). He appears to accept the appropriateness of 'cognitive preference' as a measure of 'cognitive style' (at one point he uses the phrases synonymously) but gives no theoretical interpretation of either. Indeed, it is difficult to see why he introduced the concept of 'style' into this study.

Marks (1967) also asserts the equivalence of 'cognitive style or preference' but then goes on to use the cognitive preferences to evaluate chemistry objectives rather than to identify constructs of cognitive styles. He does not make the objectives explicit but claims to have adopted these cognitive preferences 'because the "memory" and "practical application" preferences were closely related to the goals of instruction in the traditional course and the "questioning" and "principle" preferences were closely related to the goals of the CBA course of study'.

Testing a group of 24 classes following the CBA course together with 30 classes forming a Control group he was able to show that at the 0.01 level of significance the CBA group scored higher on the critical Questioning preference and the Control group scored higher on the Memory scale. At the 0.05 level of significance the CBA group scored higher on fundamental Principles, while there was no significant difference between scores for the two groups on Applications.

Marks reaches the reasonable conclusion that the CBA course is achieving at least one of its objectives, but he then goes on to say 'The cognitive style of the CBA students seem more tied to conceptual learning rather than to verbal structure'. He does not attempt theoretical justification of the leap to this statement from the observation that the science cognitive preference scores for this group tend to emphasise Q over M when compared with a Control group.

Atwood (1967a, 1967b) has developed and used a Cognitive Preference Test in chemistry. He used it purely as a measure to help characterize the CHEM Study courses and does not appear concerned with the identification of cognitive styles of pupils. He predicted that pupils showing strong preference for Memory would be at a disadvantage in these courses when compared with those showing strong preference for Principles, Questioning or Application. He was unable to isolate a group showing a strong preference for Memory, but investigation of CHEM achievement scores for groups with various cognitive preferences suggested that preference for Applications in combination with either Principles or Questioning would be most advantageous.

Kempa and Dubé (1973) have suggested that the science cognitive preference studies described above have established that different curricular treatments lead to distinct differences in 'thinking style'. Further, they identify these 'thinking styles' with 'cognitive styles' which they define as 'characteristic modes in which individuals can attend to information or tasks presented to them'. They conclude that the pupil's 'overall cognitive style in the context of science may be deduced from the general response pattern in a cognitive preference test'. The apparent effect of different science courses on pupils' cognitive preferences is viewed by them as support for Glaser's (1970) view that pupils' cognitive styles are shaped by the processes they are obliged to use in a particular course.

The identification of cognitive style is seen by Kempa and Dubé as firstly, providing insight into appropriate adjustments of teaching strategies in relation to the particular learning characteristics of the pupils, and secondly, as an additional element to achievement tests in the assessment of individual pupils and in the evaluation of 'educational processes'. Their own study is concerned with a) the determination of science cognitive preferences of post 'O' level chemistry pupils, and b) the examination of the 'effect of academic achievement in Chemistry, as measured by 'O' level grades, on students' cognitive preference profiles'.

Using inter-preference correlations and R-type factor analysis a model emerges from the study in which the cognitive preferences are represented by two independent bi-polar axes, on one of which M is contrasted with Q while on the other P is contrasted with A (*cf.* the models of Mackay, 1972, and Wish, 1964, where M is contrasted with P and Q is contrasted with A). The authors acknowledge that their statistical procedures may be inappropriate and this point will be taken up later.

The M - Q scale is identified as a 'scientific curiosity' scale; a high M (and a low Q) score indicates a pupil's preference for simply restating information, while a high Q (and a low M) score indicates a preference for extending and critically questioning the information. The A - P scale is identified as an 'applied/pure' dimension and the score on this scale indicates whether a pupil prefers the 'practical value' (A) or the 'scientific implications' (P) of the information.

One-way analysis of variance for each of the M, A, Q and P scales for four groups of 'O' level Chemistry achievement, showed differences in cognitive preferences ($P<0.01$). The mean scores indicated that high achievers preferred P and Q over M and A (*i.e.* high 'curiosity' and emphasis on 'pure' science) while low achievers preferred M and A over P and Q (*i.e.* low 'curiosity' and emphasis on 'applied' science). The authors conclude that 'comparison of the mean scores for students in different achievement categories reveals a strong dependence of cognitive preference orientation on academic achievement' (presumably they are using 'dependence' in the sense of 'correlation' and not 'causation').

Kempa and Dubé admit that their results have no predictive value for individuals but suggest that 'they provide some guide to the relationship between cognitive outcomes of 'O' level Chemistry courses and qualities which are, strictly speaking, non-cognitive'. They provide no argument for the non-cognitive nature of cognitive preferences, indeed the strong relationship with achievement may well indicate that they are cognitive.

Schmedemann and Shier (1967) view cognitive preferences as having a cognitive nature with each choice corresponding 'to a different type of knowledge or cognition' and 'each score representing a different cognitive process'; they do not indicate their reasons for choosing to administer cognitive preference tests to PSSC pupils (other than the fact that the tests had already been used for PSSC by Heath) and make no attempt to discuss what is being measured. They combine the individual weighted measures of M, A, Q and P into one single score. The weightings were selected arbitrarily after inspection of PSSC goals ($W_M = W_A = 1$, $W_P = 2$, $W_Q = 3$). The purpose of this study was to investigate the relationship between pupils' cognitive preferences and certain teacher characteristics. These characteristics were a) teachers' cognitive preferences, b) teacher 'warmth', c) teacher 'demand', and d) teacher 'use of motivation'. b), c) and d) were determined by pupil ratings on the Reed Student Inventory (1962). The analysis used the Kendall rank correlation coefficient and the authors concluded that there was

> 'no appreciable interaction...between the students' median cognitive preference score and the teacher's cognitive preference score or the student's evaluation of the teacher on the "warmth", "demand", and "use of motivation" scales'.

Atwood (1969) has continued his use of the Chemistry Cognitive Preference Test in an attempt to measure changes in cognitive preferences resulting from a physics course. His purpose in carrying out this study is not clear, he merely suggests that 'The possibility of measuring changes in cognitive preferences is an appealing one'. Presumably he was disappointed since he concludes that his pre- and post-physics course measures of chemistry cognitive preferences provide little support for the idea 'that some transfer of cognitive style might occur between closely related subject areas'. He has now developed a Cognitive Preference examination composed of items whose content is drawn from general science, social science, and closely related fields (Atwood, 1971). These, he suggests, will encompass material that will be meaningful to 'senior-high' level pupils, and will enable studies to be made on changes of pupils' cognitive preferences since this content will be familiar to the pupils at the pretest and post-test. It should also facilitate the investigation of whether or not experience in one subject area affects cognitive style in another area. In the absence of any theoretical formulation by him of what is being measured by scores on cognitive preference tests, studies of these two types lack clarity of purpose.

Barnett (1974) has investigated relationships among Biology achievement, pupils' perception of teacher style and cognitive preferences. He is concerned with the characteristics of inquiry teaching and hypothesises that 'students who show a strong preference for questioning and application could be expected to feel more comfortable, and to work more effectively using inquiry, than students who show a strong preference for memory', and 'A nondirective teaching style could be expected to promote achievement in an inquiry oriented classroom'. These relationships were explored in a study of one semester of instruction in Biology.

He administered a cognitive preference test (pretest and post-test), based on three cognitive preferences, to 1477 tenth-grade Biology students. He was able to classify 694 of these students into three categories by their strong preference for Application, Memory or Questioning. Mean Biology achievement scores (post-test) for each of these groups were compared using one-way analysis of variance, and the results indicated that pupils showing preference for Application or Questioning tend to be higher achievers than those whose preference is for Memory ($p<0.01$).

He also computed correlation coefficients, for all pupils in the sample, between achievement scores and scores on each of the three preference scales. A negative correlation of achievement with Memory ($p<0.05$), and positive correlations with Application ($p<0.05$) and Questioning (n.s.) were obtained, leading him to conclude 'that students who show a strong preference for application learn more effectively than students who display a strong preference for memory'.

Students were categorized by a second procedure, in terms of their score on a 'Student Perception of Teacher Style' scale, into a 'directive' and a 'non-directive' group. Analyses of covariance were used to compare the mean achievement scores and mean cognitive preference scores of the two groups. Corresponding pretest scores were used as covariates in these analyses. This procedure yielded F ratios from which Barnett concluded that there were significant differences at the 0.01 level between mean scores for the two groups on Biology achievement, Application and Memory, but not on Questioning.

While empirical studies in this area have been extensive, there appears to have been little attempt by these researchers to discuss the nature of pupils' cognitive preferences. This nature is not self-evident and it is disturbing to find that teachers are being encouraged (Tisher 1968) to try out cognitive preference tests under the assumption that they are identifying curriculum-related differences in cognitive style. Consideration of theoretical interpretations of the preferences is overdue.

The Nature of Cognitive Preferences

Preferences as course objectives

Work with PSSC pupils after Heath appears concerned with the achievement of course objectives, yet Mackay (1972) is the only author to make such objectives explicit. Marks (1967) and Atwood (1967a, 1967b) appear concerned with the objectives of other courses, but the vagueness of description of these objectives suggests they fit the cognitive preference instruments rather than the course. If the objectives of a particular course of study require that pupils exhibit a preference for material in the form of Questioning and Principles rather than Application and Memory, then it seems quite justifiable to construct, as Heath (1964) has done for PSSC, an instrument to measure the extent to which these objectives have been achieved. Whether it is useful to formulate course objectives in terms of cognitive preferences depends, however, on the value of such constructs in describing the cognitive activities or attitudes of pupils.

The validity of the cognitive preference constructs

Irrespective of what sort of constructs cognitive preferences are, the first question to be asked is whether or not we may assume that the M, A, Q and P measures are valid constructs for describing pupils' preferences for cognitive material in science. No validation (empirical or otherwise) of such constructs has been reported and there is no immediately obvious interpretation in terms of psychological theories of the nature of attitudes towards cognitive activities. It is very likely that a complete description of science cognitive preferences would require either more or less constructs, than the four used by Heath. He formulated these four from the literature describing the goals of new courses and comparing traditional courses (usually unfavourably) with innovations. In general this literature is written by practising natural scientists concerned with the scientific content of the course rather than the learning characteristics of the pupils. The Cognitive Preference Tests themselves cannot be used to validate MAQP scales as constructs for patterns of pupils' preferences (if, indeed, such patterns exist) since they force choices that are already formulated in terms of these constructs. However, if the constructs are valid indicators of preference patterns, we would expect high levels of internal consistency. Reported values of 0.37 to 0.68 (Kuder-Richardson Formula No. 20, Heath, 1964) and 0.28 to 0.70 (Marks, 1967) indicate low levels and suggest that MAQP are unsatisfactory as descriptive constructs of patterns of pupils'cognitive preferences. Even if this is the case, the value of 'cognitive preferences' as a construct need not depend on the appropriateness of the particular typology of preferences devised by Heath.

Preferences as indicators of pupils' cognitive styles

Kempa and Dubé (1973) claim to be concerned with cognitive style and the pupil's mode of attending to material. They outline the considerable potential that measures of cognitive style might have, but they provide no insight into the relationship between such styles and cognitive preferences. However, they identify their 'scientific curiosity' and 'pure/applied science' cognitive preference dimensions as task- or subject-specific cognitive styles of the type suggested by Glaser (1970). There is one point on which they differ from the majority of writers on cognitive styles, they see such styles as modes in which the pupil *can* attend to information or tasks while others see them as modes in which the pupil *does* attend to information or tasks. Definitions of cognitive style are usually concerned with describing the cognitive activity of the pupil. For example: 'the pre-dispositions of an individual to attend to his environment and to organise perceived elements in ways unique to him' (McCartin, 1970). In the Cognitive Preference Test, however, the pupil is not asked to do anything with, or about, his preferred statement. No evidence is provided to suggest that particular cognitive preference choices are related to particular cognitive activities. Kempa and Dubé refer to Messick's (1970) paper but give no indication of the place of cognitive preferences in his conceptualization of cognitive styles as

> 'several dimensions of individual differences in the performance of cognitive tasks...which represent a person's typical modes of perceiving, remembering, thinking, and problem solving'.

The organization of elements and characteristic behaviours seen by McCartin and Messick as necessary components of cognitive style, imply that observation of a complex system of inter-related behaviours of the pupil is needed to identify his cognitive style. It could be argued that such a behaviour system should be categorized at the level of Organization of a Value System or even Generalized Set (Krathwohl *et al.* 1964). While the determination of cognitive preferences by the procedure discussed might form one component in the assessment of such a behaviour system, it is clearly inadequate by itself as a measure of cognitive style.

Preferences as valuing or expression of an attitude

Cognitive preference choice might be identified with a verbalization of an attitude, in that it has an emotional quality (preference) associated with a cognitive element. We could then assume that a cognitive preference choice represents behaviour at the level Preference for a Value described by Krathwohl *et al.* (1964). This would indicate behaviour consistent with the individual's held values and, if we have a test that provides a valid measure of this behaviour, we would expect a high level of stability.

On this interpretation, we must ask what it is about the stimulus statements that pupils are expressing a preference for? It appears that the assumption has been made that pupils are expressing a preference for 'this process' (M, A, Q or P) which is implied by 'this' statement, rather than 'that process' which is implied by 'that' statement. The criteria that enable us to judge whether a statement demonstrates a cognitive process of Memorization, Application, critical Questioning or thinking in terms of Principles, have not, however, always been made explicit.

Heath (1964) simply states that 'In each item, each of the four options was designed to demonstrate a different form of cognitive preference in physics'. Kempa and Dubé (1973) had the information for the test items and responses 'scrutinized by an independent panel of experts' but give no details of how the relationships between the statements and the processes were established.

Atwood (1967a) used 'Three persons who are proficient chemists and familiar with the CHEM study program' to appraise the items independently. The criterion concerned with assigning M, A, Q or P to the options was that there be 'agreement by a minimum of two judges in keying each option'. It is questionable whether 'proficient chemists' are appropriate for judgment of implied cognitive processes, but more important is the point that disagreement by the judges on all four options of an item could lead to inclusion of that item in the test.

In a later study Atwood (1971) tightened this procedure. For the Cognitive Preference Examination using General Science and Social Science his panel of judges was composed of 'a person in science education, one in social science education, and one in educational research'. Critiques and revisions were carried out until 'unanimous agreement was obtained on the validity of each item'. He omitted the fundamental Principles option since he judged it to be less distinct than the others, and defined the other three as:

> '1. *Memory* or retention of factual information; is not an application; does not question or challenge. (M)
>
> 2. *Application* of information; answers what the information in the stem is "good for" or how it is useful; does not question or challenge. (A)
>
> 3. *Questioning;* challenging or expressing criticism of information in the stem. (Q)'

It would seem necessary to show that options for each item on each test could stand up to definitions at least as stringent as these before we may assume that the options demonstrate the cognitive processes.

Consideration of two options from the item example already quoted provide support for Atwood's decision to omit the P option:

> '(B) f = ma (where f = force; m = mass; a = acceleration).'
>
> '(D) this implies that the final velocity attained by a massive body is limited only by the duration of the accelerating force.'

I have suggested that (B) is a Memory statement and (D) is a statement of fundamental Principle, but it is not at all clear that there is any difference in the levels of generalization of the two. Indeed, it may be argued that any preference shown is simply for one symbolic form over another, rather than for a process at a higher or lower level of abstraction.

Even if a relationship between a statement and a cognitive process can be identified, it would seem at least as likely that it is the *content* of the statement that is being valued in a preference choice. Some statements may be preferred because the information they give is seen as interesting or important, while others are seen as unpleasant or peripheral. Let us consider a possible (but fictional) pupil response to the content of the item on Heath's Cognitive Preference Test described earlier:

> 'The acceleration of a body is directly proportional to the force acting on it.
>
> (A) when gases from a rocket produce a constant "thrust" acceleration results'.
>
> (Pupil: 'Oh no, not rockets again! Why can't they give us something on cars?')
>
> '(B) f = ma (where f = force; m = mass; a = acceleration)'.
>
> (Pupil: 'I don't see how the maths matters, it's what happens that's important').
>
> '(C) if, according to the "Law of Reaction", forces always occur in pairs (equal and opposite) how can acceleration result?'.
>
> (Pupil: 'I remember we spent a long time doing that Law of Reaction, it must be important').
>
> '(D) this implies that the final velocity attained by a massive body is limited only by the duration of the accelerating force'.
>
> (Pupil: 'How boring! I think I'll choose C').

He is familiar with the content and we would expect him to select Q again if asked to make his choice on this item a second time. There is no reason, however, to think that he will choose the Q alternative for the other items. This would lead to low values of internal consistency for the four scales, since the classification of content is probably independent of the MAQP pattern, but high values for test-retest reliability. Unfortunately researchers in this area have tended not to report both types of reliability coefficients. It has already been pointed out that, in general, internal consistency has been low. Kempa and Dubé (1973) obtained quite high values of test-retest reliability (0.69 to 0.86) but do not report values of internal consistency. While no conclusions can be drawn from such information, coming as it does from different studies, it lends support to the high face validity of the suggestion that preferences may be for content rather than process.

It is, of course, possible that in writing statements to illustrate different cognitive processes an individual test constructor may produce statements with a pattern that is perceived by the pupil as a pattern of content. For example, in Heath's original test items, statements constructed to demonstrate memory of specific facts or terms could be seen as statements containing definitions or formulae; the statements of practical applications information could be seen as containing references to technology encountered outside the science classroom; statements challenging or questioning the information could be seen as containing the sort of questions or comments that the teacher uses when she wants to lead the class on to the next section of work; statements of fundamental principle could be seen as containing the sort of conclusion that comes in the summary at the end of each chapter in the text book. Choices made in this way in terms of content could lead to high internal consistency and high stability values for M, A, Q and P scores, and there will be no way of determining from such statistical data whether the preference choice made was for processes or content.

There are innumerable ways that the pupil can classify content. Heath's statements could be classified into a) those that refer to science inside the classroom (mostly M, Q and P), b) those that refer to technology met outside the classroom (mostly A). If a pupil prefers the A alternative on the first item we would expect him to choose A on the second item and so on. This would lead to high values of internal consistency and stability for the A scale. However, no systematic distinction would be made by the pupil between M, Q and P, and we would expect a choice of M on the first item to be followed by M or Q or P on the second. On a later occasion the choice for the first item is just as likely to be Q or P as it is M. We would, therefore, expect low internal consistency and stability coefficients for the M, Q and P scales.

It is possible that pupils, in expressing preferences between statements, are attending to differences in linguistic features (formal characteristics of vocabulary and syntax) rather than differences in meaning. For example, a simple analysis of three cognitive preference tests (Table 10.1) shows that the occurrence of 'hypothetical words' ('if', 'would', *etc.*) together with 'negatives' was 19 times as frequent in Q statements as in M, A and P together. Similarly, application words ('use', 'obtain', *etc.*) were 46 times as frequent in A statements as in M, Q and P together. The pupils may recognise such phraseology, and systematic choices might reflect characteristics of the language with which they are familiar. Choice made on this basis would be expected to lead to high levels of stability and, to a lesser extent, internal consistency on some but probably not all, of the scales.

TABLE 10.1

AN ANALYSIS OF THE 'QUESTIONING' AND 'APPLICATION' STATEMENTS OF THREE COGNITIVE PREFERENCE TESTS IN TERMS OF (A) HYPOTHETICAL AND NEGATIVE WORDS, AND (B) APPLICATION WORDS

	Hypothetical or negative words (if, would, whether etc. not, never, fail, etc.).		Application words (use, application, obtain, basis, etc.).	
	Q Statements	non-Q Statements	A Statements	non-A Statements
Health (1964) (20 items)	15	2	9	0
Kempa and Dubé (1973) (40 items)	29	1	31	1
Brown (1974) (10 items)	12	0	6	0

Preferences as conformity with external influences

It is also possible that the pupil may recognize some phraseology as characteristic of his science teacher and tend to choose alternatives containing these elements, seeing them as stamped with official approval. Similarly, preferences for content or cognitive processes may be directed by what the pupil perceives as approved by authority (*i.e.* the teacher, the text book or worksheets). The conscious desire to conform to the style of the authority may well take precedence over the pupil's personal value system (if, indeed, he has one). Such behaviour would not be classified as Valuing by Krathwohl *et al.* (1964) since there is no commitment by the individual to a personal underlying value. It might rather be classified as Willingness to Respond or Satisfaction in Response (the satisfaction being the anticipation of approval or high grading from the teacher). Sometimes this conforming behaviour may be brought about by the pupil actually adopting the attitude of his teacher in order to maintain a satisfactory relationship. Through this process of 'identification' (Kelman, 1958) he actually believes in the opinions he adopts, but they are tied to, and dependent upon, the external source and so are not part of his own value-system.

It would not be possible, from the preference test responses, to distinguish valuing behaviour from conscious or subconscious conforming behaviour.

Preferences as compliance with the demands of the task

It has been suggested that the pupil may show preference for particular statements because he believes that certain features are approved by the teacher. Whether he is looking for approval or not he is in a test situation and he must indicate some preference in order to comply with the demands of the tester. He can do this by responding at the lower level (Krathwohl *et al.* 1964) of behaviour of Acquiescence in Responding. This level implies, in the hierarchical system, that the pupil is aware that there are four choices but attaches no emotional significance or value to any of them. This would lead to random choices and no pattern of selection of M, A, Q or P by a pupil as he passes from item to item. The prediction would be for zero coefficients of internal consistency and stability (since no relationship can be assumed between choices on two occasions) for all scales.

The dependence of preference on the pupil's cognitive abilities

The distinction between the choice of preference at the Preference for a Value and at the Acquiescence in Responding levels is one of valuing. If the pupil exhibits reliable behaviour at the former level and is concerned with preferences for cognitive processes (as appears to be assumed by proponents of cognitive preference tests), then it seems reasonable to assume that the pupil's own cognitive ability must be such that he can operate at the level of cognitive process implied by his choice.

If the M alternative is 'valued' it would be difficult to argue that anything more than recognition of a restatement of the introduction has occurred. A preference shown, therefore, is for psychological process of recall at the Bloom *et al.* (1956 pages 65-67) level of Knowledge of Specific Facts. It may be argued that each of the preference alternatives can be related to a cognitive behaviour corresponding to a particular level of educational objectives as proposed by the Taxonomy.

If the P alternative is chosen we could suggest that while it is essentially a restatement of the introductory sentence, it has been put on a more generalized and abstract form. As such, it could be said that for it to be 'valued' by the pupil in this context, he must have sufficient comprehension of the statement to translate it from its level of abstraction to the more concrete level of the introductory statement. The suggestion is then, that this choice may represent a preference for a level of Comprehension - Translation (Bloom, pages 91-93). The alternative level for this choice might be Knowledge of Principles and Generalizations (Bloom, pages 75-76) on the assumption that the pupil has merely recognized the particular principle.

It is likely that for the A alternative to be 'valued' by the pupil he would need to appreciate the connection between it and the introductory statement. This could require him to recognize a particular case, comprehend a generalization, and then apply the generalization in a new situation. The choice would then correspond to Bloom's Application level (Bloom, pages 120-130). However, it is also possible that the A statement has been recognized as another exemplification, already familiar to the pupil, of the introductory statement and so would be categorized as Knowledge of Specific Facts.

The Q response involves knowledge other than that supplied by the introductory statement. The 'valuing' of such a response might, therefore, be dependent on the ability of the pupil to analyse relationships between these different knowledge elements, and so to expand, develop or criticise the introductory statement. This would lead to a categorization at the level Analysis of Relationships (Bloom, pages 146-147). For some Q responses it can be argued that no more than the level of Comprehension of the situation described by the introductory statements is required.

It is, therefore, suggested that the cognitive preferences M, P, A and Q represent (at the highest estimate) preferences for statements implying cognitive behaviour at the level of Knowledge, Comprehension, Application and Analysis. This is a much more specific and narrow view of cognitive preference than that of those who make the identification with cognitive style.

It is further hypothesised that if the sample is homogeneous and all the pupils are able to operate at the level of Analysis, then we would expect equally high reliability on all scales since pupils would have no cognitive restraints on their preference. If, however, the sample is heterogeneous and contains sub-groups of pupils for whom the highest level of cognitive functioning is Knowledge, Comprehension, Application and Analysis respectively, then we would expect the scores on the M, P, A and Q scales to become less reliable as their implied cognitive processes become more complex. The pupil whose highest level of operation is Knowledge can only make consistent choices with respect to M statements. With statements corresponding to higher cognitive levels he will respond in an Acquiescence in Responding mode rather than a Preference for a Value mode leading, as discussed earlier, to zero coefficients of reliability. For pupils operating at the Comprehension level we would expect consistent M and P scores, and for these at the Application level consistent M, P and A scores. These differential effects would probably be most marked with groups of young or low ability pupils.

If the pupil is valuing content rather than process then it is not so clear that his cognitive ability must 'match' the stimulus statement. It may be that familiarity with concepts is all that is needed to make choices on the basis of content. However, this will depend on what his classification system consists of, and it seems likely that ability to operate at the appropriate cognitive level would lead to a greater feeling of security and more consistent choices.

If the pupil's choice is concerned only with vocabulary or syntax, it seems unlikely that any cognitive ability beyond that of Knowledge (recognition) is involved.

The Analysis of Cognitive Preference Data

Data collected in the form of four preference scores appears appropriate for Heath's straightforward purpose of investigating whether or not pupils following the new PSSC Course achieved its distinctive objectives at a higher level than did those pupils in a Control group. However, the data is ipsative in nature and this means that an individual's scores on the four scales are interdependent. Hicks (1970) has reviewed some of the properties of, and problems associated with, ipsative measurement. He questions the interpretations that have been put on such measures by researchers. He sees valid interpretations as being, necessarily, complex (and possibly relatively meaningless). Thus, if we are to comment on the differences in M, A, Q and P scores between pupils it would have to take a form such as:

> 'this pupil is higher on preference for Questioning relative to his preferences for Principles, Memory and Application, than are other pupils on preference for

Questioning relative to their preferences for Principles, Memory and Application'.

It may be that the absolute level of that pupil's preference for Questioning is, in fact, lower than that of all other pupils. The ipsative score can, however, neither support nor contradict a level of preference for Questioning determined by some other means.

If the data is to be used for more complex processes, such as the description of patterns of preference, severe constraints on the ways in which it may be handled are imposed by the ipsative nature.

If we consider M, A, Q and P as stimulus statements illiciting preference choice responses, then it is tempting to look at the 'between-stimulus' correlations for patterns of preference, and then to proceed to a R-type factor analysis of the responses of individuals on four variables (stimuli). However, such techniques are inappropriate for use with ipsative data (Cattell 1957, pages 492-499). Even a simple correlation matrix of this type is very difficult to interpret. As soon as preference for one stimulus is expressed the other three stimuli scores are automatically depressed, leading to negative correlations. If the distribution of scores over all areas is random then correlation coefficients of -0.33 would be expected as pointed out by Kempa and Dubé (1973). (In general the average inter-correlation will have a value of $\frac{-1}{(m-1)}$ where m is the number of variables in the ipsative test: Hicks, 1970.)

These authors have used such correlations and after stating that the normal procedure for estimating the significance of correlation coefficients cannot be applied, they proceed to conclude that the correlation between

M and A (-0.167), A and Q (-0.221), and Q and P (+0.238) 'are weak and suggest no cross-influences to exist between these areas'. They see the correlation between M and P (-0.436) as somewhat stronger but nowhere near the high level of negative correlations between M and Q (-0.714) or A and P (-0.645). They give no details of the basis on which they make these judgments. They may be using the value of r = -0.33 as 'base-line' for zero relationship between the scales. If so, then the positive relationship between Q and P could be of higher significance than the negative relationships between M and P or A and P. They quote significance levels of r-differences for $r_{MQ} - r_{MP}$ ($p<0.001$) and $r_{AP} - r_{AQ}$ ($p<0.001$), but not for $r_{PQ} - r_{PM}$.

A factor analysis was then carried out on their correlation matrix and two strong factors were extracted. One factor showed a high positive loading (0.969) on the M scale and a high negative loading (-0.847) on the Q scale, while the other factor had loadings of -0.954 and 0.824 on the A scale and P scale respectively. This leads to the neat-looking conclusion of cognitive preferences being oriented along two bi-polar independent axes.

Kempa and Dubé admit that their form of analysis is of doubtful validity in this context, but suggest that 'the strength of trends and correlations obtained by means of the conventional normative statistical procedure is such that the findings in the preceding section may be accepted without undue reservations'. Such a conclusion is highly questionable. The use of inappropriate techniques and an absence of clear statistical interpretation of both procedures and results, must lead to rejection of these conclusions however appealing they may appear.

While it may not be fruitful to consider the 'between-stimuli' correlations it should be meaningful to look at the 'between-individual' correlations. This would, hopefully, identify pupils with similar patterns of cognitive preferences. A Q-type factor analysis of the correlation matrix could carry this further and using factor loadings isolate groups of pupils characterized by particular patterns of preferences. Unfortunately in the context of cognitive preferences there are only four scores for each individual. The analysis would be analogous to carrying out a R-type factor analysis on test scores on a sample of only four individuals. Guilford (1954 page 533) observes that there are no 'rules' governing the adequacy of sample size since there are no known ways of estimating sampling fluctuations in rotated factor loadings. Errors in correlation coefficients will be reflected as errors in factor loadings and he suggests that if Pearson product moment correlations are used, a minimum of 200 for sample size is good policy. This minimum size is patently not adhered to by large numbers of researchers but sample populations as low as four do not appear in the literature. Childs (1970) suggests that the one per cent level of significance for loadings should be the criterion measure in view of the uncertainty surrounding the assessment of error in factorial work. For a sample population of four and a 'large' number of variables the loadings would have to be in excess of 0.917 to attain the one per cent level of significance (Burt and Banks, 1947).

Wish (1964) has employed unfolding analysis (Coombs, 1964) on Heath's data and has attempted to formulate models of the underlying structure of the four cognitive preferences. The multi-dimensional model of unfolding (Bennett and Hays, 1960; Hays and Bennett, 1961; Coombs, 1964 page 140) assumes that each individual and each stimulus can be represented by points in a common space. An individual's 'ideal point' in space represents his order of preference for the stimuli by the distances between that 'ideal point' and the stimuli (the distance being smallest for the most preferred stimulus and the largest for the least preferred stimulus).

Four stimuli can be ordered in 4! ways. If all these orderings occur then it is a mathematical necessity that the geometrical configuration of the four stimuli is one of four non-co-planar points on a surface of a sphere. However, if certain orderings are missing it may be that this limitation has been brought about by some underlying attribute which would restrict the four stimuli and the individuals to two dimensional space. There are a limited number of distinguishable geometrical configurations of four stimuli in two dimensional space, each requiring that 6 out of the possible 24 orderings are missing. McElwain and Keats (1961) have developed a simple method of determining the appropriate configuration from a given set of orderings.

The technique is intended to provide a near-perfect accounting for the data (since at worst 75% of the data can always be accounted for by 18 orderings). McElwain and Keats do not provide an explicit criterion for the amount of data not fitting the model that may be tolerated, but they describe an example of a non-perfect fit in which there are two discrepancies in 304 cases. Wish (1964) identified a model, locating M, A, Q and P as points in two dimensional space, that accounted for the preferences of 287 out of 324 individuals in a PSSC group, 291 out of 324 in a Control group, all the mean preference ratings of 30 PSSC classes, and 48 means out of those of 49 Control classes.

The results for individuals' preferences suggests that the criterion for a successful model implied by McElwain and Keats' example has not been achieved. A more satisfactory percentage is achieved for the group mean orderings, and Wish's conclusions are based on these. He finds that 20 per cent of PSSC group means and 0 per cent of Control group means are positioned closer to P and Q in space than M and A, while 33 per cent of PSSC and 56 per cent of Control groups are closer to M and A than to P and Q, 46 per cent of PSSC and 43 per cent of Control groups lie in an intermediate region. Arbitrarily assigning values of three, two and one to these three regions he correlates achievement with region number and finds that this correlation coefficient is higher for PSSC than for the Control group. He states:

> 'that the dimensions are related to achievement and curriculum in the expected direction simultaneously supports the PSSC curriculum claims, demonstrates the construct validity of the Cognitive Preference Test, and gives psychological meaning to the dimensions'.

His conclusions on the PSSC curriculum claims seem to add little to those reached by Heath using a much simpler analysis. Wish's concept of construct validity of the test may only be interpreted in the narrow sense that this test reflects certain objectives of the PSSC course. It has quite clearly not been shown as having construct validity for pupils' science cognitive preferences since no logical deduction or hypothesis has been presented to suggest that M, A, Q and P are a necessary or sufficient population of constructs of this purpose. There is no clear indication of what extra useful information the unfolding model supples. Consideration of group mean measures may be appropriate for questions of achievement of course objectives by teaching groups, they are not, however, suitable for the investigation of attributes (*i.e.* cognitive preferences) of individual pupils. What significance can be attached to the identification of a point in space as 'critical questioning'? What is the 'psychological meaning' of the dimensions that Wish talks of? Mackay (1972) has interpreted these dimensions as 'P - M' and 'Q - A' and uses two compound preference scores only. It seems reasonable to say that an individual located between M and P shows an increase in his P - M scores as he moves towards P. But P is a point, and when P is passed the P - M score will, thereafter, be a constant. There are, therefore, constraints on the extent to which the P/M and Q/A dimensions can be treated as axes for the Cartesian co-ordinates for individual preferences.

The considerable indeterminancy of the unfolding model (the analysis of which is based on the M, A, Q and P constructs, themselves of highly questionable descriptive value) suggests that there may be little point in pursuing this line of analysis further.

While the analysis of variance approach (Kempa and Dubé, 1973, Barnett, 1974) to the relationship between cognitive preferences and achievement in science has been appropriate for identification of significant differences between groups, it is not so clear that Barnett's use of analysis of covariance is justified.

He uses pretest scores of achievement and cognitive preferences as covariates in his analysis. The 'treatments' in this context are the pupil's perceptions of his teacher as 'directive' or 'non-directive'. The only situations in which analysis of covariance can be used without any reservations are those in which individual subjects are assigned randomly to treatment groups. In Barnett's study intact groups of subjects (biology classes) were exposed to the treatments, but the treatments were not randomly assigned to these groups. It might be marginally justifiable to use analysis of covariance if a) the treatment does not affect the covariate (*i.e.* the covariate is measured before application of the treatment, or the nature of the variables preclude this effect); and b) the covariate does not affect the treatment. (Evans and Anastasio, 1968)

We are not told if the pupils had received any instruction from these teachers prior to administration of the pretests. If they had, then the treatments would almost certainly have influenced the pretest scores (covariates). Barnett himself suggests that a 'non-directive' teaching style 'provides an informal atmosphere where student planning and questioning are encouraged'.

If, on the other hand, the pretest measures were made before the pupils came in contact with the teachers, it is still possible that the covariates may have influenced the teachers' behaviour (and so the treatments). Teachers may modify their strategies according to the abilities of their pupils - 'discovery' methods for high achievers with a critical questioning approach, but 'chalk,talk and recipe experiments' for low achievers with a preference for factual recall.

It seems unlikely that we can assume the statistical independence of covariates and treatments that is necessary for analysis of covariance to be an acceptable technique for use in this situation.

The Cognitive Preferences of S2 Pupils in Scottish Secondary Schools

This chapter has pointed to some problems associated with the nature and analysis of cognitive preferences. Data from the empirical work, designed to investigate cognitive preferences in science among the sample of S2 pupils, was also used to throw some light on some of the difficulties that had been identified.

The particular aims of the empirical study were:

1. To determine whether or not pupils following the Integrated Science course (a 'new' curriculum) have different cognitive preferences in science from those following a separate science subjects course (a more 'traditional' curriculum).
2. To discover if thirteen to fourteen year-old pupils exhibit consistent and stable cognitive preferences.
3. To provide information relevant to the hypothesis that within this age-group 'Preference for a Value' behaviour will tend to be replaced by 'Acquiescence in Responding' behaviour as one moves through M, P, A to Q (identified in the hypothesis with the cognitive levels of Knowledge, Comprehension, Application and Analysis) thus making scores progressively less reliable.

Development of the instrument and the pilot study

The questionnaire was compiled from items developed by Dr. L.D.M. Mackay, Monash University and Dr. R. Kempa, University of East Anglia for other groups of pupils. The pilot instrument was made up of twenty items. The pupil was asked to endorse the statement he found most appealing (this added +1 to the appropriate scale), and least appealing (this added -1 to the appropriate scale). The pattern for each item is thus forced into a = 1, -1, 0, 0 form. The twenty items were selected for pilot work on the basis of the following:

a) they had passed scrutiny by judges in Dr. Mackay's and Dr. Kempa's studies;

b) the information in them was covered in the Integrated Science syllabus of Curriculum Paper 7;

c) the vocabulary used appeared in either the worksheets (Heinemann 1970) or the text-book (Science for the Seventies: Mee, Boyd, Ritchie 1970) designed for this Integrated Science course;

d) three principal teachers in a comprehensive school following the Integrated Science course tried all items out informally with second year classes and found them to be acceptable to the pupils.

These twenty items were given to 141 second year pupils in another comprehensive school following the Integrated Science course. M, A, Q and P scores on each item were correlated with total M, A, Q and P scores for the test. For an item to be included among the ten items in the final version of the test, it was necessary that M, A, Q and P scores on that item correlated positively with the total M, A, Q and P respectively. This was the case in only ten of the items. Fortunately these ten items were satisfactorily distributed with regard to subject bias (two biology, three chemistry, three physics, one chemistry with biology, one physics with chemistry) since the test was to be given to 'separate subject course' pupils as well as 'integrated science' pupils (see Appendix A-23A for the final list of items and instructions to pupils). A retest was carried out with 132 of the pupils six weeks later using the ten items seen, at this point, to be working satisfactorily. From here on calculations are based on the scores of these 132 pupils on the ten items only. The group means of the test and retest scores each showed a preference ordering of PMAQ. There were no significant differences (as measured by a t-test) between means of the same category on test and retest. The coefficients of internal consistency in each case descended from M, P, A to Q (0.57, 0.49, 0.43, 0.28 and 0.59, 0.51, 0.43, 0.30) as measured by Cronbach α coefficient. Similarly the test-retest reliability coefficients followed the same pattern (0.51, 0.49, 0.35 and 0.25). Since an entire year group of thirteen year-olds had been tested, it was thought that those of lower ability might not have understood what was required of them. Twenty-nine of the pupils selected to take Latin were assumed to be of higher than average ability and had their test and retest scores examined. The reliability followed the same pattern (0.57, 0.53, 0.54, 0.28) except that the A scale seemed rather more reliable.

The sample

One school from each of the medium sized cells of the main sample of schools was selected at random (*i.e.* 3 following Integrated Science and 3 following separate science subjects).

Analyses were carried out for all (n = 445) pupils, for pupils following Integrated Science courses (n = 218), for pupils following separate science subject courses (n = 227) and for pupils in each individual school (n = 67, 68, 70, 77, 81, 82).

Results and discussion

The means of the 'integrated science' and 'separate science subjects' groups each showed a preference ordering of PMAQ. All individual school means showed the P and M preferences in first or second position, with Q and A in third or fourth position.

Cronbach α measures of internal consistency (Table 10.2) showed a decrease in value from M to P to A to Q for both groups, but were higher for the 'separate science' group on all scales. In four out of the six schools the Cronbach α value for Q was in fact negative.

Unfolding analysis was able to account for a maximum of 88% of the individual scores with a two-dimensional model where P is contrasted with M on one dimension and Q with A on the other. This, it is suggested, is too low a percentage of scores for the model to represent a useful description of cognitive preference structure despite the fact that 100% of the individual school mean preference orderings were accounted for by this model.

The inter-stimulus correlations (Table 10.3) show that the patterns for the 'integrated science' population and 'separate science' population are similar to that of the total population (there were minor discrepancies, however, in some of the individual schools).

There was a strong similarity in the proportions of pupils from the different groups choosing particular preference orderings. The correlation between the proportion of 'integrated science' pupils with a particular preference ordering and the proportion of 'separate science' pupils with same preference ordering was found to have a coefficient of $r = 0.89$ over the twenty-four possible choices.

The clear similarities between mean preference orderings, between distributions of individuals' scores among different orderings, between inter-stimulus correlation matrices and between patterns of reliability coefficients of pupils following the two different courses, lead to the conclusion that there is no discernable difference between their preferences.

TABLE 10.2

MEANS WITH STANDARD DEVIATIONS AND COEFFICIENTS OF RELIABILITY (CRONBACH α) FOR COGNITIVE PREFERENCE SCORES

Cognitive Preferences	n = 445 Total Sample		n = 218 Integrated Science Sample		n = 227 Separate Science Sample	
Memory	10.58 (2.80)	r = 0.42	10.45 (2.75)	r = 0.39	10.70 (2.85)	r = 0.45
Applications	9.38 (2.46)	r = 0.23	9.52 (2.41)	r = 0.19	9.23 (2.51)	r = 0.26
Questioning	9.19 (2.26)	r = 0.13	9.32 (2.17)	r = 0.02	9.06 (2.34)	r = 0.22
Principles	10.86 (2.51)	r = 0.27	10.71 (2.47)	r = 0.24	11.01 (2.54)	r = 0.30

TABLE 10.3

INTER-CORRELATIONS OF COGNITIVE PREFERENCES

	Total Sample n = 445				Integrated Science n = 218				Separate Science n = 227			
	M	A	Q	P	M	A	Q	P	M	A	Q	P
Memory	1.00				1.00				1.00			
Application	-0.50	1.00			-0.51	1.00			-0.50	1.00		
Questioning	-0.23	-0.26	1.00		-0.22	-0.24	1.00		-0.23	-0.28	1.00	
Principles	-0.42	-0.19	-0.39	1.00	-0.42	-0.19	-0.39	1.00	-0.42	-0.17	-0.39	1.00

(It should, however, be remembered that this is not a comparison between a random sample of pupils following an 'integrated science' course with a similar group following a 'separate science subjects' course. Any generalizations must be made in accordance with the sample of pupils being selected at random from within schools, which themselves form a stratified random sample of Scottish schools with a 'medium' size intake, stratification being by denomination and type of science course.)

All the coefficients of internal consistency are low (Table 10.2). For the Integrated Science group the M scale is the only one approaching a satisfactory level. It is therefore suggested that this group can behave in the Preference for Value mode at the level of Knowledge only. As we look at the internal consistencies of the P, A and Q scales (hypothesised as corresponding to the cognitive levels of Comprehension, Application and Analysis) the decreasing values indicate that the behaviour mode moves progressively into Acquiescence in Responding.

The separate science subjects group follows a similar pattern but with slightly higher coefficients. These suggest that Preference for a Value mode may extend to statements at the Comprehension (P) level.

The consistently higher values for Cronbach α for the 'separate science' group might suggest that a larger proportion of these pupils are able to operate at the levels of Comprehension, Application and Analysis relative to the 'integrated science' group. However, such a conclusion would be dangerous unless based on significance tests applied to the coefficients. Since α is 'an estimate of the correlation expected between two tests drawn at random from a pool of items like the items in this test' (Cronbach 1951), the coefficients were all transformed to Fisher's *z's*, and significance tests were applied to the differences between 'integrated science' and 'separate science' α's on each of the four scales. Only the Q scale showed significant differences between the α's ($p<0.05$). (It may be that this test is too stringent for testing differences between α's. Since α is the mean of all possible split-half coefficients for the test, it will not have the same sampling distribution as these coefficients but will be relatively 'stable'. Differences between α's on the M, P and A scales should therefore be described as 'small' rather than 'not significant'.)

The pattern of stability coefficients from the pilot study supports the suggestion that Preference for a Value mode is replaced by Acquiescence in Responding as one moves from M to P to A to Q. The increased value for the A scale for the higher ability group of the pilot study lends further support to this interpretation, a corollary of which is that more able pupils should exhibit Preference for Value behaviour at higher cognitive levels. Also consistent with this interpretation are the standard deviations found for the different scales, since we would expect standard deviations to decrease as the degree of randomness in pupil responses increases. In complete consistency with the hypothesis there is a decrease from M to P to A to Q in the values of internal consistency coefficients, stability coefficients and standard deviations.

All the coefficients of reliability have been low even for the M and P scales, and this may be due in part to the length of the test. This particular test is short and we can predict low reliability (Guilford 1954 page 353). If the test were doubled in length we would double the error variance but the true variance would be quadrupled so increasing the coefficient of reliability (except in the unlikely case of perfect reliability). If our reliability coefficient is 0.4 we would expect to achieve 0.7 for the reliability coefficient by increasing the length of the test by a factor n where n is given by:

$$n = \frac{0.7}{0.4} \times \frac{(1 - 0.4)}{(1 - 0.9)}$$

$$n = 10.5$$

(Guilford, 1954, page 391)

i.e. a test of about 100 items. These items would, of course, have to be homogeneous with the ones already used.

While such a test might, in theory, achieve adequate reliability, it would certainly be a formidable task for pupils of this age group.

Summary and Conclusions

It has been suggested, on the basis of a review of previous research and an analysis of the tasks with which pupils are presented in Cognitive Preference Tests:

i) that there is no evidence for the validity of Memory, Principles, Application and Questioning for describing pupils' preferences for cognitive objects in science, and that reported low levels for internal consistency cast doubt on their value;

ii) that claims, based on R-type factor analysis or unfolding analysis, that cognitive preferences assessed in terms of M, P, A and Q can adequately be described in terms of two dimensions have not been substantiated;

iii) that there is no justification for equating differences in expressed cognitive preferences with differences in cognitive style;

iv) that while pupils' expressed preferences may be for particular cognitive processes or activities, they may equally well be for particular kinds of content or for distinctive linguistic characteristics of statements; preferences of the latter two types would lead to lower internal consistency coefficients in some cases but not in others;

v) that no distinction can be made between preference choices made on the basis of conformity to external influences and those made according to a personal value system;

vi) that, insofar as pupils' choices represent preferences for cognitive activities, M, P, A and Q choices depend respectively upon the different levels of cognitive attainment of Knowledge, Comprehension, Application and Analysis; and that, with younger or less able groups of pupils, systematic 'Preference for a Value' choices will tend to be replaced by random 'Acquiescence in Responding' choices as we move through M, P, and A to Q.

The empirical findings of a study of the science cognitive preferences of 445 thirteen to fourteen year-old pupils provides support for (i) above in that the values obtained for internal consistency coefficients are uniformly low ($r = 0.02$ to 0.45).

Unfolding analysis of the data has shown that more than two dimensions are necessary to describe the preferences expressed in this investigation (*cf.* ii above).

The patterns of internal consistency and stability coefficients have proved to be consistent with the hypothesis of (vi) that predicts decreasing values for reliability as we move through M, P, and A to Q.

No significant differences were found between the cognitive preferences of pupils following an 'integrated science' course and those of pupils following a 'separate science subjects' course. However, the highly questionable nature of what is being assessed by these tests makes it dangerous to draw conclusions from these findings, and no inferences about the 'scientific attitudes' of the two groups can be made.

While consistently higher levels of internal consistency on all scales for the 'separate science' pupils over the 'integrated science' pupils might suggest that a higher proportion of 'separate science' pupils were able to operate at the higher *cognitive* levels, these differences were found to be unquestionably significant for the Questioning scale only.

In conclusion, it appears that while substantial results have been achieved from studies that have used cognitive preference tests to investigate the achievement of course objectives specifying cognitive preferences, and to explore the relationships between science achievement and cognitive preferences, these results have not been used to shed light on the nature of cognitive preferences.

It is suggested that until evidence is provided for:

a) the construct validity of M, A, Q and P as measures of pupils' cognitive preferences,
b) the relationships between cognitive preference constructs and cognitive activities or attitudes of pupils, and

c) which characteristics of the stimulus statement the pupils are expressing a preference for,

further studies using these tests can have little value.

CHAPTER 11

ANALYSIS OF SCORES ON 'ATTITUDES TO SCIENCE' SCALES AND OTHER VARIABLES

The sampling unit for the stratified random sample was the *school*. Mean scores for each school were, therefore, used as the units of analysis in the *between-schools* study. Since the effect of sex was also of interest, mean scores for boys and girls separately were usually employed.

In addition to the *between-schools* study of the 40 schools, a *within-schools* analysis was carried out in 12 schools. For this analysis the unit used was the score of the individual pupil.

Finally, an examination of scores *within-teaching-groups* was made for pupils from 38 teaching groups in 10 schools.

Between-Schools Analysis

There are two main purposes of this analysis:

a) to describe the variations in pupils' attitudes at the end of S2 among groups from different schools, and

b) to identify changes in attitudes that have occurred over the first two years of secondary school.

Attitudes at the end of S2

a) is concerned with asking the following questions:

i) To what extent do mean scores of pupils of different sexes, following different types of science course, in different types of schools, indicate significantly differing attitudes among groups of pupils at the end of S2?

ii) How much of the variance in these mean attitude scores can be predicted from sex, measures of social class of the school, pupil characteristics on entry to school (general intelligence, divergency, attitudes[1]), characteristics of the school (location, denomination, size), organization of science classes (single/mixed sex, mixed ability or not, number of periods of science, size of classes), and the type of science class followed (integrated/separate subjects)?

Question a) i) has been investigated by two 4-factor design analyses: the first analyses the mean post-test scores of pupils in medium sized schools following either the Integrated Science course or the separate science subject course, the second analyses the mean post-test scores of pupils in schools of various sizes all following the Integrated Science course. (As discussed under 'Sampling' there were insufficient 'large' and 'small' schools following a separate science subject course to maintain the full *size* x *science course* design.)

---

[1]To avoid confusion of attitude measures taken on *entry to secondary school*, and at the *end of S2*, the former will be referred to as *pretest* and the latter as *post-test*.

The first design is a four factor analysis by denomination (A), type of science course (B), sex (C) and school (S). There are three levels of denomination:

A_1 - Roman Catholic.

A_2 - non-denominational in the cities,

A_3 - non-denominational outwith the cities.

There are two types of science course:

B_1 - an Integrated Science course,

B_2 - a Separate Science subjects course.

Different levels of A and B contain scores of *different* schools. Different levels of C (C_1 - boys, C_2 - girls) correspond to the *same* schools. Therefore schools appear as an additional factor nested within A x B, and within each A x B cell the schools are crossed with C. The schools within each A x B cell are a random sample and thus constitute a random effect, the denomination, types of science course and sex effects are, however, fixed and are not subject to sampling.

The basic assumption made by this model is that the mean attitude score for boys or girls in a particular school can be regarded as a linear model made up of ten components. Thus if y_{ijkl} is the mean score in the l^{th} school, for the k^{th} sex, following the j^{th} type of science course, in the i^{th} denominational type of school, then:

$$y_{ijkl} = M + A_i + B_j + C_k + (A \times B)_{ij} + (A \times C)_{ik} + (B \times C)_{jk} + (A \times B \times C)_{ijk} + S_{ijl} + (S \times C)_{ijkl}$$

where M is a component common to all the scores;

A_i is a component common to all scores in denomination i of factor A;

B_j is a component common to all scores in type of course j of factor B;

C_k is a component common to all scores in sex k of factor C;

$(A \times B)_{ij}$ is a component resulting from the interaction of level i of factor A and level j of factor B;

$(A \times C)_{ik}$ is a component resulting from the interaction of level i of factor A and level k of factor C;

$(B \times C)_{jk}$ is a component resulting from the interaction of level j of factor B and level k of factor C;

$(A \times B \times C)_{ijk}$ is a component resulting from the interaction of level i of factor A, level j of factor B and level k of factor C;

S_{ijl} is a component common to all scores of school l of level i of factor A and level j of factor B;

$(S \times C)_{ijkl}$ is a component specific to school l (of level i of factor A and level j of factor B) interaction with sex k of factor C.

Table 11.1 shows how the variation is partitioned, Table 11.2 shows the components analysis of the four factors, one factor being nested and singly crossed.

The second design is similar to the first except that the type of *science course* factor (B) disappears (all pupils in this analysis follow the Integrated Science course) and is replaced by a *size of school* factor B'. The three levels of size are:

B'_1 - small schools (intake less than 150 pupils),

B'_2 - medium size schools (intake between 150 and 300),

B'_3 - large schools (intake greater than 300).

TABLE 11.1

BREAKDOWN OF VARIATION IN FOUR-FACTOR DESIGN

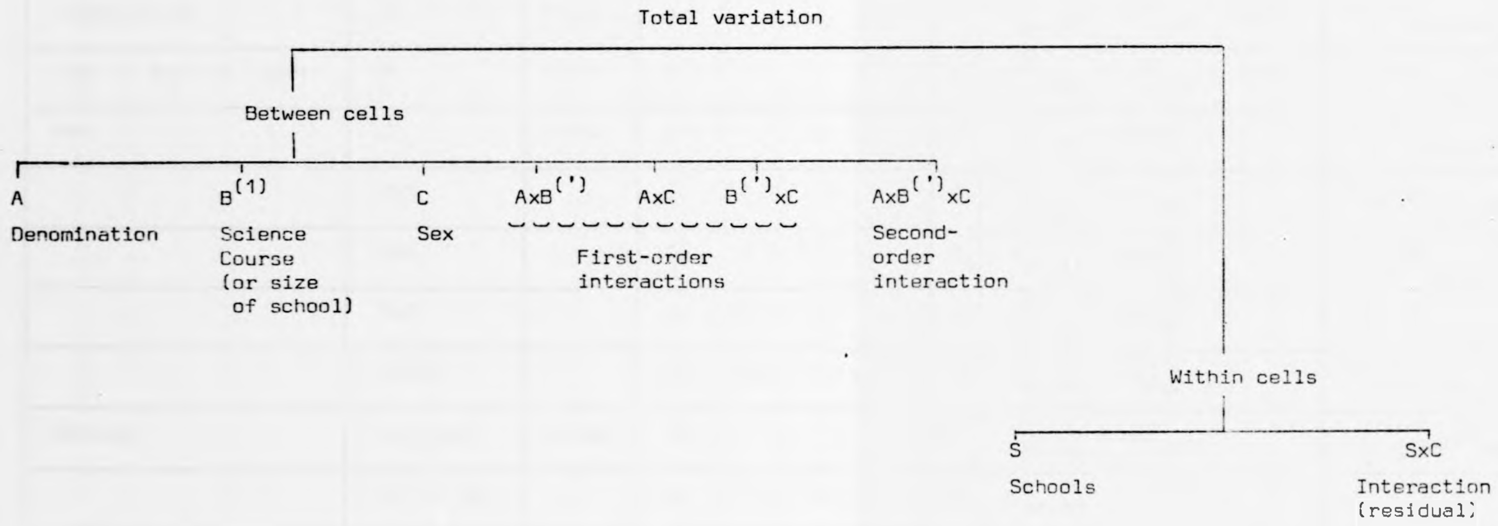


TABLE 11.2 FOUR-FACTOR COMPONENTS ANALYSIS, ONE FACTOR (S) BEING NESTED AND SINGLY CROSSED

SOURCE			DEGREES OF FREEDOM		MEAN SQUARE EXPECTATION	APPROPRIATE F-TEST
Denomination	A	fixed	a - 1	2	$\sigma^2_{sc.AB} + c\sigma^2_{s.AB} + sbc\sigma^2_A$	$MS_A/MS_{s\ w\ AB}$
Type of Science Course	B	fixed	b - 1	1	$\sigma^2_{sc.AB} + c\sigma^2_{s.AB} + sac\sigma^2_B$	$MS_B/MS_{s\ w\ AB}$
Sex	C	fixed	c - 1	1	$\sigma^2_{sc.AB} + sab\sigma^2_C$	$MS_C/MS_{sxc\ w\ AB}$
	AxB		(a - 1) (b - 1)	2	$\sigma^2_{sc.AB} + sc\sigma^2_{AB}$	$MS_{AB}/MS_{s\ w\ AB}$
	AxC		(a - 1) (c - 1)	2	$\sigma^2_{sc.AB} + sb\sigma^2_{AC}$	$MS_{AC}/MA_{sxc\ w\ AB}$
	BxC		(b - 1) (c - 1)	1	$\sigma^2_{sc.AB} + sa\sigma^2_{BC}$	$MS_{BC}/MS_{sxc\ w\ AB}$
	AxBxC		(a - 1)(b - 1)(c - 1)	2	$\sigma^2_{sc.AB} + s\sigma^2_{ABC}$	$MS_{ABC}/MS_{sxc\ w\ AB}$
Schools	S w AxB	random	ab (s - 1)	12	$\sigma^2_{sc.AB} + c\sigma^2_{s.AB}$	$MS_{s\ w\ AB}/MS_{sxc\ w\ x\ AxB}$
	SxC w AxB		ab (s - 1) (c - 1)	12	$\sigma^2_{sc.AB}$	
Total			abcs - 1	35		

The analyses were carried out using each of the five attitude measures in turn as the criterion variable. The computer program BMDX64 GENERAL LINEAR HYPOTHESIS (1972) developed at UCLA was used. The purpose of this program is to estimate parameters and test hypotheses concerning a general linear model. It automatically generates dummy variables specifying the analysis of variance and applies appropriate tests without additional hypotheses.

Certain assumptions about the experimental data must be made for the use of parametric statistics such as the F-test in this factorial analysis of variance (ANOVA) design. Gardner (Review of Educational Research, in print) reviews the literature concerned with whether or not the five requirements laid down by Siegel (1956 page 19) are in fact necessary. His conclusions are summarized as follows:

Assumption 1 (*i.e.* observations must be independent) is a necessary requirement of both parametric and non-parametric tests.

Assumptions 2 and 3 (*i.e.* observations must be drawn from normally distributed populations and these populations must have the same variance or a known ratio of variances) are not crucial in that the F-test and t-test are highly robust to non-normality and unequal variance of population measures. Provided the violations are not extreme, and several different assumptions are not violated at the one time, his conclusion (he cites fifteen studies for support) is that the F-test can withstand violations of these assumptions.

Assumption 4 (*i.e.* that variables involved must have been measured on at least an interval scale) has its validity challenged in two ways. Firstly, arguments that examine the mathematical assumptions underlying the parametric statistical procedures (he cites six studies) lead to the summary (Baker, Hardyck and Petrinovich, 1966 page 292) that:

> '...statistics apply to numbers rather than things...the formal properties of measurement scales, as such, should have no influence on the choice of statistics'.

Secondly, arguments on the basis that:

> '...if one alters the metric properties of scales and still reaches the same conclusions regardless of whether the data have been transformed or not, then for the purposes of statistical inference, it cannot matter very much what the scale properties are'.

He cites two studies in which ordinal data was transformed by various systems which demonstrated that, unless extreme departures from linear transformations were used, conclusions from parametric statistics were hardly affected.

Assumption 5 (*i.e.* the means of these normal and homoscedastic populations must be linear combinations of effects due to columns and/or rows) is, he suggests

> '...not considered necessary by some writers, and there has been much confusion concerning the interpretation of this assumption'.

In this study the method of sampling complied with Assumption 1. Inspection of the table of mean scores revealed no extreme departure from normality of distribution or lack of homogeneity of variances (Lewis 1968, page 45) relevant to Assumptions 2 and 3. The scale used was of the summated-ratings type. In the construction of this scale items were chosen to discriminate between favourable and unfavourable attitudes and to cover a limited range of popularity. Scores on the scale are the sum of a number of ratings (ordinal) and so we cannot claim that equal scale increments indicate equal increments in attitude. However, it is unlikely that the deviations from an interval scale will be extreme if care has been taken in construction, and in view of the robustness of the F-statistic with respect to violations of Assumption 4 its use with this data appears justified.

Question a) ii) has been investigated using a modified stepwise regression analysis on the mean scores, for boys and girls separately, from each of the forty schools in the study. The five attitude measures were taken in turn as criterion variable. Three groups of predictor variables were introduced into the regression equation.

Group A (in the order in which they were introduced):

i) Sex.

ii) Social class - two variables were used here a) the proportion of the school group in social classes 1 and 2 (Registrar General's Classification of Occupations 1970), b) proportion of the school group in social classes 4 and 5.

iii) General intelligence - mean scores on AH4.

iv) Five pretest attitude scores.

These variables are all characteristics that the pupils possess on entry to secondary school. The order in which they have been entered into the equation is determined by assuming that variables over which schools can be expected to have *least* control should be entered *first*. The stepwise procedure enables the *additional* variance accounted for by each predictor variable (*i.e.* the contribution over and above that made by predictors entered at an earlier stage) to be determined.

Group B

i) School location (city burgh/outwith city).

ii) Denomination (Roman Catholic/non-denominational).

iii) Size of school (large, medium or small intake, two variables were used: 'large/not large', 'small/not small').

This group represents characteristics of the school. The order is again determined by the control that educators can be expected to have over each characteristic. The location of the school is largely determined by population densities; the provision of schools of various denominations may be partly determined by a Local Education Authority policy but such decisions are subject to considerable external pressures; the size of the school is, however, mainly the choice of the Education Authority.

Group C

i) Type of science course that the pupils follow.

Groups A and B are concerned with variables that may influence the pupil's attitudes to science but which are independent of what he experiences in the science classroom. Group C introduces a predictor which enables us to ask whether the course directed specifically towards the attitude objectives is significantly more successful in achieving them than is an alternative course.

A further three predictors were added to Group C and the regression analysis was repeated using 38 of the schools (essential information was missing from two schools). The additional variables were:

Group C (continued)

ii) Mixed ability science teaching groups for S1 or not (less than 20% of pupils extracted).

iii) Mixed sex teaching groups for S1 or not (less than 20% of classes single sex).

iv) The number of periods of science allocated to S1 (arbitrarily coded '1' for 4 periods or less, '2' for 5 periods, '3' for 6 or more periods).

v) Number of pupils in science teaching groups in S1 (arbitrarily coded '1' for less than 15, '2' for 15 to 20, '3' for 20 to 25, '4' for greater than 25).

It was found to be impractical to determine divergency scores for all 2815 pupils in the study. Therefore, in order to investigate the predictions of attitudes by divergency variables a random sample of the 40 schools was taken, such that the sample contained a) a minimum of 20 schools and b) at least one school from each of the 12 'denomination x size' or 'denomination x science course' cells. The resulting sample contained 23 schools. Scores of 'fluency' on each of the three tests of divergency were determined for each pupil in these schools. Mean scores for boys and girls separately were calculated for each school.

The first problem was to ascertain that the patterns of variance for the full group of 40 schools and for the sub-set of 23 schools were similar. Initially the same predictor variables were introduced into the multiple regression equation as for the analysis of the 40 schools. The two sets of regression equations were then compared.

In general the pattern of the regressions indicated that more variance was accounted for in the smaller sample. This is to be expected since chance correlations are more likely to occur in small samples, and the stepwise technique for the multiple regression makes these chance effects cumulative. In each case multiple R's were converted to Fisher z's and confidence limits were determined (z is normally distributed while R is not). The region extending to two standard deviations on either side of the observed R (*i.e.* 95% confidence limits) was investigated for the 23-school regression and the 40-school regression so that the overlap for the two regressions could be assessed. It was not possible to do tests for significant differences between these two samples since one sample was a sub-set of the other. Tables 13D and 14D Appendix D show the range between $R \pm 2\sigma$ for each of the samples, considerable overlap in all cases is indicated.

Following this, divergency measures were introduced at two points in the regression equation and their contributions were assessed, firstly after measures of general intelligence and before pretest attitude measures, secondly after pretest attitude measures but before measures of school characteristics.

This multiple regression approach has had the advantage of being able to make use of all, or of a sub-set, of the schools in the sample. The confounding of denomination with location in the basic design has been partly disentangled. Multiple regression does not have the severe problems of analysis of variance with unequal cell sizes, and consequently estimates of the *additional* contributions from a much larger range of predictor variables is possible (Kerlinger and Pedhazur 1973 page 193). However, there are substantial correlations between the predictor variables (inevitable in non-experimental research), and no randomization of 'subjects among treatments' has been carried out. Consequently the influence of *independent* main effects and interactions cannot be determined. The *'a priori'* approach used here enters the predictor variables in some order of preference deemed appropriate (*e.g.* taking account of which variables come earlier in time, or of which one is believed to have 'caused' the other). Thus *additional* but not *unique* or *common* variance is accounted for.

Changes in attitudes over the two year period

So far the concern has been with a description of attitudes displayed at the end of S2 by different groups of pupils. These groups did not display neutral attitudes to science on *entry* to secondary school and the second important concern (see 'b' page 332) is for the *changes* in attitude that have occurred over the two years.

Three approaches to this problem were considered:

1. Analysis of *gain* scores on attitude scales over the two year period.
2. Analysis of covariance with post-test attitudes as criterion variables and pretest attitudes, general intelligence and social class as covariates.
3. Analysis of variance using a five-factor design across two occasions.

1. The 'Gain Score' is an attractive concept. The idea of estimating the

problems of analysis of variance with unequal cell sizes, and consequently estimates of the *additional* contributions from a much larger range of predictor variables is possible (Kerlinger and Pedhazur 1973 page 193). However, there are substantial correlations between the predictor variables (inevitable in non-experimental research), and no randomization of 'subjects among treatments' has been carried out. Consequently the influence of *independent* main effects and interactions cannot be determined. The *'a priori'* approach used here enters the predictor variables in some order of preference deemed appropriate (*e.g.* taking account of which variables come earlier in time, or of which one is believed to have 'caused' the other). Thus *additional* but not *unique* or *common* variance is accounted for.

Changes in attitudes over the two year period

So far the concern has been with a description of attitudes displayed at the end of S2 by different groups of pupils. These groups did not display neutral attitudes to science on *entry* to secondary school and the second important concern (see 'b' page 332) is for the *changes* in attitude that have occurred over the two years.

Three approaches to this problem were considered:

1. Analysis of *gain* scores on attitude scales over the two year period.
2. Analysis of covariance with post-test attitudes as criterion variables and pretest attitudes, general intelligence and social class as covariates.
3. Analysis of variance using a five-factor design across two occasions.

1. The 'Gain Score' is an attractive concept. The idea of estimating the

improvements in the attitudes of pupils, under the influence of their science teacher, by looking at their changes in scores is appealing. However, comparing change scores involves comparing changes at different points on an attitude scale *i.e.* we shall be comparing the change in attitudes of pupils who started with very favourable attitudes (very little room for improvement) with those who started with very poor attitudes (considerable potential for improvement). Frequently change in attitude is negatively correlated with initial attitude (advantage for those who initially score low), occasionally the correlation is positive (advantage for those who initially score high). This would be avoided if there were no significant differences between pupils at the start, but then we might as well work with final attitudes rather than gains.

Several authors have emphasised the unsuitability of raw gain scores for the measurement of changes in abilities or attitudes. Some recommend replacement of raw gain by 'true' gain (Lord 1956, 1958, 1963, McNemar 1958). 'True' gain is the difference between the 'true' post-test score (assumed to be defined by 'expected' value or mean over a number of independent observations of the same person) and the 'true' pretest score.

Other authors (Dubois 1957, Tucker, Damariu and Messick 1966) are concerned with using 'residual gain' scores or 'true residual gain' rather than raw gain. A 'residual gain' is determined by expressing the post-test score as a deviation from the regression line of post-test-on-pretest scores. Cronbach and Furby (1970) make the point that fallacious conclusions arise from consideration of raw gains primarily because of the systematic relationship between these scores and random errors of measurement. They discuss the ways of estimating 'true' change and 'true' residual change scores which they suggest are essential if the researcher feels he *must* use difference scores. However, after considering the various purposes for which gains are used they conclude:

> 'It appears that investigators who ask questions regarding gain scores would ordinarily be better advised to frame their questions in other ways'.

These authors together with O'Conner (1972) recommend the use of partial correlation and multiple regression, with the final status as the criterion variable and initial status as the covariate.

O'Conner (1972) suggests that gain scores will either give the same results as the multiple regression approach, or results that are more difficult to interpret. For example he shows that the correlation between initial status (X) and gain (G) will be positive, zero, or negative according to whether

$R_{XY}\frac{\sigma_Y}{\sigma_X}$ is >1, = 1, or <1

(R_{XY} = correlation between initial and final status)

He states:

> 'The difference between a positive and negative initial-gain correlation seems more interesting than the difference between a $R_{XY}\frac{\sigma_Y}{\sigma_X}$ of 1.05 and a $R_{XY}\frac{\sigma_Y}{\sigma_X}$ of 0.95; yet both the initial-gain correlation and $R_{XY}\frac{\sigma_Y}{\sigma_X}$ are determined by the same data. The distinction between a positive and negative initial-gain correlation appears to be artificial and misleading' (page 75).

He also shows that the relationship between gain and some other variable (W) can be most simply expressed in terms of $R_{YW.X}$ (partial correlation of final status with variable W while the initial status is held constant), so there is no need to compute change scores.

In view of these arguments analysis of gain scores was not attempted.

2. Analysis of covariance (ANCOVA) provides a method of *statistical* control for one or more independent variables which it has not been possible to control by means of the experimental design (Lindquist 1956, Chapter 14, Ferguson 1966 Chapter 20, Elashoff 1969, Winer 1962, Lewis 1968). In this study it was expected that criterion measures of post-test attitudes would be substantially influenced by the attitudes of pupils on entry to secondary school (this is born out by the intercorrelations). In addition it was hypothesized that levels of general intelligence and social class would also be influential. It might be possible to 'adjust' the criterion attitude scores (y) for these independent variable measures (x, x', x",...) using ANCOVA *provided* the necessary assumptions can be met.

Under those conditions the *adjusted* criterion measure y'_{ijkl} is given by

$$y'_{ijkl} = y_{ijkl} + \alpha_1(x_{ijkl} - \bar{x}) + \alpha_2(x'_{ijkl} - \bar{x}') + \ldots$$

where y_{ijkl} is the *unadjusted* criterion measure (see ANOVA page 334)

$\alpha_1(x_{ijkl} - \bar{x})$ is the variability due to the linear regression of y on x (similarly for x', x'', *etc*)

x, x', x'' are the independent variables, or covariates to be controlled.

Use of this statistical model provides an extension to the method of analysis of variance which allows us to estimate the significance of differences that there *would have been* among mean attitude scores of different groups, *if* those groups had all had the *same* level of pretest attitudes, general intelligence and social class.

Anderson (1963) observes that 'one may well wonder what exactly it means to ask what the data would be like if they weren't what they are'. Since we shall never have a situation where pupil groups enter secondary schools with the same attitudes, intelligence and social class levels, the results of the ANCOVA may well be of limited use only.

There are a number of assumptions (in addition to those for ANOVA) that must be justified if ANCOVA is to be a valid technique (Elashoff 1969 page 385, Lindquist 1956, page 323). Two of the most crucial of these are that the covariates must be independent of treatment effects, and that assignment of subjects to groups must be random. Evans and Anastasio (1968) suggest violation of either of these would result in linear correlation between treatment and covariate. Further, they comment that

> '...when treatment and covariate are inherently related, there is likely to be a strong linear correlation between treatment effects and covariate means. This circumstance ...makes the assumption of homogeneity of between-group and within-group regression quite generally untenable' (page 228).

(Homogeneity of regression is a third condition to be met for ANCOVA to be valid).

In order to test for the presence of any such correlations between potential covariates and 'treatments', 4-factor design analyses of variance (identical with those described earlier for post-test attitudes) were carried out using in turn five pretest attitude mean scores and mean AH4 scores as criterion variables. A 3-factor design (omitting sex) was used for the two school social class variables. Tables 1D to 12D Appendix D show the ANCOVA tables and tables of scores for each of the covariates, and the significant effects are summarized in Table 11.3. There is a problem of deciding what level of significance provides an appropriate cut-off point to avoid a Type 2 error *i.e.* retaining a null hypothesis when it is false.

Lindquist (1956, page 66) comments that the relative frequency with which a false null hypothesis is retained depends on *how far* it is from being true. For example, if the null hypothesis is *almost* true then the false null hypothesis would be retained *almost* 95% of the times it is tested. On the other hand, if the null hypothesis represents a considerable departure from the truth then it will very *seldom* be retained and the chances of a Type 2 error will be small. He suggests that if the consequences of a Type 2 error are likely to be serious, then a low level (*e.g.* 20%) of significance should be set.

For this study an arbitrary choice of the 10% level of significance was set for the covariate analyses (the 5% level was the lowest used in the attitude analyses). As Table 11.3 shows, there was at least one significant effect for all but one of the covariate analyses, indicating 'treatment' - covariate correlations and so violating conditions for the use of ANCOVA.

3. Analysis of variance with the four factors of the previous design crossed with occasions attempts to investigate the questions:

TABLE 11.3

SUMMARY OF SIGNIFICANT EFFECTS (10% or better) IN ANALYSIS OF VARIANCE, FOUR-FACTOR DESIGN USING POTENTIAL COVARIATES AS CRITERIAN VARIABLES

'Treatment' / Covariate	Denomination A	Science Course B	Sex C	AxB	AxC	BxC	AxBxC	Schools SW(AxB)
Pretest Attitude 1					0.025			0.001
Pretest Attitude 2			0.001			0.01		0.001
Pretest Attitude 3								0.01
Pretest Attitude 4	0.025		0.01					0.001
Pretest Attitude 5		0.025	0.10		0.10			0.005
AH 4								0.001
Social Class (1)	0.005							
Social Class (2)	0.01	0.10						
	A	Size of School B'	C	AxB'	AxC	B'xC	AxB'xC	SW(AxB')
Pretest Attitude 1						0.10		0.10
Pretest Attitude 2						0.025		0.025
Pretest Attitude 3				0.025				
Pretest Attitude 4			0.001					0.01
Pretest Attitude 5			0.10					0.005
AH 4	0.05							0.05
Social Class (1)	0.01	0.10						
Social Class (2)								

i) What *changes* in significance of effects (sex, type of school, type of science course) on pupils' mean attitude scores have occurred over two years?

ii) Which significant effects observed on entry to school, are still significant at the end of two years?

The linear model assumes that y_{ijkml} (the mean score for one sex in the l^{th} school) has twenty components, ten of which are defined as for those of the four-factor design and, in addition the following:

D_m is a component common to all scores on the m^{th} occasion;

$(A \times D)_{im}$ is a component resulting from the interaction of the i^{th} level of factor A and occasion m;

$(B \times D)_{jm}$ is a component resulting from the interaction of type of science course j of factor B and occasion m;

$(C \times D)_{km}$ is a component resulting from the interaction of sex k of factor C and occasion m;

$(A \times B \times D)_{ijm}$ is a component resulting from the interaction of the i^{th} level of factor A, j^{th} level of factor B and m^{th} level of factor D;

$(A \times C \times D)_{ikm}$ is a component resulting from the interaction of the i^{th} level of factor A, k^{th} level of factor C, and m^{th} level of factor D;

$(B \times C \times D)_{jkm}$ is a component resulting from the interaction of the j^{th} level of factor B, the k^{th} level of factor C and the m^{th} level of factor D;

$(A \times B \times C \times D)_{ijkm}$ is a component resulting from the interaction of the i^{th} level of factor A, j^{th} level of factor B, k^{th} level of factor C and the m^{th} level of factor D;

$(S \times D)_{ijlm}$ is a component specific to school l (of level i of factor A and level j of factor B) interaction with occasion m of factor D;

$(S \times C \times D)_{ijklm}$ is a component specific to school l (of level i of factor A and level j of factor B) interaction with sex k of factor C and occasion m of factor D.

Table 11.4 shows the partitioning of the variation and Table 11.5 the components analysis of the five factors, the school factor being nested and doubly crossed. It can be seen that different levels of A and B contain scores from *different* schools, but different occasions (D) and different sexes (C) contain scores from the *same* schools *i.e.* within each AB cell scores from the same schools appear in all cross classifications of sex and occasions. Therefore, schools are nested within each AB cell, but within each AB cell are crossed with occasions and sex.

As before the computer program BMDX64 GENERAL LINEAR HYPOTHESIS (1972) was used. However, this program imposes the limit of 90 dummy variables on this 5-factor design,which is adequate for the analysis of medium schools following two types of science course (72 dummy variables generated) but not for the analysis of various sized schools following Integrated Science (108 dummy variables generated). Since no other suitable program was readily available and it was impractical to carry out all the computations by hand it was decided that a series of 4-factor design analyses would be carried out and that inferences about missing higher order interactions would be made from lower order interactions.

TABLE 11.4

BREAKDOWN OF VARIATION IN FIVE-FACTOR DESIGN

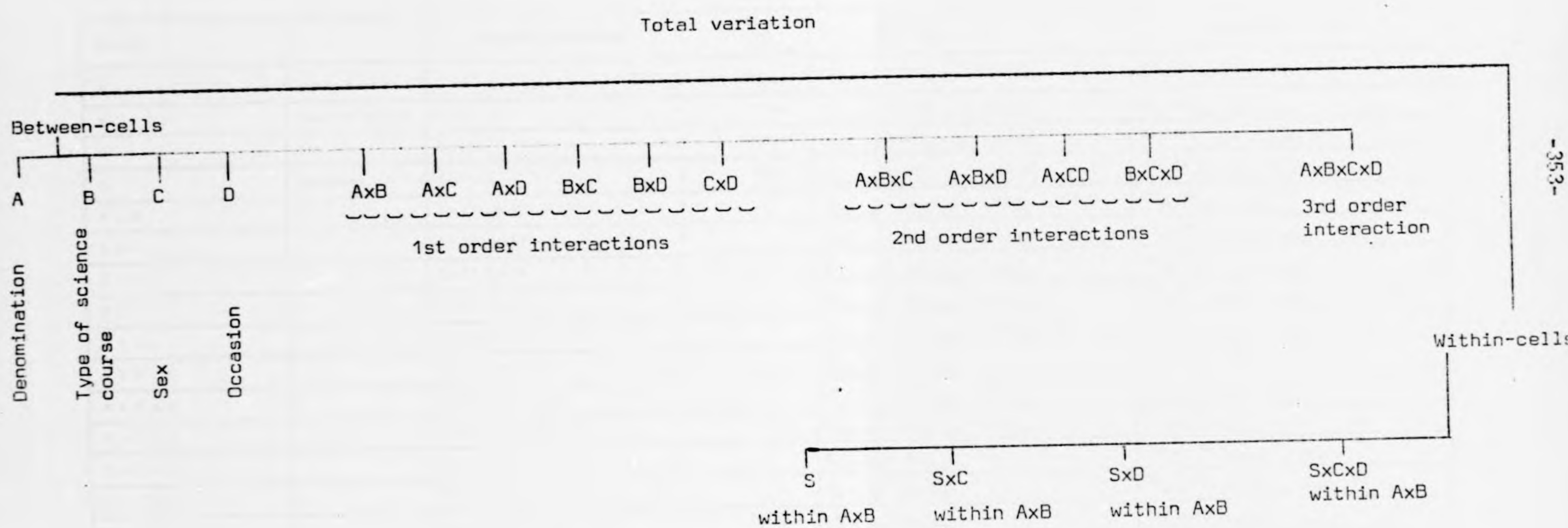


TABLE 11.5

FIVE FACTOR COMPONENTS ANALYSIS WITH ONE FACTOR (SCHOOLS) NESTED AND DOUBLY CROSSED

Source		Degrees of Freedom		Fixed or Random	Mean Square Expectation	Appropriate F Test
A	denomination		2	fixed	$\sigma^2_{SCD.AB} + d c \sigma^2_{S.AB} + b c d s \sigma^2_A$	$MS_A / MS_{sw\ (AxB)}$
B	type of course	b - 1	1	fixed	$\sigma^2_{SCD.AB} + d c \sigma^2_{S.AB} + a c d s \sigma^2_B$	$MS_B / MS_{sw\ (AxB)}$
C	sex	c - 1	1	fixed	$\sigma^2_{SCD.AB} + d \sigma^2_{SC.AB} + a b d s \sigma^2_C$	$MS_C / MS_{(SxC)\ w\ (AxB)}$
D	occasion	d - 1	1	fixed	$\sigma^2_{SCD.AB} + c \sigma^2_{SD.AB} + a b c s \sigma^2_D$	$MS_D / MS_{(SxD)\ w\ (AxB)}$
A x B		(a-1) (b-1)	2		$\sigma^2_{SCD.AB} + c d \sigma^2_{S.AB} + c d s \sigma^2_{AB}$	$MS_{AB} / MS_{sw\ (AxB)}$
A x C		(a-1) (c-1)	2		$\sigma^2_{SCD.AB} + d \sigma^2_{SC.AB} + b d s \sigma^2_{AC}$	$MS_{AC} / MS_{(SxC)\ w\ (AxB)}$
A x D		(a-1) (d-1)	2		$\sigma^2_{SCD.AB} + c \sigma^2_{SD.AB} + b c s \sigma^2_{AD}$	$MS_{AD} / MS_{(SxD)\ w\ (AxB)}$
B x C		(b-1) (c-1)	1		$\sigma^2_{SCD.AB} + d \sigma^2_{SC.AB} + a d s \sigma^2_{BC}$	$MS_{BC} / MS_{(SxC)\ w\ (AxB)}$
B x D		(b-1) (d-1)	1		$\sigma^2_{SCD.AB} + c \sigma^2_{SD.AB} + a c s \sigma^2_{BD}$	$MS_{BD} / MS_{(SxD)\ w\ (AxB)}$
C x D		(c-1) (d-1)	1		$\sigma^2_{SCD.AB} + a b s \sigma^2_{CD}$	$MS_{CD} / MS_{(SxCxD)\ w\ (AxB)}$
A x B x C		(a-1) (b-1) (c-1)	2		$\sigma^2_{SCD.AB} + d \sigma^2_{SC.AB} + d s \sigma^2_{ABC}$	$MS_{ABC} / MS_{(SxC)\ w\ (AxB)}$
A x B x D		(a-1) (b-1) (d-1)	2		$\sigma^2_{SCD.AB} + c \sigma^2_{SD.AB} + c s \sigma^2_{ABD}$	$MS_{ABD} / MS_{(SxD)\ w\ (AxB)}$
A x C x D		(a-1) (c-1) (d-1)	2		$\sigma^2_{SCD.AB} + b s \sigma^2_{ACD}$	$MS_{ACD} / MS_{(SxCxD)\ w\ (AxB)}$
B x C x D		(b-1) (c-1) (d-1)	1		$\sigma^2_{SCD.AB} + a s \sigma^2_{BCD}$	$MS_{BCD} / MS_{(SxCxD)\ w\ (AxB)}$
A x B x C x D		(a-1) (b-1) (c-1) (d-1)	2		$\sigma^2_{SCD.AB} + s \sigma^2_{ABCD}$	$MS_{ABCD} / MS_{(SxCxD)\ w\ (AxB)}$
S within A x B school		a b (s-1)	12	random	$\sigma^2_{SCD.AB} + d c \sigma^2_{S.AB}$	$MS_{S\ w\ (AxB)} / MS_{(SxCxD)\ w\ (AxB)}$
S x C within A x B		a b (s-1) (c-1)	12		$\sigma^2_{SCD.AB} + d \sigma^2_{SC.AB}$	$MS_{(SxC)\ w\ (AxB)} / MS_{(SxCxD)\ w\ (AxB)}$
S x D within A x B		a b (s-1) (d-1)	12		$\sigma^2_{SCD.AB} + c \sigma^2_{SD.AB}$	$MS_{(SxD)\ w\ (AxB)} / MS_{(SxCxD)\ w\ (AxB)}$
S x C x D within A x B		a b (s-1) (c-1) (d-1)	12		$\sigma^2_{SCD.AB}$	

The following analyses were undertaken:

1. *Size of school* x *sex* x *occasion* x *school*

   i.e. B'CDS for a) Roman Catholic schools only.

   b) Non-denominational schools in the cities only.

   c) Non-denominational schools outwith the cities only.

2. *Denomination of school* x *size* x *occasion* x *school*

   i.e. AB'DS for a) School mean scores (all pupils).

   b) Boys' mean scores.

   c) Girls' mean scores.

Neither these nor the analyses of pretest or post-test means (page 334) assess the significance of the third order interaction AB'CD and inferences will have to be made. For example if it were found that there was no significant AB'C interaction for either occasion (D) then it is unlikely that there is any significant AB'CD interaction. It is *possible* that some *nearly* significant lower order interactions could result in a low significance AB'CD interaction, but this would be marginal. This argument together with the doubtful value, in educational terms, of knowledge of such an interaction,[1] suggested that the time spent on writing a new computer program for the purpose of accommodating the extra dummy variables would not be justified.

---

[1]It might be seen as (hypothetical example) advantageous to build Roman Catholic schools small to provide an environment conducive to the reduction of the differential effects in change of attitude between boys and girls. It seems unlikely that such a result would influence school building policy.

Within-Schools Analysis

While a between-schools analysis may be appropriate for the comparison of the attitude scores of *groups* of pupils following different science courses or from different types of school, it is also of interest to investigate relationships between the attitude measures of *individual* pupils and other variables. We may ask:

1. To what extent are the attitudes to science of pupils within a school at the end of their second year, related to their sex, social class, measures of general intelligence and divergency, measures of attitude on entry to school, the number of science teachers they are exposed to, and the stream in which they are placed?
2. Is there any pattern of relationships that can be seen as common to a particular group of schools (*e.g.* large schools, or schools which follow integrated science)?

These questions were considered by analysis of measures of individual pupils from twelve different schools. One school was sampled at random from each of the nine cells of the 3 x 3 (size x denomination) table of schools following the Integrated Science course. The additional three schools were sampled from medium-sized schools (of each denomination) following the separate science subjects course. The pupil measures used were: sex, social class (2 variables), general intelligence (AH4), divergency (5 variables), attitudes on entry to secondary school (5 scales), and attitudes at the end of S2 (5 scales).

The social class variables were a) membership of social class 1 and 2 (coded 1), or below (coded 0), and b) membership of social classes 4 and 5 (coded 1), or above (coded 0). In addition to the divergency measures of 'fluency' (used in the between-schools analysis), two measures of 'uniqueness' ('circles' and 'uses of things') were added. Scores on uniqueness are achieved by the pupil providing an idea which has not been suggested by any other pupil *in that school*, and, as such, can only be used in a *within-school* analysis.

Unfortunately it was not possible to determine how many teachers each pupil had encountered in his science course. Pupils frequently moved from one teaching group to another, teachers' absences sometimes caused teaching groups to be split up between other groups (and records of such moves tend not to be kept), at other times another teacher would take over the group for a short period, and schools which have the largest turnover of staff were the least able to keep accurate records of the various changes. There were some stable teaching groups (see below), but the crucial information relating to those pupils undergoing frequent changes of teacher could not be collected. The effect of streaming on the individual pupil could not be investigated. This was due to the predominance of mixed ability science teaching groups, and also to the mixing of streams within science teaching groups (the implication of this seems to be that streaming specifically for science in S1 and S2 is rare in this group of schools).

Analysis of data

Twelve 19 x 19 correlation matrices were computed (using STATPAK, a programme developed at the Computor Unit, University of Warwick, from Hallworth and Brebner package for multi-variate analysis, Hallworth and

Brebner, 1965) of all the Pearson Product Moment intercorrelations of the following variables for pupils in each school: sex, 2 social class, AH4, 5 divergency measures, 5 pretest attitudes (on entry to secondary school), and 5 post-test attitudes (at the end of S2).

Multiple regressions were carried out for each school using each of the 5 post-test attitude scores in turn as the criterion variable. Predictor variables were added in systematically in the following order: sex, social class, AH4, and then *either* divergency (fluency), divergency (uniqueness), pretest attitudes, *or* pretest attitudes, divergency (fluency and uniqueness). This enabled the additional variance accounted for by the introduction of each new variable to be determined.

A table of R^2 (R - multiple correlation coefficient) from each regression step, for each school, on each criterion variable, was drawn up and inspected for patterns of relationships among schools with some characteristic (size, denomination, type of course) in common. As indicated in the 'Results' no such relationships were apparent, and so the twelve correlation matrices were transformed to Fisher z matrices, the individual z values were weighted and averaged to produce a combined z matrix (weighted z's were used since the numbers in the different samples were not equal, the weighting factor being N - 3 where N is the number in the sample), the z matrix was then transformed to a combined r matrix. (This averaging of coefficients has assumed that the several r's arose from a random sampling of the same population *i.e.* the population of second year pupils in Scottish secondary schools.)

Multiple regressions (using the same procedure as for the individual school correlation matrices) were carried out on the combination matrix.

Within-Teaching-Groups Analysis

So far the analysis has been directed towards the description of the variation in attitudes among pupils, and the prediction of attitude scores from other variables. However, attitudes themselves may be predictors of achievement, they may be related to the pupil's perceptions of his own achievement, and they may be related to the teacher's perceptions of the pupil. The last part of the analysis considered the following questions:

1. What level of relationship is there between teachers' ratings of pupils' interest in science and pupils' expressed level of interest?
2. How is a pupil's attitude score at the end of two years of secondary school related to his perceptions of his achievement on school examinations/tests and his teacher's rating of him on academic ability?
3. What value have measures of certain pupil characteristics (attitudes, general intelligence, divergency) measured on entry to secondary school in predicting their academic achievement at the end of the second year?

These three questions are concerned with teachers' rating of pupils and pupils' achievements in school tests and examinations. Such measures can only be strictly comparable among pupils if the pupils are all in the same teaching group. Teachers were asked to rate each pupil in their teaching group for 'interest' and 'academic ability' on a 5, 4, 3, 2, 1 scale (corresponding to the top 10%, next 20%, next 40%, next 20%, and bottom 10%

of pupils). No instructions about the criteria to be used for such ratings were given, the intent was that the teachers would develop their *own* systems of ratings (see Appendix A-17A for instructions to teachers). The teachers also provided 'achievement marks' from school examinations or tests in the pupils' second year (scores for the first year had also been collected but extensive re-organization of teaching groups at the end of S1 precluded the combining of first and second year marks). Such marks had all been made available to the pupils in school and so could be expected to contribute substantially to the pupils' perceptions of his own 'success' in science. Pupils from the three schools following separate physics, chemistry and biology courses were each given three sets of teachers' ratings and achievement marks. The mean of each set of three measures was used in the analysis.

The initial random sampling of pupils together with the re-organization of the teaching groups in most of the schools at some time during the two-year period of the study, resulted in a situation where pupils were frequently scattered throughout a large number of teaching groups rather than concentrated together. This problem was most marked in the larger schools. Analysis was carried out on data for teaching groups which contained a minimum of ten pupils from our sample. Ten of the twelve schools from the within-schools study provided at least one such group, the total number of groups being thirty-eight. The two schools that were not able to provide a suitable group were from the 'medium-sized-non-denominational-outwith-cities-separate science' and the 'large-non-denominational-outwith-cities-integrated-science' cells.

Thirty-eight within-teaching-group 19 x 19 correlation matrices were computed providing intercorrelations between 5 attitude pretest measures, 5 divergency measures, AH4, 5 attitude post-test measures, teachers' ratings of pupils' interest, teachers' ratings of pupils' academic ability, and school measures of science achievement. The median correlation coefficients were determined:

a) between teachers' ratings of interest and each of the 5 attitude post-test scores, AH4, teachers' ratings of academic ability, and school achievement;

b) between teachers' ratings of academic ability and the other eighteen variables;

c) between the schools science achievement measures and each of the other eighteen variables;

d) between AH4 and each of the 5 pretest attitude measures.

From these median coefficients estimates were made of the extent to which the teachers were able to assess their pupils on 'interest' in science as distinct from achievement. The possibility of the pupil's attitudes being related to the teacher's expectations of him was examined. Finally, the value of the attitude measures taken on entry to school as predictors of achievement in science was compared with that of AH4 and divergency scores.

CHAPTER 12

RESULTS OF MAIN 'ATTITUDES TO SCIENCE' STUDY

Each of the five attitude scales will be considered separately in terms of their 'between-schools' and 'within-schools' results of analysis. The final section of the chapter will be concerned with the 'within-teaching-group' analysis results for all five attitude measures together with the various other variables.

Attitude 1: Awareness of the Inter-relationship of the Different Disciplines of Science

Between-schools ANOVA (see Appendix F)

Table 12.1 summarizes the results of the 4-factor design analysis of the mean post-test scores in medium sized schools following the two types of science course. From the significant (1% level) effect due to type of *science course* followed (B) we can conclude that the separate science subjects course appears to be more successful than the integrated science course (Table 1F Appendix F shows $\Sigma B_1 < \Sigma B_2$) in terms of pupils' mean scores on a scale purporting to assess this attitude at the end of their second year. The significant (2.5% level) effect due to *schools* (S) within cells of the same denomination and type of science course offered, indicates that schools were not equally successful in achieving high scores on this attitude scale. Table 3F Appendix F shows the details of the 5-factor design analysis, and Table 12.1 provides a summary of the

TABLE 12.1

SUMMARY OF SIGNIFICANT EFFECTS: ATTITUDE 1, MEDIUM-SIZED SCHOOLS, TWO TYPES OF SCIENCE COURSE
(FIGURES INDICATE SIGNIFICANCE LEVEL OF EFFECT)

	Denomination A	Course B	Sex C	Occasion D	AB	AC	AD	BC	BD	CD	ABC	ABD	ACD	BCD	ABCD	School within A x B S W AB	S x C W AB	S x D W AB
5 factor		.01		.001					.025							.001		.01
4 factor post-test		.01														.025		
4 factor pretest						.025										.001		

significant effects of this and the 4-factor post-test and pretest analyses.

These figures provide strong evidence for changes over the two years in the *type of science course* and *school* effects already noted in the post-test anaylsis. From the non-significant *course* (B) effect in the pretest and the significant (1% level) *course* effect in the post-test, together with the significant *course* x *occasion* (B x D) effect in the 5-factor design, we can conclude that while there are no significant differences in mean attitude scores of pupils entering the two types of science course, by the end of two years the differential effects of the courses are considerable (separate subjects producing more favourable scores). The significant (1% level) *schools* x *occasions* (S x D within AB) effect, together with the reduction of significance level of the *school* effect from pretest to post-test (0.1% to 2.5%), indicates that differences between groups of pupils from different schools tend to get smoothed out over that time.

The analysis of the pretest scores shows a significant interaction effect between *sex and denomination* (2.5% level). (The boys entering Roman Catholic schools scored considerably higher than the girls, this difference was less marked in the non-denominational schools outwith the cities and was, in fact, reversed in the non-denominational schools in the cities.) In the post-test analysis this effect does not reach the 5% significance level, which indicates a decrease over time. This tendency, however, is not strong enough to produce an A x C x D effect that is significant at the 5% level.

Tables 6F and 7F Appendix F show details of the analysis of variance of post-test scores for various sized schools following Integrated Science courses. The main effects of *sex* and *denomination* (A) are seen to be significant at the 1% level. The highest groups of scores here were from the non-denominational schools outwith the cities, followed by the non-denominational schools inside the cities, with the Roman Catholic schools' scores being lowest. (The differences between the non-denominational schools outwith the cities and the other two groups were significant at the 0.1% level, the differences between the non-denominational schools in the cities and the Roman Catholic schools were significant at the 5% level.) This effect (A) was not significant in the analysis of pretest scores (see Table 12.2 for summary of effects), indicating an increasing influence of the denomination of the school with time. However, this is not supported by a significant *denomination* x *occasion* (A x D) effect in school means, boys' or girls' analyses.

The significant *sex* effect in the post-test scores (boys > girls) was not apparent in the pretest scores. This, together with the significant (1% level) *sex* x *occasion* effect for non-denominational schools outwith the cities, suggests that sex differences in attitudes are increasing over the two years.

For all pupil groups together, the significant *school* effect for the pretest (1% level) disappears for the post-test. The decreasing influence of the school with time (*cf.* similar result for schools following two types of course) is supported by a significant *school* x *occasion* (S x D within AB) effect for non-denominational schools outwith the cities.

The summary Table 12.2 for all the different groups shows that there is a

TABLE 12.2 SUMMARY OF SIGNIFICANT EFFECTS: ATTITUDE 1, VARIOUS SIZED SCHOOLS FOLLOWING INTEGRATED SCIENCE COURSE (FIGURES INDICATE SIGNIFICANCE LEVEL OF EFFECT)

	Denomination	Size	Sex	Occasion											School		
	A	B'	C	D	AB'	AC	AD	B'C	B'D	CD	AB'C	AB'D	ACD	B'CD	S W AB' or B'	S x C W AB'	S x D W AB'
4 factor post-test	.01		.01														
4 factor pretest															.01		
4 factor School Mean				.001											.05		
4 factor Boys				.001											.01		
4 factor Girls	.05			.001													
4 factor Roman Catholic		.05		.01													
4 factor Non-Denominational 1				.001											.025		
4 factor Non-Denominational 2				.01						.01					.001	.025	.01

permanent significant *school* effect for all but 'Roman Catholics' and 'girls'. The significant *occasion* effect (D) for all analyses reflects overall increases in scores on this attitude.

Between-schools: multiple regression (see Tables 1E to 4E Appendix E and 16F Appendix F)

Introduction of Group A predictor variables into the multiple regression carried out on data from the 40 schools, was able to account for 58% of the variance in mean school scores for attitude 1. Significant *additional* variance was accounted for by:

a) social class (36%,p<0.001, favourable attitudes correlate positively with high social class)
b) mean AH4 (6%, p<0.01, correlates positively with attitude)
c) pretest attitudes (13%, p<0.01, attitude correlates positively with all pretest attitudes)

when entered into the equation *in that order*.

The introduction of Group B predictors (school characteristics) accounted for 2% extra variance only and this contribution was non-significant at the 5% level.

The single Group C variable (type of science course) accounted for an additional 11% of the variance significant at the 0.1% level, with pupils following the separate subjects course having the higher scores. The regression was repeated for 38 schools including additional Group C variables. When each of the variables associated with sex grouping of class, number of periods of science in S1, and class size was added by itself to the regression equation *before* the addition of type of science course, the extra contribution to variance accounted for in each case was 6% (each significant at the 1% level attitude score correlated positively with each variable). The variable associated with mixed ability accounted for 4% of extra variance (significant at the 2.5% level attitude score correlated negatively with mixed ability grouping). When all four variables were added to the equation they accounted for an extra 8% of the variance (significant at the 2.5% level). If, however, these variables were added to the equation *after* the variable concerned with the type of science course, their extra contribution was 4% which did not reach the 5% significance level. One possible reason for this was the very high negative correlation (r = -0.69) between the number of periods of science and the Integrated Science course variable. A common pattern for schools doing separate subjects was for two periods to be allocated for each of the three sciences, while the most common pattern for Integrated Science was a four or five period allocation for the whole subject. (The simple correlation between this attitude score and the number of periods of science is high, r = 0.45.) The regression for 23 schools was carried out including the extra Group A variables of divergency (fluency). These variables provided no significant contribution to the variance of attitude scores accounted for.

A summary of significant effects identified in the analyses of variance and predictor variables accounting for substantial additional variance is given in Table 12.3.

TABLE 12.3

SUMMARY OF SIGNIFICANT EFFECTS AND PREDICTORS: ATTITUDE 1, BETWEEN-SCHOOLS ANALYSIS

C - analysis of schools following different courses

S - analysis of schools of different sizes

Predictors accounting for a minimum 5% variance and at significance level 5% or better in multiple regression	Effects shown as significant by analysis of post-test scores	Indication of changing (over S1 and S2) effects, supported or unsupported	Permanent effects (over S1 and S2)
	S: sex differences	S: Increasing sex differences (supported)	
Social Class AH 4		C: Decreasing sex X denomination (unsupported)	
Pretest Attitudes	S: Denominational differences	S: Increasing denominational differences (unsupported)	
	C: School differences	S/C: Decreasing school differences (supported)	S/C: School differences (except RC and girls)
Type of science course	C: Science course differences	C: Increasing course differences (supported)	

Within-schools: 12 schools analysis

Table 17F Appendix F shows the variance accounted for by the addition of the various variables to the multiple regressions for each of the twelve schools sampled. No clear relationships between denomination of school, size of school, or type of science course followed, and variance accounted for are apparent. The range of variance accounted for within the twelve schools with all predictors in the equation is 18-51%, median 27%. Table 12.4 shows the variance accounted for when the multiple regression was carried out on the *combined* r matrix. Significant additional variance in attitudes was accounted for by (see Table 18F Appendix F):

a) social class (2%, p<0.001, favourable attitudes correlate positively with high social class)
b) mean AH4 (6%, p<0.001, correlates positively with attitude score)
c) pretest attitudes (7%, p<0.001, attitude correlates positively with all pretest attitudes).

(Divergency fluency measures, but not uniqueness, accounted for significant extra variance, p<0.05, if entered *before* pretest attitudes. However, this amounts to 1% only, and the contribution was not significant if the measures were added *after* pretest attitudes.)

The multiple correlation coefficient R has a non-normal distribution and within-school estimates of the population R^2 will tend to be high. If the true value is R^2, then the expected value is $R^2 + (1 - R^2)(p - 1)/(n - 1)$ where n is the number of pupils and p is the number of predictors used (Peaker 1967). The *combined* correlations, with a far greater number of degrees of freedom, provide a better estimate of the population R, and will indicate a lower value than the median of individual schools. (This bias results from the fact that the addition of any predictor variable always

TABLE 12.4

% Variance accounted for by modified stepwise regression using combined r matrix for 12 schools for Attitudes 1, 2, 3, 4 and 5 (for key see Table 55)

	Attitude 1	Attitude 2	Attitude 3	Attitude 4	Attitude 5
B	2	1	2	5	2
C	8	6	8	6	11
D	15	17	21	13	27
E	(9)	(6)	(9)	(7)	(15)
F	(9)	(7)	(9)	(8)	(15)
G	16	17	21	15	29

makes *some* contribution to the predicted variance, and these small contributions add up.)

The total variance accounted for by the predictors in the combined analysis was 16%. Both this and the median value (27%) are considerably below that of the between-schools regression (58%). Variance unaccounted for in each case is due, in part, to errors of measurement and, in part, to variables that we have not taken into account. For example pupils' attitudes may well be related to parental attitudes, or teachers' attitudes, or abilities we have not considered, or specific peer-group influences. These individual differences between pupils could be expected to be 'averaged out' when we consider *school* measures (*i.e.* mean values for pupils in that school), and this may account for being able to account for more than half the variance in the between-schools analysis and only one sixth of the variance in the within-schools analysis.

Despite the within-schools analysis accounting for less variance, the significant contributions come from the same variables as those of Group A in the between-schools analysis. In both analyses AH4 and pretest attitude scores are substantial predictors, and the best single predictor of this attitude, from the combined r matrix, was its own pretest measure (r = 0.28, see Table 6E Appendix E).

Attitude 2: Awareness of the Relationship of Science to Other Aspects of the Curriculum

Between-schools: ANOVA (see Appendix G)

Analysis of the post-test scores of the medium sized schools following two types of course (Tables 12.5 and 12.7) indicates significant differences between the *sexes* (boys > girls, $p<0.01$), significant differences between *schools* ($p<0.025$) and a weak ($p<0.05$) *denomination* x *science course* interaction effect. The total scores for A x B cross classifications (Table 1G Appendix G) suggest that while Roman Catholic schools obtain higher mean scores on this attitude if they follow a separate science subject course rather than an Integrated Science course, this tendency is not apparent for non-denominational schools in the cities, and the trend is in fact reversed for non-denominational schools outwith the cities.

The pretest and 5-factor design analyses (see summary Table 12.5) show that the *sex* effect is a permanent one through the two year period ($p<0.01$ and $p<0.001$) but this must be interpreted in conjunction with the significant *science course* x *sex* (B x C) interaction effects ($p<0.01$ for both). These suggest that the sex differential is significant for the pupils following the integrated course but not the separate sciences course (Table 4G). The absence of a significant B x C interaction at the post-test stage implies that this effect becomes reduced over the two year period *i.e.* girls in the integrated course are able to make up some of the differences that existed on entry. However, this is not supported by a significant B x C x D effect.

The significant (but weak) *denomination* x *science course* effect does not appear in the pretest analysis suggesting that this effect was increasing over the two years, but this suggestion is not supported by a significant A x B x D effect.

Tables 12.6 and 12.7 show the analysis of mean scores of schools of different sizes following the Integrated Science course. It can be seen that

TABLE 12.5 SUMMARY OF SIGNIFICANT EFFECTS: ATTITUDE 2, MEDIUM-SIZED SCHOOLS FOLLOWING TWO TYPES OF SCIENCE COURSE

(Figures indicate significance level of effect)

	Denomination A	Course B	Sex C	Occasion D	AB	AC	AD	BC	BD	CD	ABC	ABD	ACD	BCD	ABCD	School within A x B S w AB	SxC w AB	SxD w AB
5 factor			.001	.001				.01								.01		
4 factor post-test			.01	/	.05		/		/	/		/	/	/	/	.025	/	/
4 factor pretest			.01	/			/	.01	/	/		/	/	/	/	.001	/	/

TABLE 12.6 SUMMARY OF SIGNIFICANT EFFECTS: ATTITUDE 2, VARIOUS SIZED SCHOOLS FOLLOWING INTEGRATED SCIENCE COURSE (FIGURES INDICATE SIGNIFICANCE LEVEL OF EFFECTS)

	Denomination A	Size B'	Sex C	Occasion D	AB'	AC	AD	B'C	B'D	CD	AB'C	AB'D	ACD	B'CD	AB'CD	School within AB' or B' S w AB' or B'	SxC w B'	SxD w B'
4 factor post test			.01													.01		
4 factor pre test								.025								.025		
4 factor school means				.001														
4 factor boys				.001												.025		
4 factor girls	.001			.001														
4 factor RC				.001													.025	
4 factor non-denominational 1				.025				.05								.025		
4 factor non-denominational 2				.001												.01		

the only significant effects are those of *sex* ($p<0.01$) and the *school* ($p<0.05$). We can conclude from this that at the end of two years of secondary school the boys score significantly higher on this attitude than do the girls, and that schools are not equally successful in attaining mean scores on this attitude.

The significant *sex* effect is not apparent on entry to secondary school but there is a significant *size of school* x *sex* interaction effect ($p<0.025$). Table 8G shows that while boys in medium sized schools enter secondary school with more favourable attitudes than girls, there is no difference between the sexes in small schools, and in large schools the girls' attitudes on entry are more favourable than the boys'. The reduction of the *size* x *sex* effect with time is not supported by any significant B x C x D effect, and the analysis of scores from non-denominational schools in the cities shows B x C as an enduring effect for their pupils ($p<0.05$).

The suggested increasing *sex* effect is not supported by any significant C x D effects.

As with Attitude 1, analyses of the different groups of pupils showed significant *school* effects for all but Roman Catholics and girls (see Table 12.6), and the general increases in attitude scores over the two years are reflected in consistent significant *occasion* effects.

Between-schools: multiple regression (see Tables 1E to 4E Appendix E and 16G Appendix G)

In the 40 school multiple regression analysis Group A variables were able to account for 44% of the variance in attitude scores. Significant *additional* variance was accounted for by:

a) social class (15%, $p<0.01$, favourable attitudes correlate positively with high social class)

b) mean AH4 (6%, $p<0.025$, correlates positively with attitude)

c) pretest attitudes (19%, $p<0.01$, attitude correlates positively with all pretest attitudes),

when entered into the equation *in that order*.

Two of the variables of Group B also provided significant *additional* prediction (Table 16G Appendix G):

a) location of the school (8%, $p<0.01$, highest scores outwith the city burghs)

b) size of school (8%, $p<0.01$, highest scores from small schools).

None of the Group C variables provided any further prediction significant at the 5% level or better.

When the divergency (fluency) variables were added to Group A in the 23 school regression, significant *additional* variance was accounted for:

a) variables introduced *after* AH4 but *before* pretest attitudes - 19%, $p<0.01$ (attitude scores correlate positively with all divergency measures)

b) variables introduced *after* both AH4 and pretest attitudes - 9%, $p<0.05$.

Table 12.7 provides a summary of significant effects identified in the analyses of variance and predictors accounting for substantial variance.

TABLE 12.7

SUMMARY OF SIGNIFICANT EFFECTS AND PREDICTORS: ATTITUDE 2

C - analysis of schools following different courses

S - analysis of schools of different sizes

Predictors accounting for a minimum 5% variance and at significance level 5% or better in multiple regression	Effects shown as significant by analysis of post-test scores	Indication of changing (over S1 and S2) effects, supported or unsupported	Permanent effects (over S1 and S2)
Social class	S/C: Sex differences	S: Increasing sex differences (unsupported)	C: Sex
AH 4			
Pretest Attitudes		C: Decreasing sex X course (unsupported)	
Divergency		S: Decreasing sex X size (unsupported)	
Location of school			
Size of school	C: Denomination X type of course	C: Increasing denomination X course (unsupported)	
	C/S: School differences		C/S: School differences (except RC and girls)

Within-schools: 12 schools analysis

As in the case of attitude 1, no patterns of the variance accounted for were observed that were related to schools with common characteristics (see Table 17G Appendix G). The total variance accounted for in individual schools ranged from 19-44%, median 32.5%. The total variance of scores on this attitude accounted for by the multiple regression on the *combined* r matrix with all predictor variables included was 17%. Significant *additional* variance in the regression was accounted for by (Table 18G Appendix G):

a) social class (1%, p<0.025, favourable attitudes correlate with high social class)

b) mean AH4 (5%, p<0.001, correlates positively with attitude)

c) pretest attitudes (11%, p<0.001, correlates positively with all pretest attitudes).

(Divergency uniqueness, but not fluency, accounted for significant extra variance, p<0.025, if entered *before* but not *after* pretest attitudes. This amounted to only 1%.)

The total variance accounted for, as for attitude 1, was less than that of the between-schools analysis (44%), but again the patterns of significant predictors were similar. In both cases AH4 and pretest attitudes accounted for substantial variance, and the best single predictors were the pretest measures of this attitude (r = 0.30) and attitude 3 (r = 0.32) (Table 6E Appendix E).

Attitude 3: An Awareness of the Social and Economic Implications of Science for the Community

Between-schools: ANOVA (see Appendix H)

From Table 12.8 it can be seen that for the analysis of medium sized schools post-test scores, the only significant effect is that of *sex* ($p<0.01$, boys' > girls').

The absence of a *sex* effect that reaches the 5% level in the pretest scores (Table 12.8) indicates that the differences between boys' and girls' attitudes are increasing over the two years, and this is supported by the significant *sex* x *occasion* ($p<0.01$) effect in the 5-factor design analysis. A suggested decrease in *school* effect from pretest ($p<0.01$) to post-test (N.S) is not supported by a significant *schools* x *occasions* effect.

Tables 12.9 and 12.10 show significant *sex* differences (boys > girls, $p<0.001$), *school* differences ($p<0.001$), and a *size of school* x *sex* interaction effect (the sex difference is clear for medium sized schools, reduced for small schools and disappears in large schools) in the post-test means of schools of various sizes.

None of these three effects reached significance level in the pretest scores analysis (Table 12.9), but here a *denomination* x *size* interaction reached the 2.5% level of significance (for small schools highest scores are from non-denominational schools outwith the cities, medium sized schools highest scores are from non-denominational city schools, for large schools highest scores are from the Roman Catholic schools). The implied increasing *sex* differences, increasing *sex* x *size* interaction effect, and decreasing

TABLE 12.8 SUMMARY OF SIGNIFICANT EFFECTS: ATTITUDE 3, MEDIUM-SIZED SCHOOLS FOLLOWING TWO TYPES OF SCIENCE COURSE

(Figures indicate significance level of effect)

	Denomination A	Course B	Sex C	Occasion D	AB	AC	AD	BC	BD	CD	ABC	ABD	ACD	BCD	ABCD	School within A x B S w AB	SxC w AB	SxD w AB
5 factor			.025	.01						.01						.01		
4 factor post-test			.01															
4 factor pretest																.01		

TABLE 12.9 SUMMARY OF SIGNIFICANT EFFECTS: ATTITUDE 3, VARIOUS SIZED SCHOOLS FOLLOWING INTEGRATED SCIENCE COURSE (FIGURES INDICATE SIGNIFICANCE LEVEL OF EFFECT)

	Denomination A	Size B	Sex C	Occasion D	AB'	AC	AD	B'C	B'D	CD	AB'C	AB'D	ACD	B'CD	AB'CD	School within AB' or B' S w AB' or B'	SxC w B'	SxD w B'
4 factor post test			.001					.05								.001		
4 factor pre test					.025													
4 factor school means				.01														
4 factor boys				.001												.01		
4 factor girls				.05	.01													
4 factor RC				.01														
4 factor non-denominational 1				.01												.025		
4 factor non-denominational 2																.01		.01

TABLE 12.10

SUMMARY OF SIGNIFICANT EFFECTS AND PREDICTORS: ATTITUDE 3

C - analysis of schools following different courses

S - analysis of schools of different sizes

Predictors accounting for a minimum 5% variance and at significance level 5% or better in multiple regression	Effects shown as significant by analysis of post-test scores	Indication of changing (over S1 and S2) effects, supported or unsupported	Permanent effects (over S1 and S2)
Sex	C/S: Sex differences	Increasing sex differences C - supported S - unsupported	
Social Class	S: Sex x size of school	S: Increasing sex x size (unsupported)	
Pretest Attitudes		S: Decreasing denomination x size (unsupported)	
	S: School differences	C: Decreasing school differences (unsupported) S: Increasing school differences (supported)	S/C: School differences (except RC, girls, one pretest, one post test)

denomination x *size* interaction effect are not supported by corresponding significant C x D, B x C x D and A x B x D interaction effects. However, the significant *schools* x *occasions* (S x D within AB) interaction effect (p<0.01) for non-denominational schools outwith the cities supports the increasing *schools* effect over the two years.

Again Table 12.9 shows significant *schools* effects for boys and non-denominational groups but not for girls or Roman Catholics. General increases in attitude scores with time are indicated by significant *occasions* effects for all analyses but the non-denominational outwith the cities (and here the effect just fails to reach the 5% level of significance).

Between-schools: multiple regression (see Tables 1E to 4E Appendix E and 16H Appendix H)

Three predictor variable from Group A accounted for significant *additional* variance when introduced into the 40 school multiple regression:

i) sex (8%, p<0.025, boys having more favourable attitudes then girls)

ii) social class (15%, p<0.01, favourable attitudes correlate positively with high social class)

iii) pretest attitudes (28%, p<0.001, attitude correlates positively with all pretest attitudes)

when entered into the equation *in that order*. (Table 16H Appendix H)

No variables from Groups B or C provided any significant contribution to additional variance accounted for.

(It is interesting, however, that in the analysis of 38 schools the size of class correlated substantially, r = -0.33, with this attitude score.)

Within-schools: 12 schools analysis

The total variance accounted for by individual school multiple regressions varied between 21 and 52% with a median of 38% (Table 17H Appendix H), and no patterns for particular types of school were apparent.

The multiple regression carried out on the *combined* r matrix accounted for 21% variance when all predictors were included in the equation. Significant *additional* variance was accounted for when the following predictors were added to the equation (Table 18H Appendix H):

a) social class (2%, $p<0.001$, favourable attitudes correlate with high social class)

b) AH4 (6%, $p<0.001$, correlates positively with attitude scores)

c) pretest attitudes (13%, $p<0.001$, attitude correlates positively with all pretest attitude scores)

(Divergency fluency, but not uniqueness, accounted for 1% of variance, $p<0.05$, if introduced into the regression *before*, but not *after*, pretest attitudes.)

Again, the total variance accounted for is less than the between schools analysis (51%), but for this attitude the predictors providing significant contributions were not the same in the two analyses (sex was a significant predictor in the between-schools only, AH4 in the within-schools only). However, in both cases the largest amount of additional variance was accounted for by pretest attitudes. The best single predictor of an individual attitude score was the pretest score on this attitude (median $r = 0.40$).

Attitude 4: Interest and Enjoyment in Science

Between-schools: ANOVA (see Appendix I)

Analysis of post-test scores (Table 12.11) indicates that in medium sized schools following the two types of science course the only significant effect is that of *sex* ($p<0.001$, boys > girls).

The pretest scores (Table 12.11) also show a *sex* effect but at a lower level of significance ($p<0.01$). This suggests an increasing *sex* effect which is supported in the 5-factor design analysis (Table 12.11) by a significant *sex* x *occasion* ($p<0.001$) effect. From the summary table it can be seen that the *school* effect present at the pretest ($p<0.001$) is decreasing (post-test N.S) as indicated by a significant *schools* x *occasions* effect in the 5-factor design ($p<0.05$).

Table 12.12 shows a significant *sex* effect ($p<0.001$, boys > girls), and a significant *school* effect ($p<0.01$) in post-test scores for schools of various sizes following Integrated Science. The significant *sex* x *occasions* interaction effects for each denomination analysis indicates that the *sex* effect is increasing with time, as was the case for the medium sized schools analysis. The significant ($p<0.001$) *schools* x *occasions* effect in non-denominational schools outwith the cities supports the suggestion that, like the medium sized schools, the *school* effect is decreasing with time.

The significant *school* effect is permanent for boys ($p<0.025$), non-denominational schools in and outwith the cities ($p<0.01$, $p<0.001$), but, as with the other attitudes, not for the girls or Roman Catholic schools.

TABLE 12.11 SUMMARY OF SIGNIFICANT EFFECTS: ATTITUDE 4, MEDIUM-SIZED SCHOOLS FOLLOWING TWO TYPES OF SCIENCE COURSE

(Figures indicate significance level of effect)

	Denomination A	Course B	Sex C	Occasion D	AB	AC	AD	BC	BD	CD	ABC	ABD	ACD	BCD	ABCD	School within A x B S· w AB	SxC w AB	SxD w AB
5 factor	.05		.001	.001						.001						.001	.05	.C5
4 factor post-test			.001															
4 factor pretest	.025		.01													.001		

TABLE 12.12 <u>SUMMARY OF SIGNIFICANT EFFECTS: ATTITUDE 4, VARIOUS SIZED SCHOOLS FOLLOWING INTEGRATED SCIENCE COURSE (FIGURES INDICATE SIGNIFICANCE LEVEL OF EFFECTS)</u>

	Denomination A	Size B'	Sex C	Occasion D	AB'	AC	AD	B'C	B'D	CD	AB'C	AB'D	ACD	B'CD	AB'CD	School within AB' or B' S w AB' or B'	SxC w B'	SxD w B'
4 factor post test			.001													.01		
4 factor pretest			.001													.01		
4 factor school means				.001														
4 factor boys				.01												.025		
4 factor girls				.001														
4 factor RC			.025	.01						.05								
4 factor non-denominational 1			.01	.01						.05						.01		
4 factor non-denominational 2			.025							.01						.001	.01	.001

The significant *occasions* effect in all but non-denominational schools outwith the cities reflects the general *decrease* in interest scores over the two years.

Between-schools: multiple regressions (see Tables 1E to 4E Appendix E and 16I Appendix I)

Group A variables accounted for 66% of the variance on this attitude in the 40-schools regression. The following variables provided significant *additional* contributions (Table 16I Appendix I):

i) sex (25%, p<0.001, boys > girls)

ii) pretest attitudes (31%, p<0.001, all pretest attitudes except *attitude* 5 correlate positively with this attitude).

None of the variables in Groups B or C in this or in the 23- and 38-school analyses accounted for any significant extra variance, except for the mixed ability teaching group variable. This accounted for 3% additional variance which just reaches the 5% level of significance (with pupils in mixed ability classes showing the less favourable attitudes). (The 'Interest' scores had a simple correlation coefficient r = -0.36 with size of class, *i.e.* 13% common variance.) Significant effects are summarized in Table 12.13.

Within-schools: 12 schools analysis

The total variance in scores on this interest scale accounted for in the multiple regressions for the individual schools varied from 20 to 40% with a median of 32% (Table 17I Appendix I). The multiple regression

TABLE 12.13

SUMMARY OF SIGNIFICANT EFFECTS AND PREDICTORS: ATTITUDE 4

C - analysis of schools following different courses

S - analysis of schools of different sizes

Predictors accounting for a minimum 5% variance and at significance level 5% or better in multiple regression	Effects shown as significant by analysis of post-test scores	Indication of changing (over S1 and S2) effects, supported or unsupported	Permanent effects (over S1 and S2)
Sex	C/S: Sex differences	C/S: Increasing sex differences (supported)	C/S: Sex differences
Pretest Attitudes		C/S: Decreasing school differences (supported)	C/S: Schools (except RC, girls and one post test)

carried out on the *combined* r matrix accounted for a total of 15% of the variance in interest. Significant additional variance was accounted for by:

a) sex (4%, p<0.001, boys scored higher than girls)

b) social class (1%, p<0.025, interest scores correlate with high social class)

c) AH4 (1%, p<0.01, correlates positively with interest scores)

d) divergency - fluency and uniqueness (2%, p<0.01, correlates positively with interest scores)

e) pretest attitudes (7%, p<0.001, interest correlates positively with all pretest attitudes). (Table 18I Appendix I)

The total variance accounted for is far less than that for the between-schools analysis (66%), but in both cases sex and pretest attitudes are the two predictors accounting for substantial variance on interest. The best single predictor of interest is the pretest measure of interest (r = 0.28).

<u>Attitude 5: An Objectivity in Observation and in Assessing Observations</u>

Between-schools: ANOVA (see Appendix J)

The only significant effect apparent in the analysis of medium sized school post-test scores (Table 12.14) is that of *school*. The significant *science course* effect in the pretest scores together with the significant *science course* x *occasion* effect in the 5-factor design analysis (see Table 12.14) indicate that differences between pupils entering the two types of course have disappeared after two years of science. The 5-factor analysis also indicates a *sex* effect (p<0.05) but this is not strong enough to appear in either pre- or post-test analyses. The analysis of the schools

TABLE 12.14 SUMMARY OF SIGNIFICANT EFFECTS: ATTITUDE 5, MEDIUM-SIZED SCHOOLS FOLLOWING TWO TYPES OF SCIENCE COURSE

(Figures indicate significance level of effect)

	Denomination	Course	Sex	Occasion												School within A x B	SxC w AB	SxD w AB
	A	B	C	D	AB	AC	AD	BC	BD	CD	ABC	ABD	ACD	BCD	ABCD	S w AB		
5 factor			.05	.001					.01							.001		
4 factor post-test																.01		
4 factor pretest		.025														.05		

of various sizes (Table 12.15) also shows significant *school* effects at both pretest and post-test (this time for *all* groups).

The significant *occasions* effect for all groups reflects the increases in scores on this attitude over the two years.

Between-schools: multiple regression (see Tables 1E to 4E Appendix E and 16J Appendix J)

Three variables from Group A accounted for significant *additional* variance in the 40-schools analysis (Table 16J Appendix J):

i) social class (32%, p<0.001, favourable attitudes correlate positively with high social class)

ii) AH4 (13%, p<0.001, attitude score correlates positively with AH4)

iii) pretest attitudes (13%, p<0.001, all pretest attitudes correlate positively with this attitude).

Neither the predictors from Groups B or C, nor those in the 23-school analysis accounted for any significant further variance. In the 38-school analysis significant *additional* variance was accounted for by the following Group C predictors:

i) Mixed ability grouping (3%, p<0.025, correlates negatively with attitude).

ii) Mixed sex grouping (3%, p<0.025, correlates positively with attitude).

iii) Number of science periods in S1 (3%, p<0.025, correlates positively with attitude).

TABLE 12.15 SUMMARY OF SIGNIFICANT EFFECTS: ATTITUDE 5, VARIOUS SIZED SCHOOLS FOLLOWING INTEGRATED SCIENCE COURSE (FIGURES INDICATE SIGNIFICANCE LEVEL OF EFFECTS)

	Denomi-nation	Size	Sex	Occa-sion	AB'	AC	AD	B'C	B'D	CD	AB'C	AB'D	ACD	B'CD	AB'CD	School within AB' or B' S w AB' or B'	SxC w B'	SxD w B'
4 factor post test																.05		
4 factor pretest																.05		
4 factor school mean				.001												.001		
4 factor boys				.001												.001		
4 factor girls				.001												.001		
4 factor RC				.001												.01	.05	
4 factor non-denominational 1				.001												.01		
4 factor non-denominational 2				.001												.025		

TABLE 12.16

SUMMARY OF SIGNIFICANT EFFECTS AND PREDICTORS: ATTITUDE 5

C - analysis of schools following different courses

S - analysis of schools of different sizes

Predictors accounting for a minimum 5% variance and at significance level 5% or better in multiple regression	Effects shown as significant by analysis of post-test scores	Indication of changing (over S1 and S2) effects, supported or unsupported	Permanent effects (over S1 and S2)
Social class AH 4 Pretest Attitudes	C: Sex (weak) C/S: School differences	C: Decreasing course differences (supported)	C/S: School differences

Within-schools: 12 schools analysis

The total variance accounted for in the multiple regressions for the individual schools (Table 17J Appendix J) varied from 34 to 59% (median 42%). The total variance accounted for in the multiple regression carried out on the *combined* r matrix was 29%. Significant *additional* variance was accounted for by the following predictor variables:

a) social class (2%, p<0.001, favourable attitudes correlate with high social class)
b) AH4 (9%, p<0.001, correlates positively with attitude score)
c) pretest attitudes (16%, p<0.001, this attitude correlates positively with all pretest attitudes)
d) divergency fluency and uniqueness (2%, p<0.001, correlates positively with attitude). (Table 18J Appendix J)

Apart from (d) this list corresponds to that of the between-schools analysis, although the total variance accounted for is far less than the 62% of the between-schools. In both cases substantial variance is accounted for by AH4 and pretest attitudes. The best single predictor of this attitude is the pretest measure of this attitude (r = 0.47).

Teaching Group Analysis

The areas of concern here, as described in Chapter 11 page 359, are:

1. the relationship between teachers' ratings of pupils' interest and pupils' expressed level of interest;
2. the relationship between pupils' attitudes to science at the end of S2 and

a) their perceived science achievement,

b) the teacher's rating of academic ability;

3. the value of certain pupil characteristics (including attitudes) measured on entry to secondary school as predictors of science achievement.

The 38 teaching groups investigated varied in both size and composition and were not distributed evenly or randomly over the 10 schools from which they were selected. It was not, therefore, possible to apply significance tests to the results obtained for them. Moreover, the small sizes of the groups (minimum n = 10) led to correlations between any two variables that tended to vary widely across the groups. It was assumed that general trends of relationships between variables would be best reflected by median correlations (see Table 12.17); the stability of these trends may be inferred from the fact that on average, the upper quartile correlations were 0.16 above, and the lower quartile correlations were 0.18 below, the median correlations.

The teachers' ratings of pupils' level of interest was only able to account for 16% of the variance of pupils' expressed level of interest in science (median r = 0.39). These ratings had a lower correlation with the other attitude scales (0.28, 0.29, 0.32 and 0.26) but shared 72% and 66% common variance with the teachers' ratings on pupils' *academic ability* and school science *achievement* measures respectively.

The median correlation coefficients between each of the attitude scores and the school science achievement mark showed a maximum common variance of 13% (median r's = 0.33, 0.34, 0.36, 0.36 and 0.32). A similar pattern was

	Interrelationship sc. (pretest)	Relationship sc. with rest of curriculum (pretest)	Social & Economic Implic. of science (pretest)	Interest & Enjoyment in science (pretest)	Objectivity (pretest)	Circles (fluency)	Uses (fluency)	Meanings	Circles (uniqueness)	Uses (uniqueness)	Interrelationship sc. (post test)	Relationship sc. with rest of curriculum (post test)	Social & Economic Implic. of science (post test)	Interest & Enjoyment in science (post test)	Objectivity (post test)	Teacher's rating on interest	Teacher's rating on academic ability	School achievement measure	AH4
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Teacher's rating of interest											.28	.29	.32	.39	.26	-	.85	.82	.38
Teacher's rating of academic ability	.13	.20	.22	.20	.19	.28	.41	.44	.24	.31	.32	.35	.33	.36	.32				.45
School achievement measures	.18	.16	.20	.22	.27	.24	.39	.44	.28	.28	.33	.35	.36	.36	.32				.49
AH4	.23	.17	.19	.03	.26														

TABLE 12.17 Within Teaching Groups: median correlations between Teacher's ratings and measures of various pupil characteristics.

obtained for the relationships between teachers' ratings on academic ability and attitudes (median r's = 0.32, 0.35, 0.33, 0.36 and 0.32). These do not seem very strong relationships, but the correlation coefficients between the teachers' ratings on academic ability and the attitude measures taken on *entry* to secondary school are even lower (median r's = 0.13, 0.20, 0.22, 0.20 and 0.19). It might seem then, that the pupils' attitudes *had moved into line* with their teachers' ratings of them on academic ability over the two year period *i.e.* the pupil's attitudes are related to his teacher's expectations of him. The objection to this explanation is that the teachers' ratings were made at a *different time* from assessment of *pretest* attitudes. However, the teachers could not produce ratings at the time of pretest since they had not yet met the pupils. We could hypothesise that if the teachers had been forced to indicate their expectations of pupils' ability, the best estimate they could have made would have been from measures of pupils' intelligence received from the primary school. The AH4 scores correspond to such measures. It is hypothesised, therefore, that the relationships between teachers' ratings of pupils on *academic ability on entry* to school and *pretest attitude* measures, correspond to the AH4/attitude relationships, for which median r's = 0.23, 0.17, 0.19, 0.03 and 0.26. These relationships are substantially less strong than the rating on academic ability/attitude relationships at the end of S2, the difference being most striking between the correlations with the 'Interest and enjoyment in science' measure (0.36 and 0.03).

None of the pretest attitude measures were able to predict more than 7% of the variance of science achievement scores in S2 (median r's = 0.18, 0.16, 0.20, 0.22 and 0.27). This is considerably lower than the 24%, 19% and 15% predicted by AH4 scores, 'Meanings of Words' scores and 'Uses of Things' scores respectively.

CHAPTER 13

DISCUSSION OF RESULTS AND CONCLUSIONS

The first section of this chapter is concerned with discussions of the influence of the various independent variables on attitude scores as revealed by the empirical study. The order of these discussions follows a similar pattern to that of the entry of predictor variables into the multiple regression equations. Firstly, the influence of characteristics of pupils will be considered, followed by those of the school and finally the science class.

Following this, the next two sections will summarize the empirical findings related to the substantive hypotheses developed in Chapters 4, 5 and 7 and to the research questions posed in Chapter 7.

The final section draws together the general conclusions from the two strands of this thesis - the evaluation of the worth of the attitude objectives and the study of the factors influencing their achievement.

Discussion of Empirical Results - The Influence of the Independent Variables on Attitude Scores

Tables 13.1 and 13.2 summarize the variance in attitudes accounted for by pupil characteristics in the multiple regression equations of the between-schools and within-schools analyses. It can be seen from Table 13.1 that if we rank the five attitude scales in order of *increasing* variance accounted for by *sex* and *pretest attitudes*, and in order of *decreasing* variance

TABLE 13.1

	Independent Variables (characteristics of pupil groups)			
	Sex (increasing variance)	Social Class (decreasing variance)	AH4 (decreasing variance)	Pretest Attitudes (increasing variance)
Attitude Scales	5 (0)	1 (36)	5 (13)	5 (13)
	1 (0)	5 (32)	1 (6)	1 (13)
	2 (0)	2 (15)	2 (6)	2 (19)
	3 (8)	3 (15)	3 (0)	3 (28)
	4 (25)	4 (0)	4 (0)	4 (31)

Some independent variables that account for significant additional variance in mean scores for groups of pupils on each of the attitude scales. Figures in brackets indicate the % variance accounted for.

TABLE 13.2

	Independent Variables (characteristics of individual pupils)			
	Sex (increasing variance)	Social Class (decreasing variance)	AH4 (decreasing variance)	Pretest Attitudes (increasing variance)
Attitude Scales	5 (0)	5 (2)	5 (9)	5 (2)
	1 (0)	1 (2)	1 (6)	1 (7)
	2 (0)	3 (2)	3 (6)	4 (7)
	3 (0)	2 (1)	2 (5)	2 (11)
	4 (4)	4 (1)	4 (1)	3 (13)

Some independent variables that account for significant additional variance in the scores of individual pupils on each attitude scale. Figures in brackets indicate % variance accounted for.

accounted for by *social class* and *AH4*, then the rank order of scales is almost identical in each case (5 or 1, 2, 3, 4). If we examine position statements (see Chapter 5) to which the items of each scale are supposedly related, then we could suggest that there is some relationship between this rank order and the cognitive demands implied by the corresponding statements (the least cognitive demands being associated with Scale 4).

Table 13.2 provides a similar pattern for the within-schools regressions except for one odd placement of scale 4 (5, 1, 2 or 3, 4).

Sex differences in attitude scores

Sex differences can be observed to a greater or lesser extent in some part of the analysis of each of the five attitude scales. In all cases the scores indicate that boys have more favourable attitudes to science than girls.

The sex effect is clearest for Scale 4 where the ANOVA of various sized schools and the ANOVA of medium sized schools each indicate that at the end of S2 boys' mean scores are greater than girls' ($p<0.001$) and that the differences are increasing over the two year period ($p<0.001$). A substantial 25% of the variance in these mean scores is accounted for by sex, and in the regression analysis of individual scores of pupils within schools 4% of the variance is accounted for by sex.

Scale 3 exhibits sex differences in scores for medium sized schools and the differential increases from S1 to S2, but these differences are reduced in small schools and do not exist in large schools. The overall sex effect is sufficient for it to account for 8% of the variance in mean scores, but not sufficient to account for significant variance in individuals' scores.

Each of the two ANOVA's show significant differences between the sexes on post-test scores for Scale 2.

On Scale 1 there are significant differences on post-test scores of schools of various sizes following Integrated Science, and these differences are seen to be increasing with time, but there is no significant sex effect among medium sized schools following two different courses.

The fifth scale provides evidence for a very weak sex effect for medium sized schools only.

If we are to explain these sex differences which are clear for attitudes with strong affective components such as 'Interest and enjoyment in science' but much less striking for attitudes with stronger cognitive components, then we must:

a) examine the ways in which boys and girls differ in personality, or more generally what Heim (1970b) calls 'temperament', and intellectual abilities, and

b) identify which of these differences, if any, might be relevant to the development of the attitudes with which we are concerned.

Gardner (1974a) and Kelly, A (1974) have reviewed the evidence on science-related sex differences in detail. These differences are very clear but to what extent they are determined genetically and to what extent they are the results of social pressures is unknown. It is not difficult to demonstrate that much of the influence in the United Kingdom comes from the forces of our society. For example Kelly provides statistics to support her statement that 'other countries are utilizing their talented women to a far greater extent than Britain'. Such concern for the wastage of scientific womanpower in this country is not new. A decade ago Hunt (1965) was telling us that

> 'In our treatment of women we are not only falling behind countries such as the Soviet Union where 29% of qualified engineers are women, compared with 0.1% in Great Britain, but we are also falling behind newly emergent countries, and behind countries such as Egypt and Turkey where women were until recently, in a completely subordinate position'.

And Meyer (1963) reports that

> 'Australian teachers, unlike those in England, were not faced with tremendous resistance to science found amongst most of the English girls'.

These national differences in the differential sex effect suggest that the influence of socialization processes, probably in both home and school, is crucial and characteristic of particular countries.

It is still possible, as Heim (1970b) suggests, that a major part of

> 'The male-science/female-arts difference may, in all seriousness be a congenital difference rather than a social artefact, but it is hard to tell as long as society continues to treat women as intellectually different from men....the truth of this basic psychological hypothesis cannot be determined until women have had strictly equal vocational, social and educational opportunities *for many generations*' (page 140).

There is one area of sex-differences, spatial ability, where it does seem possible that *some* of the effect may be genetically determined. The consistent empirical findings of boys' higher average test scores are supported by studies of other animals, by studies of the influence of sex hormones on the activation and inhibition of neural processes, and by arguments relating male 'traits' to the greater genetic information available in a set of male chromosomes than in a set of female chromosomes (Stafford 1961, Broverman, Klaiber, Kobayashi and Vogel 1968, Buffery and Gray 1972, Hutt 1972). However, if this difference is to be used to explain the results of this study it would be necessary to demonstrate that spatial ability is related to the development of the particular attitudes being considered. Such a relationship is by no means obvious.

The different social roles imposed on girls are exemplified by Gardner (1974a, page 25)

> 'The connection between social forces and science becomes evident when one considers that (1) the study of science is frequently justified (certainly in the eyes of students) on vocational grounds; (2) the importance of preparation for a vocation has, traditionally at any rate, been perceived as more important for boys than for girls; (3) the popular image of science, particularly physical sciences, as abstract, theoretical, objective and impersonal and relatively unconcerned with humanitarian, artistic and social ends, is inconsistent with the image of girls as dependent, affiliative and nurturant people who ought to be more interested in other people than in things and (4) for boys, however, self-reliance, independence, and achievement-motivation are perfectly consistent with the popular image of science'.

The social environment is clearly a strong potential influence on the attitudes to science of the two sexes.

Whatever the origins of the differences, the general empirical findings on personality and ability are fairly consistent. Boys are found, among other things, to be self-reliant, inclined to act on practical and logical evidence, more interested in theoretical, economic and political matters, and less submissive to authority than are girls who are more imaginative, subjective, acting on intuition, interested in social and aesthetic matters and people, submissive to authority and conforming. The stereotype we have of science does not then reflect the general interests of girls and it is not surprising that they enter secondary school with lower scores on 'Interest and enjoyment in science' than do boys. However, this difference is increased over S1 and S2 and there is an overall decrease in scores, which suggests that school science may have not only done little to improve the image of science for girls but may have added a further antipathetic element.

If we look closely at Scale 3 we see that the position statements and the items (*e.g.* 'Scientists do nothing for me' and 'Everyone in the modern world needs to learn science') are concerned with the need for 'everyone' to feel involved in science. The lack of correspondence between the profile of 'females' and that of 'science' makes such identification with the subject difficult for girls, and, as expected, they score lower than boys on 'awareness of the contribution of science to the economic and social life of the community'.

The influence of sex on scores of Scales 1 and 2 is less than on Scales 3 and 4, nevertheless it is still significant. Since both scales (as defined by the position statements and items) are concerned with science as a school subject, the general negative affect towards science may be reflected in these scales. For example, Scale 2 items, as seen in the factor analysis of Chapter 6, are predominantly concerned with how useful science is to other subjects.

Failure to admit the usefulness of science to the rest of their curriculum appears in line with the girls' general hostile feelings. The fact that the sex differences on a scale of 'awareness of the relationship of science to other aspects of the curriculum' are not as great as those on 'Interest and enjoyment in science' reflects the increased *cognitive* element which seems less likely to differentiate between the sexes.

For Scale 1 the differences are again at a lower level and again the influence of the cognitive component can be seen. It may be the case that the lower interest of girls in the sciences has led to a lower level of attention. As pupils approach the time at which they have to make choices about which subjects they will continue to study they may pay more attention to those of their choice (predominantly science for boys, predominantly arts for girls). It would be expected that pupils giving the more attention to science would be more likely to be aware of what constituted the separate disciplines and the nature of their inter-relationships. The higher scores of boys of 'awareness of the inter-relationships of the different disciplines of science' and the increase of the sex differences as the pupils near the end of S2 (when subject choices are made) is therefore not unexpected.

Their conformity, greater readiness to submit to authority and the tendency to act on intuition rather than on practical or logical evidence, suggests that girls might have lower mean scores on 'objectivity in observation and in assessing observations' than boys. The position statements (Chapter 5) explicitly characterize this objective in terms of the sorts of evidence that are acceptable and the scepticism with which 'true explanations' from 'good scientists' and other authorities (such as the teacher) should be received.

Surprisingly, perhaps, the sex differences on this objective were small. It may be the case that it is the cognitive demands of this objective that are important, and that it is, say, verbal ability that is the crucial influence on responses to the scale items. It is also possible that girls' general submissiveness to authority does not extend to the intellectual sphere in the science classroom. Gardner (1974a) cites an unpublished study (Kent-Smith, 1972) which reported such a finding.

The individual differences among boys and girls are very great and this is reflected in the disappearance on Scale 3, and great reduction on Scale 4, of the variance in attitude predictable by sex when we consider multiple regressions on individual pupil scores in the within-schools analyses, rather than mean scores of boys and girls in the between-schools analyses.

Social class differences

It can be seen from Table 13.2 that the influence of his social class on the attitudes of the individual pupil is significant, but accounts for very little variance (1 or 2%). However, the between-schools analysis (Table 13.1) indicates that the socio-economic status of the catchment area of the school is a very important predictor of mean scores on four out of the five scales, and that it is the cognitively-oriented scales for which this effect is particularly important. The contrast between the within-schools and the between-schools results is striking when these results are compared with those for AH4, the other pupil characteristic which relates most closely to the cognitively-oriented scales.

On the one hand, social class, while only accounting for 1 to 2% of variance on the cognitive-related scales in the within-school analysis, accounts for 15 to 36% variance on these scales in the between-schools analysis. On the other hand AH4 accounts for 5 to 9% of variance on the same scales in the within-schools analysis, and only 0 to 13% on the between-schools analysis.

The social class effect on Scales 1 and 5 (and possibly Scales 2 and 3) which accounts for 32 and 36% of the variance in school mean scores, cannot be interpreted as a mathematical consequence of social class differences among individual pupils for which the corresponding variances accounted for were each 2%. It is necessary then to look for other explanations of the between-school differences that have been accounted for by the school social class variables. In the absence of any further evidence to suggest what the alternative explanations might be we can only speculate.

Differences may arise from teachers' differential expectations of, or behaviour towards, groups of predominantly middle class pupils and groups of predominantly working class pupils. This differential treatment might arise from perceived differences in the pupils' abilities, disciplinary problems, career or 'O' grade aspirations, or cognitive styles. There is no basis on which, from the data of this study, we can identify which of these pupil characteristics might be influential.

Group differences may result from peer group reinforcement. Although the individual pupil's social class only accounts for a small amount of attitude variance, his attitudes may also be influenced by the social class of other pupils with whom he is in contact. For example, if we consider a pupil from social class 4 or 5, from our significant social class within-schools effect, we would predict low scores on Scales 1 and 5. The pupil would not be expected (Scale 1) to distinguish clearly among the separate science subjects (and this, it has been argued, is related to 'O' grade subject choice, see page 407 and Chapter 5, page 138), and his picture of science (Scale 5) will be one of 'true explanations' to be accepted from 'good scientists'. Reinforcement of such attitudes can be expected to depend on whether he is surrounded by pupils with similar attitudes, characteristic of low socio-economic status, or by pupils with personal expectations relating to goals such as 'O' grade and with an image of science as explanations based on objective evidence which we expect to question and modify in the light of new information.

This group reinforcement may not be restricted to peer influences. There may be cultural differences among catchment area communities and these may establish differences in attitudes before entry to secondary school. If this were the case, then we would expect that introduction of pretest attitude scores into the between-schools multiple regression *before* social class would remove a substantial part of the variance now apparently accounted for by the social class variables.

It seems very likely that there are complex interaction effects among these three sorts of influence (and possibly others) and that such interactions lead to distinctive school and classroom climates that are related to the socio-economic status variables.

However, a further possibility is suggested by the social class/AH4 comparison. Whereas the between-schools variance accounted for by social class is greater (on cognitive-oriented scales) than that predicted from the within-schools results, that for AH4 is, if anything, less than expected. The between-schools analysis suggests that social class extracts substantially more variance than AH4 (which was entered later into the regression), but the within-schools analysis strongly indicates the reverse. The correlations between social class and AH4 for *individuals* are low.[1] Therefore it is to be expected that after extracting the social class effect in the within-schools analysis most of the variance attributable to intelligence is left and can be accounted for by AH4.

In contrast the between-schools correlations of mean AH4 with social class variables are more substantial.[2] In this case the social class variables will be expected to take out much of the variance attributable to intelligence leaving little to be accounted for by AH4. It should be possible to evaluate this explanation by carrying out the regressions again but introducing the AH4 variable *before* social class. The variance extracted by AH4 in the *between-schools* analysis may then be increased while that extracted by social class may be reduced. Since the correlations for the *within-schools* case are low there is likely to be little change there. This may well result in removal of the inconsistencies between the variance extracted in the two sets of analyses for these two predictors. Such a situation can arise as a result of entering social class into the regression before AH4, AH4 being a considerably better predictor of cognitive-oriented attitudes than social class, and AH4 being only marginally correlated with social class for individuals but substantially correlated with social class for groups.

---

[1] When social class is assumed to be an interval scale it correlates 0.19 with AH4, see Chapter 8, page 249. The combined r matrix for the within-schools analysis Table 6E Appendix E shows that AH4 correlates 0.12 and -0.09 with the two social class variables.

[2] Table 2E Appendix E shows that for the analysis of 40 schools' mean scores the correlation coefficients between AH4 and the social class variables are 0.40 and -0.29.

Pupils' ability measures (AH4 and divergency)

From Table 13.1 it can be seen that the additional variance accounted for by mean AH4 scores is greatest for Scale 5 (13%). For Scales 1 and 2 this is reduced to 6%, and to 0% for Scales 3 and 4. A similar pattern is found if the variance accounted for in the within-schools analysis of individual pupils is examined, except that substantial variance on Scale 3 is predicted.

The pattern of variance accounted for by AH4 follows the expected path with most variance accounted for on the Scales with strong cognitive components and least on the affective Scale 4.

Scale 3 is something of a problem. Substantial variance (6%) in individual attitude scores was accounted for by AH4 in within-schools analysis. This finding was supported by a similar one in the pilot study (Chapter 5). It was suggested that the *cognitive* component of this attitude was strong. Yet mean scores of groups of pupils on Scale 3 have no substantial prediction from AH4, moreover it was argued in a previous section that girls would view Scale 3 items as a personal identification (*i.e. affective*) with science.

Since mean AH4 scores for groups of pupils did not predict significant variance on this scale we can conclude that:

a) mean achievement of this attitude objective by groups of pupils from different schools is independent of the mean levels of intelligence of those groups, and therefore

b) the achievement of this attitude objective by individual pupils (across all schools) is not related to level of intelligence.

However, the fact that AH4 does predict significant variance on individual pupils' scores on this scale *within* schools suggests that:

> achievement of this attitude objective is related to the level of intelligence of the pupil *relative* to the intelligence of those other pupils in his school making up this sample.

Since the teaching groups in these schools were predominantly mixed ability, we can perhaps go further and suggest that the achievement is related to the pupil's level of intelligence relative to the rest of his science teaching group. We must then presume that the science teacher, by some means during the course, differentiates pupils of various abilities within her class and communicates to those of relatively high intelligence the feeling that science is of concern to them and to the general population.

This general argument may also be used in the interpretation of the significant correlation between AH4 and Scale 3 scores for S2 pupils in the pilot study. The absence of a similar relationship between scores for S1 pupils in that study is consistent with the progressive differentiation of, and communication of feeling to, the more able pupils within the class.

This explanation removes the necessity for the implied hypothesis of the pilot study that Scale 3 must have a strong *cognitive* component. It is consistent with the *post hoc* rationalization of a previous section that sex differences on this attitude resulted from the *affective* nature of the scale. However, it would be desirable to have some other evidence regarding the nature of the scale.

If we look back to Chapter 6 we find that Scale 3 failed to correspond to any dimension of attitudes on which pupils can be seen to differ. However,

Table 6b Appendix B shows that of the six items which define Factor 6 - 'Alienation from Science' - with loadings of 0.30 or more, four are negative items from Scale 3 and each of these displays appropriate moderate correlations with the other items. (The only other factor on which there was a loading above 0.30 for more than one Scale 3 item was the unsatisfactory and uninterpretable Factor 5.)

It appears then that while no overall description of the items of Scale 3 has been provided, there is some evidence that 4 of the Scale's 6 negative items help to define an 'Alienation' or 'Negative Identification with Science' scale of predominantly *affective* nature.

Mean scores on divergency were unable to account for any additional variance (over and above that of AH4 and pretest attitudes) in the between-schools analysis except for Scale 2. Here the three fluency measures predicted an extra 9% of variance.

For the within-schools analysis the introduction of uniqueness in addition to fluency measures resulted in significant, but very small (2%), additional predicted variance on Scale 4 and Scale 5.

In general it seems that measures of divergency are not adding very much to our prediction of attitudes beyond what traditional measures, criticised for their narrowness, are able to do. However, the results for Scale 2 do suggest that groups of pupils who have shown they are able to produce a large number of different ideas in response to a given stimulus, are more able than other groups to show 'awareness of the relationship of science to other aspects of the curriculum'. It is possible that pupils are not having the links between the sciences and other subjects brought to their attention and that these links are not self-evident from the worksheets or text-books. In that event we could surmise the groups with the highest mean divergency scores would be those that, when presented with an item such as 'Geography provides examples of things we learn about in science', would be the most likely to be able to come up with ideas exemplifying that statement, and so to agree with the statement.

Pretest attitude measures

As seen in the results section the best single predictors of post-test attitude measures are the pretest attitude measures. When all the pretest measures are added into the regression equations for each scale after sex, social class and AH4, they account for substantial extra variance in each case. Table 13.1 shows that most variance is accounted for, in the between-schools analysis, on Scale 4 followed by Scales 3, 2, 1 and 5. Table 13.2 indicates that a similar pattern is followed in the within-schools analysis except that substantially less Scale 4 variance is predicted.

These patterns show an increasing effect as we move from those objectives with weaker to those with stronger affective components. For achievement of the latter then, the 'feelings' with which the pupil arrives at school are relatively more important than for the former. However, the variance in 'Interest and enjoyment' accounted for in the within-schools analysis by pretest attitudes was rather less than expected from the between-schools results. It may be that a pupil's interest level is not so much influenced by his *own* attitudes on entry to school as it is by the *general* level of attitudes of the other pupils around him.

The differences between pretest attitude measures and post-test measures are reflected in the consistent ANOVA *occasions* effect. This effect was significant in 33 out of 35 analyses. The two exceptions were both analyses of non-denominational schools outwith the cities (on Scales 3 and 4). For the rest there was clear evidence for increasing attitude scores over the two years on Scales 1, 2, 3 and 5, and decreasing scores on Scale 4.

These changes suggest that where attitude development involves substantial *cognitive* learning, S1 and S2 science teaching is achieveing a measure of attitude improvement. If we consider the significant decreases on Scale 4 scores and the two non-significant *occasions* results for Scales 3 and 4, we have to conclude that where *feelings* about science are concerned we have evidence that the subject has not lived up to the expectations that the pupils had on entry to secondary school.

The within-teaching-groups analysis examined the value of pretest attitude measures as predictors of school (S2) science achievement scores. Not more than 7% achievement variance was predicted (this was accounted for by Scale 5 scores) suggesting that the predictive validity of the attitude scales (for science achievement) is substantially lower than that of intelligence measures such as AH4, 'Meanings of Words' and 'Uses for Things'.

School characteristics

Table 13.3 summarizes the significant effects of the demographic school variables for the various attitude scales. The influence of these variables is marginal except in the case of Scale 2. Here the multiple regression analysis suggests it is the small schools outwith the cities that show the most favourable attitudes. From the ANOVA results there is some evidence

TABLE 13.3

Attitude Scale	School Characteristics	
	Size of School	Denomination/Location of school
1	-	ANOVA (Integrated Science only)
2	M.R. 8%	a) ANOVA: *denomination* x *course* b) M.R. (location of school) 8%
3	ANOVA: *sex* x *size*	-
4	-	-
5	-	-

Summary of significant effects of school characteristics on scores of 5 attitude scales that Roman Catholic Schools (mostly within the cities) that follow separate science courses attain higher mean scale scores than those that follow Integrated Science, while for non-denominational schools outwith the cities higher scores are achieved by those following Integrated Science.

It appears then that groups of pupils in small country schools following Integrated Science are more aware than other pupils of the extent of transfer among their various school subject disciplines. In such schools it is possible that the science teachers actually participate in the teaching of several different subjects, and it is very likely that they would be in closer contact with teachers from other disciplines in the small 'general' staffroom than are the teachers in big schools where the staffroom habitually used may be in the 'science wing' of the school. These factors would be expected to influence the day-to-day opportunities for transfer.

A related, but less substantial, result is indicated by the significant effects for Scale 1 where the noteworthy differences are between those groups attending non-denominational schools outwith the cities and the others. This suggests that the pupils in the country schools following Integrated Science are more aware of the extent of transfer among the sciences.

Differences between individual schools

The significant differences among individual schools were apparent in almost every ANOVA analysis *except for Roman Catholics and girls.* Only for Scale 5 was there evidence of a *school* effect when scores for these two groups were analysed.

These results do not necessarily suggest that, for groups of pupils in Roman Catholic schools, schooling (or science teaching) has no effect on attitudes to science, but what we can say is that there is no significant variation among the individual Roman Catholic Schools as compared with differences between groups of different sex within schools. The groups of pupils entering Roman Catholic schools exhibit a uniformity of attitudes (possibly reflecting pressures from their religious or social backgrounds), and this uniformity is maintained through to the end of S2, suggesting that the science teaching, insofar as it influences attitude development, exhibits a high degree of homogeneity.

The lack of a *school* effect on the scores of girls together with the significant effect for boys suggest that the variations between schools are unimportant compared with the major determinants of girls' attitudes to science. It appears that groups of girls enter secondary school with a uniformity of attitudes which the science teaching of S1 and S2 does little to change.

For groups of boys and pupils in non-denominational schools, the differences between individual schools are clear. However, examination of the post-test ANOVA tables for each attitude scales shows that only for Scale 4 scores is more than 3% of the variance among groups accounted for by unsystematic between-school differences (for 'Interest and enjoyment' scores in various sized schools following Integrated Science this reaches the 10% level).

Type of science course

The only clear *course* effect appears in the scores of Scale 1. The ANOVA analysis shows differences developing over the two years with separate science pupils displaying the more favourable attitudes at the end of S2. In the between-schools multiple regression substantial (11%) additional variance was accounted for by this variable.

The objective 'awareness of the inter-relationship of the different disciplines of science' was included in Curriculum Paper 7 with the intention that pupils following the integrated course would be at no disadvantage, compared with those following the separate science courses, with regard to recognizing what constituted biology, chemistry and physics. Furthermore, it was intended that these pupils would have an advantage over the separate subject pupils in that the early sections of Curriculum Paper 7 would enable them to appreciate the inter-relationships between the sciences (see Chapter 5, page 139). From the observed outcome, it must be concluded that Integrated Science pupils have not developed this awareness to the same extent as the separate science pupils, and it seems likely that achievement of this objective has rested largely on an appreciation of what part of science makes up biology or chemistry or physics. Apparently such an appreciation is developed more adequately in a separate subjects course, which is not very surprising. However, it does indicate that those teaching the integrated course are not teaching adequately for transfer despite their greater number of opportunities. It has already been argued (Chapter 4, page 125) that transfer between school subjects is unlikely to 'just happen', a strategy of conscious and explicit attention to the feature to be transferred must be adopted. It must be concluded that teachers of Integrated Science are not taking adequate advantage of the potential transfer situations that Curriculum Paper 7 sees in the early sections of the syllabus, in order to bring out the inter-relationships between the sciences.

Ability grouping, sex grouping, number of science periods and number of pupils in teaching group

Table 13.4 shows that scores on Scale 5 are the only ones for which significant extra variance (3%) is accounted for by more than one of these variables. It appears that pupils in mixed sex, but not mixed ability, classes with larger numbers of periods display the most faviourable attitudes.

The only other scale for which significant extra variance is accounted for by any of these variables is Scale 4. Here it seems that pupils in mixed ability classes show less interest in science than the others.

There is no evidence that the objectives of Curriculum Paper 7 have been achieved to a greater degree by pupils in mixed ability classes than those in banded classes despite the fact that the course was designed specifically for the former. There is indeed some evidence for the reverse effect, for 'Interest and enjoyment in science' and 'an objectivity in observation and in assessing observation', although the variance accounted for by these variables is not very great.

Sex grouping (mixed sex groups displaying higher scores) and the number of periods allocated to science also account for a small amount of variance in 'objectivity', but again the value of these variables in predicting class attitudes is somewhat limited.

The number of pupils in the teaching group was unable to account for significant extra variance on any scale. We can conclude that, within the range of class sizes in this study (largely between 10 and 30), insofar as the teachers are attempting to teach towards attitudes, the number of pupils in their teaching group in itself does not influence the affective

TABLE 13.4

	Teaching Group Variables			
Attitude Scale	Ability Grouping	Sex Grouping	No. of Science Periods	No. of Pupils in Teaching Groups
1	-	-	-	-
2	-	-	-	-
3	-	-	-	-
4	3%	-	-	-
5	3%	3%	3%	-

Additional variance accounted for by introduction of teaching group variables to multiple regressions corresponding to each attitude scale.

achievement.

Teachers ratings and school achievement measures

It was argued in Chapter 12 that the relationships between attitude scores at the end of S2 and teachers' ratings of pupils' 'academic ability' while not very strong:

a) were better than chance,

b) were considerably stronger than the relationships between corresponding variables measured on *entry* to secondary school, and

c) showed the most striking change in median correlation (0.03 to 0.36) over the two year period for the relationship between the ratings and pupils' scores on Scale 4 (Interest and enjoyment in science).

If we assume that the teachers' ratings are a measure of their expectations of the pupils in their teaching group, then it appears that pupils' attitudes have moved into line with their teachers' expectations of them over the two year period. The change is most marked for the attitude that is least related to 'academic ability'. These findings are interesting in the light of other studies (Rosenthal and Jacobson 1966, Beez 1968) which appear to have shown that pupils' performances on cognitive tasks are influenced by their teachers' expectations, and that the behaviour is in accordance with the expectation.

The extent to which teachers were able to assess the level of pupils' interest is reflected in the correlation coefficient of 0.39 between teachers' ratings of pupils on 'interest in science' and pupils' score on the 'Interest and enjoyment in science' scale. Correlations of these ratings with the other attitude scales were somewhat lower. In contrast the correlation coefficients between 'interest' ratings and 'academic ability' ratings or school achievement marks were very high (r = 0.85 and 0.82). This suggests that teachers base their assessments of pupils' *interest* in science largely on *academic performance* rather than other behaviour that might be indicative of pupils' attitudes.

Jackson (1968 Chapter 2) reports similar findings. He found that teachers' predictions of pupils' responses to a questionnaire assessing attitudes towards school when correlated with those responses yielded a coefficient of 0.35. However, when he related the teachers' predictions to attainment in reading, arts and arithmetic he found that

> 'the teachers' estimates of their students' responses to a school opinionnaire turn out to be more closely related to the students' academic standing than to their actual responses to the questionnaire'.

The results of both this study and that of Jackson indicate that teachers can predict *some* of the variation in interest scores and that some teachers are very much better at doing this than others. However, there is no reason to suggest that, at this time, teachers are able to provide *adequate* assessment of pupils' attitudes.

It may be,as Jackson suggests, that these ratings by teachers are based on the belief that 'doing well' and 'having favourable attitudes' covary. However, while this may be the case for extreme levels of achievement and extreme attitudes, the correlations of the five attitude scores with science achievement in this study (0.32 to 0.36) suggest that such a relationship is by no means established across the whole range of achievement and attitude. The concurrent validity for these attitude scales, as measures of science achievement, could not be considered as adequate.

Substantive Hypotheses and Related Findings

The following summarizes the findings related to the various hypothesised differences in achievement of the affective objectives.

Differences hypothesised on the basis of theoretical differences between the Integrated Science course and separate science subject courses (see Chapter 4, page 135)

Hypothesis 1a: Pupils following the Integrated Science course will achieve *awareness of the contribution of science to the economic and social life of the community* to a greater degree than will pupils following three separate science subject courses.

No empirical evidence was found to support this hypothesis.

Hypothesis 2a: Pupils following the Integrated Science course will achieve a higher level of *Interest and enjoyment in science* than will pupils following three separate subject science courses.

No empirical evidence was found to support this hypothesis.

Differences hypothesised on the basis of a review of the empirical findings of other studies (see Chapter 7, page 218)

Hypothesis 1b: Boys will achieve a higher level of *Interest and enjoyment in science* than will girls.

There was clear evidence (summarized on page 402) to support this hypothesis.

Differences hypothesised on the basis of results from the pilot study (see Chapter 5, pages 168 and 169)

Hypothesis 1c: Pupils following the Integrated Science course will achieve a higher level of *awareness of the relationship of science to other aspects of the curriculum* than will pupils following three separate science subject courses.

No empirical evidence was found to support this hypothesis.

Hypothesis 2c: Pupils following the Integrated Science course will achieve a higher level of *objectivity in observation and in assessing observations* than will pupils following three separate science subject courses.

No direct empirical evidence was found to support this hypothesis. (However, there was some indication that while pupils entering Integrated Science courses had less favourable attitudes on this scale than those entering the separate science courses, by the end of S2 these differences had disappeared.)

Hypothesis 3c: Boys will achieve a higher level of *awareness of the* *contribution of science to the economic and social life of the community* than will girls.

Substantial evidence (summarized on page 402) was found for higher achievement by groups of boys in all but large schools following Integrated Science. However, no significant variance in individual pupil's scores is accounted for by sex.

Hypothesis 4c: Boys will achieve a higher level of *Interest and enjoyment in science* than will girls.

There was clear evidence (summarized on page 402) to support this hypothesis.

Hypothesis 5c: Pupils' scores on a test of general intelligence (AH4) will predict substantial variance at the end of S2 in their scores on *awareness of the contribution of science to the economic and social life of the community*.

There was no evidence that achievement of this objective by an individual pupil was related to intelligence scores. There was some evidence (see page 412) that the achievement was related to the pupil's level of intelligence *relative* to that of the rest of his group, this relationship accounting for 6% of the variance of attitude scores.

Hypothesis 6c: Pupils' scores on a test of general intelligence (AH4) will predict substantial variance at the end of S2 in their scores on *objectivity in observation and in assessing observations*.

AH4 mean scores were able to account for 13% of the variance in mean scores on this attitude among groups of pupils of each sex from different schools. Individual AH4 scores were able to account for 9% of variance in individual pupil's attitude scores. The hypothesis is, therefore, supported.

Research Questions and Related Empirical Findings

The following provides a brief summary of those findings from the empirical study that are related to the research questions identified in Chapter 7.

(i) How does the achievement of affective objectives over S1 and S2 relate to the different types of science course experienced by the pupils?

The way in which the knowledge is organized (as Integrated Science or separate disciplines) appears, in itself, to be largely unrelated to the achievement of affective objectives or to the patterns of cognitive preferences exhibited by pupils. The only exception to this general finding is that pupils following separate sciences are more aware of the inter-relationships between the sciences, probably because they have a clearer picture of the distinctions between biology, chemistry and physics.

(ii) How is the achievement of affective objectives related to characteristics of the school such as its size, its location, its denominational status, the socio-economic level of its catchment area, and the average level of ability of its intake?

The first three of these characteristics seem to be unrelated to the formation of attitudes to science except that awareness of the relationships of science with other subjects of the curriculum is most likely to develop among groups of pupils in small country schools.

The socio-economic level of the school's catchment area is related to achievement of all but the 'purely affective' objective. On the cognitively-oriented scales it is the groups of higher socio-economic status that display the most favourable attitudes.

The average level of intelligence is substantially related to those objectives that appear to have strong cognitive elements, but not to interest. We would predict that a group of pupils with a high mean I.Q. would have a high level of awareness of the relationships among the sciences and other subjects in the curriculum, and would display an objectivity in observation, but would not necessarily show a high level of enjoyment or interest in science.

The level of divergent ability of the group provides useful further information on affective achievement only for the group's awareness of the relationship of science to other aspects of the curriculum.

(iii) How are the achievement of affective objectives related to characteristics such as class size, ability grouping, and number of periods of science?

Class size (range approximately 10-30) appears to have no influence on the development of attitudes to science.

Pupils in banded classes have marginally better scores on 'Interest and enjoyment in science' and 'objectivity' than pupils in mixed ability groups but the effect is very weak.

The number of periods allocated to science (most schools have 4, 5 or 6) does not appear to influence affective achievement except for the development of the scientific attitude of objectivity (once again the effect is very small).

(iv) *How is the achievement of affective objectives related to pupils' sex, social class, general intelligence measures, divergency measures, perceived achievement in science, and teachers' ratings?*

Only for scores on interest in science is there significant (4%) variance of individual pupils' scores (within-schools) accounted for by sex. However, *group mean* scores are greater for boys on all the scales. Pupils' social class accounts for significant variance in individual scores on all five attitude scales but this is very low (1 or 2%).

Pupils' levels of intelligence appear to influence their attitudes related to those scales with strong cognitive components but not to interest and enjoyment in science. The pupil's personal indentification with science seems dependent not on his absolute level on intelligence but his level relative to the rest of his group.

Information on pupils' levels of divergent ability provides little further useful information on attitude achievement, except for minimal (2%) variance accounted for on scores for interest and objectivity.

Pupils' attitude scores are not closely related to their achievement in S2 science. Nor is there a strong relationship between teachers' ratings on pupils' academic ability and pupils' attitudes, but there is evidence that these attitudes (particularly interest) are moving into line with the teachers' expectations during the two years.

Pupils' attitude scores are not strongly related to teachers' ratings of pupils' interest (however, these *interest* ratings are strongly related to pupils' *achievement*).

(v) How are teachers' behaviour and attitudes related to pupils' achievement of affective objectives?

No clear attitude dimensions, closely related to behaviour in the science classroom, and on which teachers could be seen to differ were identified. This together with the weak or zero relationships found between teachers' and pupils' attitudes in other studies suggest that there may be no general ideological dimensions by which science teachers can be characterized and which also exert strong influences on pupils' affective achievement.

(vi) Do pupils' attitudes to science change over the first two years in secondary school and, if so, in what direction? To what extent are attitudes assessed on entry to secondary school valuable as predictors of attitudes held at the end of two years?

In general pupils' attitudes to science show significant changes over the two year period. For attitudes involving substantial *cognitive* learning, pupils' scale scores increase, for those elements concerned with *feelings* about science the scores decrease. Attitudes assessed on entry to secondary school appear to be the best predictors we have of the attitudes of pupils at the end of S2.

General Conclusions

Unless a curriculum incorporates a very large number of objectives, it is usually not possible to comprehend from the objectives alone what the course is designed to do. The five attitude objectives of Curriculum Paper 7 do not, in isolation, disclose the purposes of, or the value of, the general affective goals that the Working Party had in mind. We therefore have to look to the Curriculum Paper itself for expansion and explanation of their intentions in this area of science education.

A close examination of the Paper has revealed no explicit rationale for inclusion of these objectives. It appears that the Working Party considered that what they were attempting to achieve with each objective would be self-evident. Unfortunately this is not the case, and it is difficult to say which it was, of the various categories of argument for teaching towards affective goals, that they had in mind. It is even more difficult, in the absence of an explicit justification for these broad attitude goals, to assess their worth.

In the absence of an explicit rationale for the affective goals, the rationale for the presentation of science in an integrated form and the meanings ascribed to integration were examined to see if there was evidence for an implicit justification for the five attitude objectives.

Insofar as Curriculum Paper 7 adopts a meaning of integration in the 'strong-thesis' sense (*i.e.* science as an integral part of the single conceptual structure of all knowledge), it has been argued that the following objectives describe attitudes that are logically related to that meaning:

a) awareness of the inter-relationship of the different disciplines of science,

b) awareness of the relationship of science to other aspects of the curriculum,

c) awareness of the contribution of science to the economic and social life of the community.

If integration if viewed in the 'weak thesis' sense (*i.e.* science as a discipline distinct from other forms of knowledge but without component disciplines) then only (a) above is logically related to its meaning. If integration is conceptualized as the 'interdisciplinary study' of several component sciences, then none of the objectives relate logically to the meaning.

Curriculum Paper 7 implies each of the three meanings of integration at different places in the text and seems undecided about what conceptual structure for science it is adopting. Such a decision would not be easy, appropriate structures are not found on trees. Eggleston (1974) points to the possibility

> 'that the struggles for survival, which successive versions of general science, combined science, and integrated science have faced might be related to the difficulties of providing an adequate map [conceptual structure] without making impossible demands on the navigator' (page 119).

If the content of the syllabus of the course is examined it should be possible to draw some implications about the conceptual structure adopted. In this case the content reflects 'interdisciplinary study' and so fails to provide a justified framework for any of the attitude objectives.

The arguments put forward for the adoption of the integrated approach were also scrutinized, and three of the intended outcomes for which some justification was provided were found to be directly related to three of the attitude objectives:

a) awareness of the contribution of science to the economic and social life of the community,

b) Interest and enjoyment in science,

c) an objectivity in observation and in assessing observations.

It appears that by expending some effort most of the attitude objectives can be fitted somewhere into the rationale for the Integrated Science course (provided we forget that the content of the course operationalizes 'integration' as 'interdisciplinary study').

However, this by no means provides a theoretical framework that is adequate for clarifying the purposes and meanings of the attitude objectives.

In an attempt to provide such clarification the five objectives were related to Krathwohl *et al's* (1964) classification of affective objectives and Klopfer's (1971) categories of science-related attitudes, interests and orientation. Both these systems aim to classify and clarify affective behaviours and objectives that are commonly used by curriculum developers and teachers.

Each of the first three objectives can be readily placed in Krathwohl's lowest affective category of Awareness, but although each corresponds in some way to Klopfer's Orientation categories, lack of specification of the *cognitive* demands of the objectives prevents precise placement.

The fourth objective conflates three categories in each of Krathwohl's and Klopfer's system. This reflects the lack of precision in the objective which fails to distinguish among behaviours from Willingness to Respond to Commitment, among interest in science as a school subject, other science activities and science as a career. The fifth objective is directly related to one of Klopfer's categories 'Adoption of scientific attitudes'. This category itself may imply behaviour at any of about six Krathwohl levels (the minimum level being that of Acceptance of a Value).

The findings of the empirical study illustrate the relative usefulness of various types of independent variables in predicting attitude scores. In general, the demographic and organizational variables of school, science course and science class (size, location and denomination of school, mixed ability, mixed sex and integrated or separate science arrangements of classes, number of periods allocated to science and number of pupils in the teaching group) are unrelated, or weakly related, to attitude scores. Although teachers and others habitually blame schools that are too large, or that are situated in inner-city areas, or that give too few periods to science, or that have over-large classes, for unfavourable attitudes among pupils, it appears that these factors *per se* do not exert any substantial influence.

Much more important are variables related to characteristics of the pupil such as sex, social class, intelligence (but not divergency) and attitudes developed before entry to secondary school.

Although the relationship between teachers' attitudes and pupils' attitudes, if it exists at all, is elusive, there is an indication that teachers' expectations of their pupils academic performance have an influence on attitude development (particularly interest in science). That the teachers are specifically teaching towards attitude objectives seems very doubtful. They are not able to distinguish between pupils' academic ability and level of interest. Their inability to recognize pupil behaviours that reflect levels of interest suggests that they are not consciously teaching towards interest goals.

This study was not designed to investigate the influence of the teacher on pupils' attitudes to science, and the general plan militates against observation of any such effects since the intention was to look at the differential affective achievement among groups of pupils characterized by large scale demographic and organizational variables. The latter has been anything but a fruitful area, and the most striking aspect of the differences in attitudes among *schools* is the very low proportion of variance that is accounted for by those variables controlled by the system of secondary education. The major part of the predicted variance is accounted for by factors outwith the control of the secondary school *i.e.* characteristics that the groups of pupils have on entry to the school. If we consider the differences among *pupils*, however, the proportion of variance that we can account for by these various characteristics is low. The major part of the variation in attitudes is still unaccounted for.

If our interest is in the influence of *educational* factors on attitudes to science, then it appears that we must now look in detail at those factors that have an impact on the pupil inside the science classroom rather than the general and organizational variables of schools and classes. It is an examination of the nature of the teacher/pupil interaction in the science classroom, the teacher's patterns of communication with individual pupils and groups of pupils, the transmission of the teacher's expectations to the pupils, the particular topics that are covered in the lessons, the strategies and tactics within strategies adopted by the teacher, that we can be hopeful will provide educational variables that will be more profitable in explaining differences in attitudes to science among pupils than have those used in this study.

With hindsight it can be said that selection of appropriate independent variables on the basis of the 'intuition' of, and 'what looked interesting' to, teachers and HMI's was probably an unwise procedure. In an area as uncharted as that of achievement of affective objectives, either a substantial exploratory study or a guiding theoretical framework is essential if the most fruitful paths for investigation are to be mapped out.

In conclusion, let us return to White's criteria for curriculum evaluation on page 1 of this thesis, and see if the attitude objectives of Curriculum Paper 7 have 'got their tick'.

The clarity of the objectives and the adequacy of the specification of 'appropriate means to achieve them' are doubtful. However, the process of relating them to Krathwohl's and Klopfer's categories suggests that by specifying some affective and cognitive behaviours and making some distinctions between categories, a group of clear objectives, readily related to classroom procedures, could be provided.

The educational respectability of the objectives is not necessarily in doubt, but up till now, this has not been established by adequate justification. This reflects the lack of a theoretical framework for the objectives, or, indeed, for the integrated course.

The changes in attitudes of pupils over the two years suggests that the objectives, to some extent, 'connect with the abilities of those pupils for whom it is designed'.

There is no evidence to suggest that this curriculum (Integrated Science) is 'more efficient than rivals [separate sciences] in the field' of attitude objectives.

Regrettably, at present - no tick.

REFERENCES

Aiken, L.R. and Aiken, D.R. (1969) 'Recent Research on Attitudes Concerning Science' *Science Education*, 53, pp. 295-305.

Alexander, D. (1972) *Nuffield Secondary Science Evaluation*, Schools' Council, London.

Allport, G.W. (1961) 'Values and our Youth', *Teachers College Record*, 63, pp. 211-219.

Allport, G.W., Vernon, P.E. and Lindzey, G. (1951) *Study of Values: A Scale for Measuring the Dominant Interests in Personality* (1951 rev.), New York: Houghton Mifflin.

Alvord, D.J. (1972) 'Achievement and Attitude', *The Science Teacher*, 39 (4), pp. 36-38.

Anderson, G.J. and Walberg, H.J. (1968) 'Classroom Climate and Group Learning', *International Journal of the Educational Sciences*, 2, pp. 175-180.

Anderson, K.E., Page, T.C. and Smith, H.A. (1958) 'A Study of the Variability of Exceptional High School Seniors in Science and other Academic Areas', *Science Education*, 42, pp. 42-59.

Anderson, N.H. (1963) 'Comparison of Different Populations: Resistance to Extinction and Transfer', *Psychological Review*, 70, pp. 162-179.

Atwood, R.K. (1967a) 'A Cognitive Preference Examination using Chemistry Content', *Journal of Research in Science Teaching*, 5, pp. 34-35.

Atwood, R.K. (1967b) 'CHEM Study Achievement Among Groups Classified by Cognitive Preference Scores', *Journal of Research in Science Teaching*, 5, pp. 154-159.

Atwood, R.K. (1969) 'Change in Cognitive Preferences of High School Physics Students', *School Science and Mathematics*, 69, pp. 697-699.

Atwood, R.K. (1971) 'Development of a Cognitive Preference Examination Utilizing General Science and Social Science Content', *Journal of Research Science Teaching*, 8, pp. 273-275.

Ausubel, D.P. (1963) 'Some Psychological Considerations in the Objectives and Design of an Elementary School Science Program', *Science Education*, 47 (3), pp. 278-284.

Ausubel, D.P. (1968) *Educational Psychology*, New York: Holt, Rinehart and Winston.

Baker, B.O., Hardyck, C.D. and Petrinovich, L.F. (1966) 'Weak Measurements vs. Strong Statistics: An Empirical Critique of S.S. Steven's Prescriptions on Statistics', *Educational and Psychological Measurement*, 26, pp. 291-309.

Barber, B. (1962) *Science and the Social Order*, New York: Collier.

Barnett, H.C. (1974) 'An Investigation of Relationships Among Biology Achievement, Perception of Teacher Style, and Cognitive Preferences', *Journal of Research in Science Teaching*, 11, pp. 141-147.

Beatty, W.H. (1969) 'Emotion: the Missing Link in Education', in Beatty, W.H. (ed.) *Improving Educational Assessment and an Inventory of Measures of Affective Behaviour*, Washington: NEA.

Beez, W.V. (1968) 'Influence of Biased Psychological Reports on Teacher Behavior and Pupil Performance' in Morrison, A. and McIntyre, D. *Social Psychology of Teaching*, Harmonsdworth, Middlesex: Penguin Books.

Bennet, J.F. and Hays, W.L. (1960) 'Multidimensional Unfolding: Determining the Dimensionality of Ranked Preference Data', *Psychometrika*, 25, pp. 27-43.

Bernstein, B. (1971) 'On the Classification and Framing of Educational Knowledge' in Young, M.F.D. (ed.) *Knowledge and Control*, London: Collier-Macmillan.

Bixler, J.E. (1958) *The Effect of Teacher Attitudes on Elementary Children's Science Information and Science Attitude*, unpublished Ph.D. thesis, Stanford University.

Bloom, B.S., Englehart, M.D., Furst, E.J., Hill, W.H. and Krathwohl, D.R. (1956) *Taxonomy of Educational Objectives Handbook I Cognitive Domain*, London: Longmans.

Bloom, B.S., Hastings, J.T. and Madaus, G.F. (1971) *Handbook on Formative and Summative Evaluation of Student Learning*, New York: McGraw Hill.

Blum, A. (1973) 'Towards a Rationale for Integrated Science Teaching' in *New Trends in Integrated Science Teaching*, Volume 2, Paris: UNESCO.

BMDX64 General Linear Hypothesis (1972) *Biomedical Computer Programs*, Berkeley: University of California Press.

Bradbury, F.R. (ed.) (1969) *Words and Numbers. A Students' Guide to Intellectual Methods*, Edinburgh: The University Press.

Brandwein, P.F. (1951) 'Selection and Training of Future Scientists', *Science Education*, 35, pp. 251-253.

Broudy, H.S. (1971) 'Democratic Values and Educational Goals', in *70th Yearbook of NSSE*, McClure, R.M. (ed.), Chicago: University of Chicago Press.

Broverman, D.M., Klaiber, E.L., Kobayashi, Y. and Vogel, W. (1968) 'Roles of Activation and Inhibition in Sex Differences in Cognitive Abilities', *Psychological Review*, 76, pp. 328-331.

Brown, M.M. (1953) 'Some Educational Influences on the Choice of Science Career by Grammar School Girls', *British Journal of Educational Psychology*, 23, pp. 188-195.

Bruner, J.S. (1960) *The Process of Education*, Cambridge Mass: Harvard University Press.

Buffery, A.W.H. and Gray, J.A. (1972) 'Sex Differences in the Development of Perceptual and Linguistic Skills', in Ounsted, C. and Taylor, D.C. (eds) *Gender Differences: Their Ontogeny and Significance*, Churchill.

Burt, C. and Banks, C. (1947) 'A Factor Analysis of Body Measurements for British Adult Males', *Annals of Eugenics*, 13, pp. 238-256.

Butcher, H.J. (1965a) *Sampling in Educational Research*, Statistical Guides in Educational Research No. 3, Manchester: Manchester University Press.

Butcher, H.J. (1965b) 'The Attitudes of Student Teachers to Education', *British Journal of Social and Clinical Psychology*, 4, pp. 17-24.

Butcher, H.J. (1968) *Human Intelligence: Its Nature and Assessment*, London: Methuen and Company.

Butcher, H.J. (1972) 'Divergent Thinking and Creativity' in *Advances in Educational Psychology I*, Wall, W.D. and Varma, V.P. (eds) London: University of London Press.

Butcher, H.J., Ainsworth, M. and Nesbitt, J.E. (1963) 'Personality Factors and School Achievement: a Comparison of British and American Children', *British Journal of Educational Psychology*, 33, pp. 276-285.

Butcher, H.J. and Pont, H.B. (1968) 'Opinions about Careers Among Scottish Secondary School Children of High Ability', *British Journal of Educational Psychology*, 38, pp. 272-279.

Cattell, R.B. (1957) *Personality and Motivation: Structure and Measurement*, New York: World Book Company.

Campbell, D.T. (1963) 'Social Attitudes and Other Acquired Behavioural Dispositions', in S. Koch (ed.) *Psychology: a Study of a Science, Volume 6, Investigations of Man as Socius: Their Place in Psychology and the Social Sciences*, McGraw-Hill.

Carlson, E.R. (1956) 'Attitude Change Through Modification of Attitude Structure', *Journal of Abnormal and Social Psychology*, 52, pp. 256-261.

Cattell, R.B. (1966) *Handbook of Multivariate Experimental Psychology*, pp. 174-243, Chicago: Rand McNally.

Childs, D. (1970) *The Essentials of Factor Analysis*, London: Holt, Rinehart and Winston.

Christie, T. (1969) 'Contrary Imaginations at Keele: an Alternative Interpretation', *Universities Quarterly*, 24, pp. 91-92.

Clarke, C.O. (1972) 'A Determination of Commonalities of Science Interests held by Intermediate Grade Children in Inner-city, Suburban and Rural Schools', *Science Education*, 56, pp. 126-136.

Cline, V.B., Richard, J.M. and Abe, C. (1964) 'Predicting Achievement in High School Science with a Biographical Information Blank', *Journal of Experimental Education*, 32, pp. 393-399.

Comber, L.C. and Keeves, J.P. (1973) *Science Education in Nineteen Countries*, New York: Wiley.

Cook, W.W., Leeds, C.H. and Callis, R. (1951) *The Minnesota Teacher Attitude Inventory*, New York: Psychological Corporation.

Cooley, W.W. and Reed, H.B. (1961) 'The Measurement of Science Interests: an Operational and Multidimensional Approach', *Science Education*, 45, pp. 320-325.

Coombs, C.H. (1964) *A Theory of Data*, New York: John Wiley and Sons Inc..

Cronbach, L.J. (1951) 'Coefficient Alpha and the Internal Structure of Tests', *Psychometrika*, 16, pp. 297-333.

Cronbach, L.J. (1960) *Essentials of Psychological Testing*, New York: Harper.

Cronbach, L.J. and Furby, L. (1970) 'How We should Measure "Change" - or should We?', *Psychological Bulletin*, 74, pp. 68-80.

Cropley, A.J. and Clapson, L. (1971) 'Long Term Test-retest Reliability of Creativity Tests', *British Journal of Educational Psychology*, 41, pp. 206-208.

Crowther Report 1959 Central Advisory Council for Education (England), *15 to 18*, p. 263, London: HMSO.

Culbertson, F.M. (1957) 'Modification of an Emotionally Held Attitude Through Role Playing', *Journal of Abnormal and Social Psychology*, 54, pp. 230-233.

Daniels, S.L. (1956) 'Search for Science Talents Among High School Freshmen', *Science Education*, 50, pp. 385-391.

De Fleur, M.L. and Westie, F.R. (1963) 'Attitude as a Scientific Concept', *Social Forces*, 42, pp. 17-31.

Du Bois, P.H. (1957) *Multivariate Correlational Analysis*, New York: Harper.

Educational Testing Service (undated) *Cognitive Preference Test, High School Physics*, Princeton and Berkley.

Edwards, A.L. (1957) *Techniques of Attitude Scale Construction*, New York: Appleton-Century-Crofts.

Edwards, A.L. and Kenney, K.C. (1946) *A Comparison of the Thurstone and Likert Techniques of Attitude Scale Construction*, *Journal of Applied Psychology*, 30, pp. 72-83.

Eggleston, J.F. (1974) 'Principles of Curriculum Design', in Sutton, C.R. and Haysom, J.T. (eds) *The Art of the Science Teacher*, Science Teacher Education Project, London: McGraw-Hill.

Eiss, A.F. and Harbeck, M.B. (1969) *Behavioural Objectives in the Affective Domain*, National Science Supervisors Association, Washington DC.

Elashof, J.D. (1969) 'Analysis of Covariance: a Delicate Instrument', *American Educational Research Journal*, 6, pp. 383-401.

Elliott, W.E. (1971) 'Perceptions of High School Physics and Physics' Teachers', *Physics Teacher*, 9, pp. 33-36.

Evans, S.H. and Anastasio, E.J. (1968) 'Misuse of Analysis of Covariance when Treatment Effect and Covariate are Confounded', *Psychological Bulletin*, 69, pp. 225-234.

Ferguson, G.A. (1966) *Statistical Analysis in Psychology and Education*, London: McGraw-Hill.

Festinger, L. (1957) *A Theory of Cognitive Dissonance*, Evanston, Illinois, Row, Peterson.

Festinger, L. (1964) 'Behavioural Support for Opinion Change', *Public Opinion Quarterly*, 28, pp. 404-417.

Festinger, L.A. and Carlsmith, J.M. (1959) 'Cognitive Consequences of Forced Complients', *Journal of Abnormal and Social Psychology*, 58, pp. 203-210.

Fiasca, M.A. (1970) 'Integration of the Sciences', *American Biology Teacher*, 32 (4), pp. 225-226.

Field, T.W. and Cropley, A.J. (1969) 'Cognitive Style and Science Achievement', *Journal of Research in Science Teaching*, 6, pp. 2-10.

Finlayson, D.S. and Cohen, L. (1967) 'The Teacher's Role: a Comparative Study of College of Education Students and Head Teachers', *British Journal of Educational Psychology*, 37, pp. 22-31.

Fishbein, M. (1967) 'Attitude and the Prediction of Behaviour' in M. Fishbein (ed.) *Readings in Attitude of Theory and Measurement*, Wiley.

Gagné, R.M. (1970) *The Conditions of Learning*, New York: Holt, Rinehart and Winston.

Gardner, P.L. (1974a) 'Sex Differences in Achievement Attitudes and Personality of Science Students: a Review', paper presented at the Fifth Australian Science Education Research Association Conference.

Gardner, P.L. (1974b) 'Research on Teacher Effects: Critique of a Traditional Paradigm', *British Journal of Educational Psychology*, 44, pp. 123-130.

Gardner, P.L. (1974c) 'Pupil Personality, Teacher Behaviour and Attitudes to a Physics Course', in Musgrave, P.W. (ed.) *Contemporary Studies in the Curriculum*, Sydney: Angus and Robertson.

Gardner, P.L. (1975) 'Attitudes to Science: A Review', *Studies in Science Education*, 2, pp. 1-41.

Gardner, P.L. (in print) 'Scales and Statistics: A Review', *Review of Educational Research*.

Gauld, C.F. (1973) 'Science, Scientists and "Scientific Attitudes"', *The Australian Science Teachers Journal*, 19 (3), pp. 25-32.

Getzels, J.W. (1969) 'A Social Psychology of Education' in Lindzey, G. and Aronson, E. (eds) *The Handbook of Social Psychology*, Volume 5, Don Mills, Ontario: Addison-Wesley.

Getzels, J.W. and Jackson, P.W. (1962) *Creativity and Intelligence*, New York: Wiley.

Getzels, J.W. and Jackson, P.W. (1963) 'The Teacher's Personality and Characteristics' in Gage, N.L. (ed.) *Handbook of Research on Teaching*, Chicago: Rand McNally.

Glaser, R. (1970) 'Discussion of Messick's (1970) Paper' in Wittrock, M.C. and Wiley, D.E. (eds) *The Evaluation of Instruction - Issues and Problems*, p. 218, Chicago: Holt, Rinehart and Winston.

Glick, O. (1970) 'Sixth Graders Attitudes toward School and Interpersonal Conditions in the Classroom', *Journal of Experimental Education*, 38, pp. 17-22.

Goldman, R.J. (1964) 'The Minnesota Tests of Creative Thinking', *Educational Research*, 7, pp. 3-14.

Greenblatt, E.L. (1962) 'An Analysis of School Subject Preferences of Elementary School Children in the Middle Grades', *Journal of Educational Research*, 55, pp. 554-560.

Gribble, J.H. (1970) 'Pandora's Box: The Affective Domain of Educational Objectives', *Journal of Curriculum Studies*, 2 (1), pp. 11-24.

Guilford, J.P. (1954) *Psychometric Methods*, New York: McGraw-Hill.

Guilford, J.P. (1956) 'The Structure of the Intellect', *Psychological Bulletin*, 53, pp. 267-293.

Guttman, L. (1945) *Questions and Answers about Scale Analysis*, Research Branch, Information and Education Division, Army Service Forces, Report D-2.

Guttman, L. (1950) 'The Basis for Scalogram Analysis' in Stouffer, S.A. (ed.) *Measurement and Prediction*, Princeton, New Jersey: Princeton University Press.

Hallworth, H.J. and Brebner, A. (1965) *A System of Computer Programs*, British Psychological Society Booklet.

Hamilton, D. (1971) *Teacher and Pupil Responses to Integrated Studies*, invited paper read to symposium on *Integrated Studies in the Secondary School*, Leeds University Institute of Education, October 1971.

Harbeck, M.B. (1970) 'Instructional Objectives in the Affective Domain', *Educational Technology*, 10, pp. 49-52.

Harmin, M., Kirschenbaum, H. and Simon, S.B. (1970) 'Teaching Science with a Focus on Values', *The Science Teacher,* 37 (1), pp. 16-20.

Harvey, O.J. and Beverly, G. (1961) 'Some Personality Correlates of Concept Change Through Role Playing', *Journal of Abnormal and Social Psychology,* 63, pp. 125-130.

Hasan, P. and Butcher, H.J. (1966) 'Creativity and Intelligence: a Partial Replication with Scottish Children of Getzel's and Jackson's Study', *British Journal of Psychology,* 57, pp. 129-135.

Hays, W.L. and Bennett, J.F. (1961) 'Multidimensional Unfolding: Determining Configuration from Complete Rank Order Reference Data', *Psychometrika,* 26, pp. 221-238.

Heath, R.W. (1964) 'Curriculum, Cognition, and Educational Measurement', *Educational and Psychological Measurement,* 24, pp. 239-253.

Heim, A.W. (1970a) *AH4 Group Test of General Intelligence* (Revised Edition), Windsor, Berks: NFER.

Heim, A.W. (1970b) *Intelligence and Personality,* Harmondsworth: Penguin Books.

Heinemann (1970) *Integrated Science Worksheets,* Edinburgh: Heinemann.

Hicks, L.E. (1970) 'Some Properties of Ipsative, Normative, and Forced-Choice Normative Measures', *Psychological Bulletin,* 74, pp. 167-184.

Hilton, T.L. and Berglund, G.W. (1974) 'Sex Differences in Mathematics Achievement - A Longitudinal Study', *Journal of Educational Research,* 67, pp. 231-237.

Hirschlein, B.M. and Jones, J.G. (1971) 'The Function of Stated Objectives in Teaching for Affective Learning', *Educational Technology,* 11, pp. 47-49.

Hirst, P. (1965) 'Liberal Education and the Nature of Knowledge' in *Philosophical Analysis and Education,* (ed.) Archambault, R.D., London: Routledge and Kegan Paul.

Hirst, P. (1967) 'The Curriculum' in *The Educational Implications of Social and Economic Change*, The Schools Council Working Paper No. 12, London: HMSO, pp. 74-84.

Hirst, P.H. and Peters, R.S. (1970) *The Logic of Education*, London: Routledge and Kegan Paul.

Hollander, E.P. (1971) *Principles and Methods of Social Psychology*, London: OUP.

Holroyd, C. (1972) 'Natural Science' in Dixon, K. (ed.) *Philosophy of Education and the Curriculum*, Oxford: Pergamon Press, pp. 29-73.

Hovland, C.I., Janis, I.L. and Kelley, H.H. (1963) *Communication and Persuasion*, Yale University Press.

Hovland, C.I. and Mandell, W. (1952) 'An Experimental Comparison of Conclusion Drawing by the Communicator and by the Audience', *Journal of Abnormal and Social Psychology*, 47, pp. 581-588.

Hudson, L. (1966) *Contrary Imaginations*, London: Methuen.

Hughes, E.F. (1971) 'Role Playing as a Technique for Developing a Scientific Attitude in Elementary Teacher Trainees', *Journal of Research in Science Teaching*, 8, pp. 113-122.

Hunt, E.A. (1965) 'Discrimination against Girls', *Symposium on Science Shortfall*, Sir John Cass College, London.

Hutt, C. (1972) *Males and Females*, London: Penguin.

Insko, C.A. and Schopler, J. (1967) 'Triadic Consistency: A Statement of Affective-Cognitive-Conative Consistency', *Psychological Review*, 74, pp. 361-376.

Jackson, P.W. (1968) *Life in Classrooms*, New York: Holt, Rinehart and Winston.

Jacobson, W.J. (1969-70) 'Integrated Science Programs for Schools and Colleges', *New Trends in Integrated Science, Volume 1,* Paris: UNESCO.

James, A.N. and Pafford, W.N. (1973) 'The Relationship between Academic Achievement in Science and Father's Occupation', *Science Education,* 57, pp. 37-41.

James, C. (1968) *Young Lives at Stake: a Reappraisal of Secondary Schools,* London: Collins.

Janis, I.L. and Gilmore, J.B. (1965) 'The Influences of Incentive Conditions on the Success of Role-Playing in Modifying Attitudes', *Journal of Personality and Social Psychology,* 1, pp. 17-27.

Janis, I.L. and King, B.T. (1954) 'The Influence of Role-Playing on Opinion Change', *Journal of Abnormal and Social Psychology,* 49, pp. 211-218.

Jevons, F.R. (1969) *The Teaching of Science,* London: Allen and Unwin Ltd..

Katz, D. (1960) 'The Functional Approach to the Study of Attitudes', *Public Opinion Quarterly,* 24, pp. 163-204.

Keeves, J. (1973) 'Differences between the Sexes in Mathematics and Science Courses', *International Review of Education,* 19, pp. 47-63.

Kelman, H.C. (1953) 'Attitude Change as a Function of Response Restriction', *Human Relations,* 6, pp. 185-214.

Kelman, H.C. (1958) 'Compliance, Identification and Internalization, Three Processes of Attitude Change', *Journal of Conflict Resolution,* 2, pp. 51-60.

Kelman, H.C. (1961)'Three Processes of Social Influence', *Public Opinion Quarterly,* Princeton, U.S.A., 25, pp. 57-78.

Kelly, A. (1974) *An Unfair Profession: A Review of the Position of British Women in Science,* Centre for Educational Sociology, University of Edinburgh.

Kelly, A.V. (1974) *Teaching Mixed Ability Classes: an Individualized Approach,* London: Harper and Row.

Kelly, E.L. (1967) 'Transfer of Training: An Analytic Study' in *Psychological Concepts in Education,* Komisar, B.P. and MacMillan, C.J.B. (eds), Chicago: Rand McNally and Company.

Kempa, R.F. and Dubé, G.E. (1973) 'Cognitive Preference Orientations in Students of Chemistry', *British Journal of Educational Psychology,* 43, pp. 279-288.

Kerlinger, F.N. (1969) *Foundations of Behavioural Research,* London: Holt, Rinehart and Winston.

Kerlinger and Pedhazur, (1973) *Multiple Regression in Behavioural Research,* New York: Holt, Rinehart and Winston.

Kerr, J.F. (1966) 'Science Teaching and Social Change', *School Science Review,* 47, No. 162.

Khan, S.B. and Weiss, J. (1973) 'The Teaching of Affective Responses' in Travers, R. (ed.) *Second Handbook of Research on Teaching,* pp. 759-804, Rand McNally.

King, B.T. and Janis, I.L. (1956) 'Comparison of the Affectiveness of Improvised Versus Non-improvised Role-Playing Producing Opinion Change', *Human Relations,* 9, pp. 177-186.

Kirk, G. (1973) 'A Critique of Some Arguments in the Case for Integrated Studies', *Scottish Educational Studies,* 5 (2), pp. 95-102.

Klopfer, L.E. (1971) 'Evaluation of Learning in Science' in *Handbook on Formative and Summative Evaluation of Student Learning,* pp. 559-641, Bloom, B.S., Hastings, J.T., Madans, G.F., London: McGraw-Hill Book Company.

Knox, R.E. and Inkster, J.A. (1968) 'Post Decision Dissonance at Post Time', *Journal of Personality and Social Psychology,* 8, pp. 319-332.

Koestler, A. (1964) *The Act of Creation*, London: Hutchinson.

Krathwohl, D.R. (1965) 'Stating Objectives Appropriately for Programme, for Curriculum and for Instructional Materials Development', *Journal of Teacher Education*, 16 (2), pp. 83-92.

Krathwohl, D.R., Bloom, B.S. and Masia, B.B. (1964) *Taxonomy of Education Objectives Handbook II: Affective Domain*, London: Longman.

Kuder, G.F. (1953) *Examiner Manual for the Kuder Preference Record, Vocational Form C* (5th ed.), Chicago: Science Research Associates.

Kuhn, D.J. (1973) 'Value Systems in Life Science Instruction', *Science Education*, 57 (3), pp. 343-352.

La Piere, R.T. (1934) 'Attitudes VS Actions', *Social Forces*, 13, pp. 230-237.

Lawton, D. (1973) *Social Change, Educational Theory and Curriculum Planning*, London: U.L.P..

Lewis, D.G. (1968) *Experimental Design in Education*, London: University of London Press.

Lewy, A. (1968) 'The Empirical Validity of Major Properties of a Taxonomy of Affective Educational Objectives', *The Journal of Experimental Education*, 36 (3), pp. 70-77.

Likert, R. (1932) 'A Technique for the Measurement of Attitudes', *Archives of Psychology*, No. 140.

Lindquist, E.F. (1956) *Design and Analysis of Experiments in Psychology and Education*, Houghton Mifflin.

Lord, F.M. (1956) 'The Measurement of Growth', *Educational and Psychological Measurement*, 16, pp. 421-437.

Lord, F.M. (1958) 'Further Problems in the Measurement of Growth', *Educational and Psychological Measurement*, 18, pp. 437-454.

Lord, F.M. (1963) 'Elementary Models for Measuring Change' in Harris, C.W. (ed.) *Problems in Measuring Change*, Madison: University of Wisconsin Press.

Lowery, L.F. (1967) 'An Experimental Investigation into the Attitudes of Fifth Grade Students Toward Science', *School Science and Mathematics*, 67, pp. 569-579.

McCartin, R. (1970) 'The Cognitive and Affective Learning of Children' in Bacon, E. (ed.) *40th Yearbook National Council for the Social Studies*, Washington: NEA.

McElwain, D.W. and Keats, J.A. (1961) 'Multidimensional Unfolding: Some Geometrical Solutions', *Psychometrika*, 26, pp. 325-332.

McIntyre, D. and Morrison, A. (1967) 'The Educational Opinions of Teachers in Training', *British Journal of Social and Clinical Psychology*, 6, pp. 32-37.

Mackay, L.D. (1972) 'Changes in Cognitive Preferences During Two Years of Physics Study in Victorian Schools', *Australian Science Teachers' Journal*, 18, pp. 63-66.

McKeachie, W.J. (1963) 'Research on Teaching at the College and University Level' in Gage, N.L. (ed.) *Handbook of Research on Teaching*, pp. 1118-1172.

McLean, M.S., Gowan, M.S. and Gowan, J.C. (1955) 'A Teacher Selection and Counselling Service', *Journal of Educational Research*, 48, pp. 669-677.

McNemar, Q. (1958) 'On Growth Measurement', *Educational and Psychological Measurement*, 18, pp. 437-454.

Mager, R.F. (1968) *Developing Attitude Toward Learning*, Fearon Publishers, Palo Alto, California.

Mallinson, G.G. and Van Dragt, H. (1952) 'Stability of High School Students' Interest in Science and in Mathematics', *School Review*, 60, pp. 362-367.

Mann, L. and Abeles, R.P. (1970) 'Evaluation of Presidential Candidates as a Function of Time and Stage of Voting Decision', *The Journal of Psychology*, 74, pp. 167-173.

Marks, R.L. (1967) 'CBA High School Chemistry and Concept Formation', *Journal of Chemical Education*, 44, 8, pp. 471-474.

Medawar, P.B. (1967) *The Art of the Soluble*, London: Methuen.

Mee, A.J., Ritchie, D. and Boyd, P. (1971) *Science for the Seventies*, London: Heinemann.

Meister, M. (1966) 'Strengthening Science Education for Youth and Industry: Factors in the Early Motivation of Scientists', *Science Education*, 50, pp. 409-410.

Messick, S. (1970) 'The Criterion Problem in the Evaluation of Instruction: Assessing Possible, not just Intended Outcomes' in Wittrock, M.C. and Wiley, D.E. (eds) *The Evaluation of Instruction -Issues and Problems*, pp. 183-202, Chicago: Holt, Rinehart and Winston.

Meyer, G.R. (1961a) 'Factors Accompanying the Scientific Interests of a Selected Group of English Secondary School Pupils', *Australian Journal of Education*, 5, pp. 27-40.

Meyer, G.R. (1961b) 'Factors Accompanying the Scientific Interests of a Selected Group of English Secondary Pupils Part II', *Australian Journal of Education*, 5, pp. 105-115.

Meyer, G.R. (1963) 'Factors Related to Scientific Attitudes within the Secondary Schools of an Australian City', *Australian Journal of Education*, 7, pp. 21-40.

Miles, D.T. (1972) 'Affective Priorities and Educational Objectives', *Educational Technology*, 12, pp. 33-35.

Moore, R.W. and Sutman, F.X. (1970) 'The Development, Field Test and Validation of an Inventory of Scientific Attitudes', *Journal of Research in Science Teaching*, 7, pp. 85-94.

Morshead, R.W. (1965) 'Taxonomy of Educational Objectives Handbook 2: Affective Domain', *Studies in Philosophy and Education*, 4 (1), pp. 164-170.

Musgrove, F. (1973) 'Power and the Integrated Curriculum', *Journal of Curriculum Studies*, 5 (1), pp. 3-12.

Neale, D.C. and Proshek, J.M. (1967) 'School-related Attitudes of Culturally Disadvantaged Elementary School Children', *Journal of Educational Psychology*, 58, pp. 238-244.

Neidt, C.O. and Hedlund, D.E. (1969) 'Longitudinal Relationships between Cognitive Learning and Affective Learning Outcomes', *Journal of Experimental Education*, 37, pp. 56-60.

Newsom Report (1963) Central Advisory Council for Education (England), *Half our Future*, London: HMSO.

Nisbet, J.D. and Entwistle, N.J. (1970) *Educational Research Methods*, London: University of London Press.

Nuttall, D. (1971a) *Modes of Thinking and Their Measurement*, Ph.D. Thesis, University of Cambridge.

Nuttall, D. (1971b) *Administrator's Manual for Science Attitude Questionnaire*, Slough:NFER.

O'Conner, E.F. (1972) 'Extending Classical Test Theory to the Measurement of Change', *Review of Educational Research*, 42, pp. 73-97.

O'Conner, D.J. (1957) *Philosophy of Education*, pp. 104-110, London: Routledge and Kegan Paul.

Oliver, R.A.C. and Butcher, H.J. (1962) 'Teachers' Attitudes to Education: the Structure of Educational Attitudes', *British Journal of Social and Clinical Psychology*, 1, pp. 56-69.

Oliver, R.A.C. and Butcher, H.J. (1968) 'Teachers' Attitudes to Education', *Durham Research Review*, 22, pp. 358-364.

Omerod, M.B. (1971) 'The Social Implications' Factor in Attitudes to Science', *British Journal of Educational Psychology*, 41, pp. 335-338.

Oppenheim, A.N. (1966) *Questionnaire Design and Attitude Measurement*, London: Heinemann.

Osgood, C.E., Suci, G.J. and Tannenbaum, P.H. (1957) *The Measurement of Meaning*, Urbana, Illinois: University of Illinois Press.

Peak, H. (1959) *The Affects of Aroused Motivation on Attitudes*, Technical Report to the Office of Naval Research. (Mimeo)

Peaker, G.F. (1967) 'The Regression Analyses of the National Survey' in *Children and Their Primary Schools Volume 2*, London: HMSO.

Peck, R.F. and Havinghurst, R.J. (1960) *The Psychology of Character Development*, New York: Wiley.

Perrodin, A.F. (1966) 'Children's Attitudes toward Elementary School Science', *Science Education*, 50, pp. 214-218.

Peters, R.S. (1962) 'Theory Policy and Technology' in *Brett's History of Psychology*, Peters, R.S. (ed.), pp. 25-36, London: Allen and Unwin.

Phenix, P.H. (1964) *Realms of Meaning*, New York: McGraw-Hill.

Popham, W.J. and Baker, E.L. (1970) *Systematic Instruction*, Englewood Cliffs, New Jersey: Prentice-Hall Inc..

Popper, K.R. (1963) *Conjectures and Refutations*, London: Routledge and Kegan Paul.

Popper, K. (1972) 'A Realist View of Logic, Physics and History' in *Objective Knowledge*, pp. 285-318, Oxford: Clarenden Press.

Pring, R. (1971a) 'Curriculum Integration', *Proceedings of the Philosophy of Education Society of Great Britain*, V (2), pp. 170-200.

Pring, R. (1971b) 'Curriculum Integration' in *The Curriculum: Context, Design and Development*, Edinburgh: Oliver and Boyd.

Raths, J. (1962) 'Clarifying Childrens' Values', *The National Elementary Principal*, 42 (2), pp. 35-39.

Raths, L.E., Harmin, M. and Simon, S.B. (1966) *Values and Teaching: Working with Values in the Classroom*, Columbus, Ohio: Charles E. Merrill.

Raven, J., Ritchie, J. and Baxter, D. (1971) 'Factor Analysis and Cluster Analysis: Their Value and Stability in Social Survey Research', *Economic and Social Review*, 2 (3), pp. 367-392.

Reed, H.B. (1961) 'Teacher Variables of Warmth, Demand and Utilization of Intrinsic Motivation Related to Pupils' Science Interests: A Study Illustrating Several Potentials of Variance-Covariance', *Journal of Experimental Education*, 29, pp. 205-229.

Reed, H.B. (1962) 'Implications for Science Education of a Teacher Competence Research', *Science Education*, 46, pp. 473-486.

Registrar General (1970) *Classification of Occupations*, London: HMSO.

Robinson, J.T. (1969) 'Philosophical and Historical Bases of Science Teaching', *Review of Educational Research*, 39, pp. 459-471.

Roe, A. (1963) 'Psychological Approaches to Creativity in Science' in Coler, M.A. (ed.) *Essays on Creativity in the Sciences*, New York University Press.

Rogers, E.M. (1972) *The Dominance of Cognitive Processes in Science Teaching: Possible Measures to Overcome this Bias'*, Paper presented to IPN Symposium on *Cognitive Processes* and Science Education, Kiel.

Rosenberg, M.J. (1956) 'Cognitive Structure and Attitudinal Affect', *Journal of Abnormal and Social Psychology*, 53, pp. 367-372.

Rosenberg, M.J. (1960) 'A Structural Theory of Attitude Dynamics', *Public Opinion Quarterly*, 24, pp. 319-341.

Rosenthal, R. and Jacobson, L. (1966) 'Teacher Expectancies: Determinants of Pupils' I.Q. Gains', *Psychological Reports*, 19, pp. 115-118.

Rothman, A.I. (1967) 'The Factor Analysis of a Science Related Semantic Differential Instrument', *Journal of Educational Measurement*, 5, pp. 145-149.

Rothman, A.I. (1969) 'Teacher Characteristics and Student Learning', *Journal of Research in Science Teaching*, 6, pp. 340-348.

Rothman, A.I., Welch, W.W. and Walberg, H.J. (1969) 'Physics Teacher Characteristics and Student Learning', *Journal of Research in Science Teaching*, 6, pp. 340-348.

Rutherford, J. and Gardner, M. (1969-70) 'Integrated Science Teaching' in *New Trends in Integrated Science Teaching*, Paris: UNESCO.

Schmedemann, G. and La Sheir, W.S. (1967) 'Cognitive Preferences of Students and Selected Characteristics of their PSSC Teachers', *Journal of Research in Science Teaching*, 5, pp. 40-42.

Schools Council Integrated Science Project (1971) *Bulletin 3*, 90 Lillie Road, London S.W.6.

Schools' Council Working Paper No. 2 (1965) *Raising the School Leaving Age*, pp. 52-53, London: HMSO.

Schools' Council Working Paper No. 11 (1967) *Society and the Young School Leaver*, London: HMSO.

Schulman, L.S. and Tamir, P. (1973) 'Research on Teaching in the Natural Sciences' in Travers, R.M.W. (ed.) *Second Handbook of Research on Teaching*, Chicago: Rand McNally.

Schwab, J.J. (1960) 'What Do Scientists Do?', *Behavioral Science*, 5, pp. 1-27.

Schwab, J.J. (1962) 'The Teaching of Science as Enquiry' in *The Teaching of Science*, (ed.) Schwab, J.J. and Brandwein, P.F., Cambridge Mass: Harvard University Press, pp. 3-103.

Schwab, J.J. (1964a) 'Structure of the Disciplines: Meanings and Significances' in *The Structure of Knowledge and the Curriculum*, (ed.) Ford, G.W. and Puguo, L., pp. 31-49, Chicago: Rand McNally.

Schwab, J.J. (1964b) 'The Structure of the Natural Sciences' in *The Structure of Knowledge and the Curriculum*, (ed.) Ford, G.W. and Puguo, L., Chicago: Rand McNally.

Schwirian, P.M. (1968) 'On Measuring Attitudes toward Science', *Science Education*, 52, pp. 172-179.

Schwirian, P.M. and Thomson, B. (1972) 'Changing Attitudes toward Science: Undergraduates in 1967 and 1971', *Journal of Research in Science Teaching*, 9, pp. 253-259.

Science 5/13 (1971) *Report for Teachers on the Evaluation of the First Set of Units*, Schools' Council Publications, University of Bristol.

Scott, W.A. (1968) 'Attitude Measurement' in Lindzey, G. and Aronson, E., *The Handbook of Social Psychology*, Volume 2, pp. 207-273, London: Addison-Wesley Publishing Company.

Scottish Certificate of Education Examinations Board (1968) *Biology: Ordinary and Higher Grades Syllabuses and Notes*, Edinburgh.

Scottish Certificate of Education Examinations Board (1969a) *Chemistry: Ordinary and Higher Grades Syllabuses and Notes*, Edinburgh.

Scottish Certificate of Education Examinations Board (1969b) *Physics: Ordinary and Higher Grades Syllabuses and Notes*, Edinburgh.

Scottish Education Department (1969) *Curriculum Paper 7: Science for General Education*, Consultative Committee on the Curriculum, HMSO Edinburgh.

Scriven, M. (1966) 'Student Values as Educational Objectives', *Proceedings of the 1965 Invitational Conference on Testing Problems*, Princeton, New Jersey: Educational Testing Service.

Shaw, M.E. and Wright, J.M. (1967) *Scales for the Measurement of Attitudes*, London: McGraw-Hill.

Sherwood, J.J., Barron, J.W. and Fitch, H.G. (1969) 'Cognitive Dissonance: Theory and Research' in Wagner, R.V. and Sherwood, J.J. (eds) *The Study of Attitude Change*, Brooks/Cole Publishing Company, Belmont, California.

Siegal, S. (1956) *Non-Parametric Statistics for the Behavioral Sciences*, New York: McGraw-Hill.

Skurnick, L.S. and Jeffs, P.M. (1971) *Science Attitude Questionnaire*, National Foundation for Educational Research, Slough, Bucks.

Smith, B.O. (1966) 'Teaching and Testing Values' in *Proceedings of the 1965 Invitational Conference on Testing Problems*, Princeton, New Jersey: Educational Testing Service, 1966, pp. 50-59.

Stafford, R.E. (1961) 'Sex Differences in Spatial Visualization as Evidence of Sex-limited Inheritance', *Perceptual and Motor Skills*, 13, p. 428.

Stenhouse, L. (1968) 'The Humanities Curriculum Project', *Journal of Curriculum Studies*, 1 (1), pp. 26-30.

Stephens, J.M. (1963) 'Transfer of Learning' in *Transfer of Learning*, pp. 103-121, edited by Grose, R.F. and Birney, R.C., New York: D. Van Nostrand Inc.

Stoops, J.A. (1953) 'Stability of the Measured Interests of High School Pupils Between Grades Nine and Eleven', *Educational Outlook*, 27, pp. 116-118.

Strong, E.K. (1943) *Vocational Interests of Men and Women*, Stanford: University Press.

Suchman, J.R. (1962) *The Elementary School Training Programme in Scientific Enquiry*, Urbana, Illinois: University of Illinois.

Sumner, R.J. and Wilson, N.L. (1972) 'A Survey of Science Interests of Secondary School Students', *Australian Science Teachers' Journal*, 54, pp. 64-67.

Taylor, A.L. (1965) *Teacher Attitudes, Pupil Behaviour and Content Attributes in Relation to the Use of Programmed Science Materials at Fourth Grade Level*, Unpublished Ed.D. dissertation, University of Virginia.

Taylor, P.H., Christie, T. and Platts, C.V. (1970) 'An Exploratory Study of Science Teachers' Perceptions of Effective Teaching', *Educational Review*, 23, pp. 19-32.

Terman, L.M. (1954) 'Scientists and Non-Scientists in a Group of 800 Gifted Men', *Psychological Monograph 68*, Whole Number 378, pp. 1-44.

Thurstone, L.L. and Chave, E.J. (1929) *The Measurement of Attitudes*, Chicago: University of Chicago Press.

Tisher, R.P. (1968) 'Evaluation of New Curricula', *Australian Science Teachers' Journal*, 14, pp. 54-56.

Tisher, R.P. and Power, C.N. (1973) *The Effects of Teaching in Mini-teaching and Micro-teaching Situations where Australian Science Education Project Materials are used,* Faculty of Education, University of Queensland.

Torrance, E.P. (1962) *Guiding Creative Talent,* Englewood Cliffs, New Jersey: Prentice-Hall.

Tucker, L.R., Damarin, F. and Messick, S. (1966) 'A Base-free Measure of Change', *Psychometrika,* 31, pp. 457-473.

Tyler, L.E. (1964) 'The Antecedents of Two Varieties of Vocational Interest', *Genetic Psychological Monographs,* 70, pp. 177-227.

University of Bristol School of Education, (1970) *Science 5/13: Some Results from the Computer Analysis of the Evaluation Data,* Wynne Harlen.

Vernon, P.E. (1953) *Personality Tests and Assessments,* London: Methuen.

Wagner, R.V. (1969) 'The Study of Attitude Change: An Introduction' in Wagner, R.V. and Sherwood, J.J., *The Study of Attitude Change,* Brooks/Cole Publishing Company, Belmont, California.

Walberg, H.J. (1967) 'Dimensions of Scientific Interests in Boys and Girls Studying Physics', *Science Education,* 51, pp. 111-116.

Walberg, H.J. (1969) 'Predicting Class Learning: An Approach to the Class as a Social System', *American Educational Research Journal,* 6, pp. 529-542.

Welch, W.W. (1969) 'Correlates of Course Satisfaction', *Journal of Research in Science Teaching,* 6, pp. 54-58.

Welch, W.W. and Walberg, H.J. (1967) 'Are the Attitudes of Teachers Related to Declining Percentage Enrollments in Physics?', *Science Education,* 51, pp. 436-442.

White, J.P. (1971) 'The Concepts of Curriculum Evaluation', *The Journal of Curriculum Studies*, 3, pp. 101-112.

Whitfield, R.C. (1971) *Disciplines of the Curriculum*, London: McGraw-Hill.

Wick, J.W. and Yager, R.E. (1966) 'Some Aspects of the Student's Attitude in Science Courses', *School Science and Mathematics*, 66, pp. 269-273.

Wicker, A.W. (1969) 'Attitudes Versus Actions: The Relationship of Verbal and Overt Behavioural Responses to Attitude Objects', *Journal of Social Issues*, 29, pp. 41-78.

Wilson, P.S. (1971) *Interest and Discipline in Education*, London: Routledge and Kegan Paul.

Winder, P.A. (1966) 'Interest Test', *Science Education*, 50, pp. 371-373.

Winer, B.J. (1962) *Statistical Principles in Experimental Design*, New York: McGraw-Hill.

Wish, M. (1964) 'An Unfolding Analysis of Cognitive Preferences', *Educational Testing Service*, Princeton, New Jersey.

Wynn, D.C. and Bledsoe, J.C. (1967) 'Factors Related to Gain and Loss of Scientific Interest During High School', *Science Education*, 51, pp. 67-74.