MODES OF INFORMATION PERCEPTION AND

TRANSFORMATION IN SCIENCE STUDENTS

A thesis submitted for the Degree of Doctor of Philosophy of the University of Keele

August 1980 **J.N. King**

Original copy tightly bound.

Some text is bound into the spine.

This work is dedicated to my family in gratitude for endless forbearance.

"Science is knowledge. Knowledge is not wisdom. Wisdom is knowledge tempered by judgement."

Lord Ritchie-Calder

ABSTRACT

A model is proposed in which the learning of a physical science is related to the reception and perception of information through behaviour expressed in terms of cognitive preference and to the transformation of information which is examined in terms of judgement ability. The nature of cognitive preference and of judgement is defined, and the validity and other statistical credentials of the instruments used to measure these concepts is reported. The nature, stability and structure which is found to underlie cognitive preference supports the contention that cognitive preference may be classified correctly as a cognitive style. The relationships between cognitive preference and the usual parameters of psychological measurement are examined and the whole body of data is related to a tentative model of learning behaviour in order to determine whether deeper insight is available than is possible without the model.

ACKNOWLEDGEMENTS

It is with much gratitude that I acknowledge my debt to Professor R.F. Kempa of the Department of Education at the University of Keele, Keele, Staffordshire, U.K. He initiated the work and sustained it with much inspiration. He was invariably patient and encouraging.

I am indebted, too, to the students, Head Teachers and Departmental Teachers of the following ten secondary schools:-

Colliers School, Horsham, Sussex.

Hurstpierpoint College, Haywards Heath, Sussex.

The Judd School, Tonbridge, Kent.

Kent College for Girls, Pembury, Kent.

Maidstone Grammar School for Girls, Maidstone, Kent.

The Skinners' School, Tunbridge Wells, Kent.

Tonbridge Grammar School for Girls, Tonbridge, Kent.

Tonbridge School, Tonbridge, Kent.

Walthamstow Hall, Sevenoaks, Kent.

Weald of Kent School, Tonbridge, Kent.

They gave generously of their time and energy.

I am profoundly grateful to the Directors of the Reed International Paper Co. Ltd. and to Mr. J.A. Lewis and particularly to Mr. B.E. Blythe, both of that company. I was permitted to make unlimited use of the computing facilities at Aylesford, Nr. Maidstone, and the care and tolerance of all the staff made this research possible. I was given much valuable help by my colleagues in both the Chemistry Department and Mathematics Department at Tonbridge School. Mr. B.W. King's help with some of the computing was much appreciated. Finally, I must thank Mrs. A.D. Braithwaite for typing the manuscript; her care created a fine manuscript from a difficult draft.

INDEX

Chapter 6 : The Structure of Cognitive Preference and the Evaluation of Pupils

Chapter 7 : The Stability of Cognitive Preference Behaviour

Chapter 8 : The Judgement Exercises

Chapter 9 : Information Perception and Information Transformation

Chapter 10 : Cognitive Preferences, Judgement Behaviour and Classroom Learning Behaviour

CHAPTER 1 : LEARNING ACTIVITIES

Section 1.0 : Introduction

Learning concerns the incorporation of information into existing cognitive structure; *cognition* refers to the processes by which knowledge is acquired. It has been customary in psychological research to relate and evaluate learning against constructs like ability (i.e. levels of skill), personality and valuing. Cognition is intimately concerned with innate capacity and information processing habits but its study has been essentially theoretical and far removed from classroom learning.

A fruitful and important area for research was seen to lie in the relationships between practical classroom learning and the manner and form of the cognition. The principal purpose was to study both learning and cognition in adolescent minds while they were actively engaged in learning in an academic discipline. Throughout this study the term 'learning' is considered to be the acquisition of knowledge. A three-stage sequence of classroom learning was hypothesised. In this sequence a stimulus statement was thought to initiate a *reception* response. Reception was thought to be followed by a *perception* response and then by *transformation*. It was further postulated that the three stages occur before incorporation into existing cognitive structure is possible.

This research is a logical development of the work which has already been done on *cognitive styles*. These have been variously defined. Messick (1970) defined them as "information processing habits"; Scott (1973) preferred "the patterns that an individual may use in coping with environmental stimuli" and Field (1972) simply described them as "the pattern of thinking". In each case, the pattern is assumed to be relatively stable and consistent within an individual. One common feature of the work in this field has been the absence of situations which might be embraced by the syllabus of an academic discipline; i.e. of material which might be described as classroom learning material. The logical development, then, was to seek to trace the relationships, if any, between established cognitive style theory and the learning of material with curricular content.

Witkin et al. (1971) described cognitive styles in terms of two types. The first type of style was concerned with "the extent of the development of particular capacities, and when used in this way it is a *capacity* variable." The second type embraced the notion of *choice* and they were related to "the manner in which a person chooses to use his capacity." Witkin related the stylistic behaviour of individuals to scales with poles of extreme behaviour at the extremes. Both these two types of style and the differentiation of the individuals in a population are of great significance in the work which follows. This significance stems from the claim that the manner and form of individual response to the factual material of an academic discipline may have many of the characteristic attributes of cognitive styles which lack curricular content. Two cognitive styles of learning were postulated. The first of these was concerned with the first two stages of the learning sequence (information reception and information perception) and the third to the final stage of the sequence (information transformation).

In 1964 Heath described work which was concerned with choice behaviour in response to the learning of factual material in elementary physics. Heath differentiated behaviour in terms of four scales and he gave the title *cognitive preference behaviour* to the patterns which he observed. It was one of two principal aims of this research to relate the first two stages of the learning sequence, in so far as they can be measured in terms of cognitive preference behaviour, to a choice style of cognitive learning. The second aim was to examine the transformation stage of the learning sequence, which was measured in an exercise in judgement, in terms of a capacity style of cognitive learning.

Now the analysis of psychological development views behaviour in terms of component pieces and then sums the pieces with the expectation that the sum will be more significant than the isolated pieces. The conjectural learning sequence was proposed in order to divide classroom learning behaviour into some of its possible components. Of necessity, other components of proven importance like short and long term memory, motivation and intellectual abilities were largely omitted. The first of the three stages of the sequence, information *reception*, was thought of as a passive activity, i.e. as the acquisition of information without it being altered intellectually by the individual. The rote learning of poetry may be thought of as an exemplar of this activity. Information *perception* was conceptualised as an intuitive recognisory action; it was thought of as the internal response by the active cognitive structure of the individual to the stimuli from elements of new information. The final stage, information transformation, was seen as a precondition of information utilisation particularly in problem solving and, in the context of this research, of judgement.

The implications of cognitive style theory which is now well established, are closely related to the sequence; in particular, it is style of information perception and style of information utilisation that are of relevance here. The former is considered further in Section 1.1 and the latter in Section 1.2.

Section 1.1 : Information Reception, Information Perception and Cognitive Preference

The first two stages of the proposed sequence are information reception and information perception. The reception of information may be considered to be a somewhat "mechanical" process which was not affected by the prevailing cognitive structure of the learner. At this stage the learner does not impose ordering or the selection of material. The perception of the learner to new material maybe postulated as being subject to the material which is already in the learner's cognitive structure. Thus the distinction between the first two stages may be thought to reside in the rôle played by existing cognitive structure.

It is reasonable to suggest that the perception of new information inevitably entails its comparison with information or knowledge which is already firmly embedded in the student's cognitive structure. If the new material has characteristics which conflict with the existing cognitive structure, it is likely that intellectual tensions will appear, with the result that the new material may be difficult to assimilate. If, however, the new material harmonises with the learner's existing cognitive structure, no major intellectual tensions will occur and the new material will be assimilated readily. A further suggestion that may be advanced is that the learner, when confronted with different stimuli, will choose from the set the one or ones which least conflict with his existing cognitive structure. In doing so he would keep the intellectual conflict or tension to a minimum. Likewise in a process requiring different stimuli to be ordered according to the preference felt by the student for them, it may be held that the order at which the student ultimately arrives reflects the order of increased cognitive conflict so that the least preferred stimulus would be the one which produces within him the highest intellectual conflict.

It may be supposed that in expressing a preference in any field the person concerned is revealing an aspect of his cognitive structure. It may thus be argued that cognitive preference orderings or measurements may legitimately be interpreted in terms of a respondent's cognitive structure.

The test which was devised by Heath (1964) for the measurement of cognitive preference consisted of items each of which contained a certain amount of factual information, followed by four options reflecting four different response modes. These four modes were construed by Heath as being representative of four of the ways of approaching factual information and they are described in detail in Chapter 2. It may suffice here merely to label the four modes of response as recall of information, application, questioning and principles and leave the description of them until later. Heath's instrument consists of items in multiple-choice format; in each item there are four options, one for each of the four modes of response. Respondents are asked to indicate which one of the four modes of response they found "most appealing".

The intellectual process of selecting the one mode in each item is believed to be an exercise in comparing modes with one another until the one which causes the least conflict with the intellectual material which is already in the respondent's cognitive structure has been chosen. This is the "most appealing" option for that item and in selecting it the respondent has expressed a preference for one of the four modes of response and rejected the other three modes. The same activity is performed on all the items of the instrument and the sum of the preferences cast in favour of each mode reveals the respondent's orientation towards that mode of response; the relative scores of the four modes are called his 'cognitive preference orientation'.

Heath, in his review of his findings concluded that

> "the generally positive results obtained suggest that this type of instrument can identify, in a meaningful context, curriculum-related differences in cognitive *style*."
>
> Heath (1964)

This statement by Heath implies a relationship to exist between cognitive preferences and cognitive styles, albeit in a context related to a curriculum. Although a number of workers followed Heath in studying cognitive preference, the notion that cognitive preferences could be related to cognitive styles was not taken up until the research by Kempa and Dube

(1973). They suggested overtly that cognitive preferences might be seen as an aspect of cognitive style. In this argument they relied on Messick's definition of cognitive style as "information processing habits " (1970). They saw cognitive preference tests as measuring certain traits but they left open the question whether these were acquired through curricular influences or whether they were latent in the students themselves. One particular finding in support of the notion of styles, reported by Kempa and Dubé, was the existence of two dichotomous relationships between the four modes which had been suggested by Heath. These relationships are discussed later (Chapter 7).

Brown (1975) discussed anew the relationships between cognitive preference measurements and cognitive styles. In doing so she used a wider definition of cognitive style compared with the definition given by Messick, of the type given by Witkin et al (1971) and by Kagan, Moss and Siegel (1963). The latter for example suggested that cognitive styles represent ...

> "Stable individual preferences in modes of perceptual organisation and conceptual organisation of the external environment."

She questioned the validity of Kempa and Dubé's argument on the grounds that a cognitive style is a complex system of interrelated behaviours which cannot be described adequately by measurements based on the constructs for the modes which were used by Heath. (She also questioned the validity of the constructs). Tamir (1975), responding to Brown's critique, argued in favour of cognitive styles, but he suggested the term "cognitive preference style" as an alternative. This he conceptualised as having three elements, viz: (a) a general characteristic of the individual student which may consist of two components, one inherited and the other acquired by experience; (b) a discipline or subject dependent element; and (c) a subject matter specific element. This effectively constitutes an elaboration and refinement of the view which was held by Heath and by Kempa and Dubé.

The reception of information cannot be measured in isolation from the perception of information because the two cannot be distinguished in operational terms for independent measurements but this first stage is still considered to be an important part of the learning sequence. In all the work which follows it is proposed that an assessment of a student's cognitive preference embraces both the notion of information reception and of perception.

Section 1.2 : Information Transformation and Judgement

It will be recalled that the third stage in the suggested leaving sequence was *information transformation*. The concept of information transformation is seen in the following way. It starts with the premise that a learner has available to him a body of information in his cognitive structure. Any new piece of information which is presented to him for learning or testing has to be related to some part or parts of this existing cognitive structure. If the new piece of information is already embodied there, no transformation is required. Alternatively it may be that the information is recognised as an exemplar of a generalisation which is already established in the learner's mind; in this case, the new information is assimilated by the learner without his cognitive structure being affected. This may be said to be a situation in which rote learning ensues. The situation is different if the learner's cognitive structure does not embrace the conceptual framework to which the piece of information can be readily related. In this case a meaningful response to the new information must involve some modification of the learner's cognitive structure: the activity which it invokes is seen to be an activity of *judgement*.

The act of judgement requires first that the learner shall fail to find a reference in cognitive structure to the item of information which he encounters. He must then conceive as many hypotheses as possible which will serve to link the new information with as much of the existing framework as is available. It will then be necessary for the learner to scan his cognitive structure in order to establish which hypothesis is most tenable and which items of previously accepted information could be unified with the new item by the same hypothesis. If all the hypotheses are rejected, the information is not transformed or it is either lost completely, or perhaps, retained in some less meaningful manner. If one (or more) of the hypotheses makes no contradiction with any of the facts or principles which are already in the cognitive structure of the learner, the new information is transformed and bound into the cognitive structure in association with the relevant material that is already present. In practice it is inevitable that in the scrutiny of most hypotheses there will be a greater or lesser degree of contradiction and the hypothesis which is accepted is the one which is found to produce minimum cognitive dissonance. Thus the correctness or accuracy of the accepted hypothesis will often fail to concur with accepted scientific fact. It should also be noted that classroom teaching of science is often concerned with judicious application of facts which are unified by hypotheses of unimpeachable quality, i.e. by laws; the practice of using unfamiliar situations for which no hypotheses have been supplied is well tried in teaching situations (and is referred to as heuristic teaching) but is uncommon in testing.

An example may clarify the transformation. Suppose that a student is told that a soft green solid is commonly found in laboratory solutions of manganese (II) sulphate and that he is asked to explain the cause or origin of this material. He will know that laboratory solutions do not normally contain insoluble materials and therefore he is unable to relate this instance to his previous experience of liquid reagents. To "explain" the phenomenon, the student would have to generate new hypotheses which would include, e.g. the possibility of a reaction between manganese (II) salts with glass, or of contamination through careless use, or of impurity in the water which is used to prepare the reagent, of a reaction between the salt and the air, and of the existence of an air-borne holophytic organism flourishing in this apparently inhospitable habitat. (See footnote). The student will be in one of three situations. He may already know the cause and therefore no element of judgement arises in his selection of the correct hypothesis. His scientific experience may be devoid of all information on plants in inhospitable habitatand so he is unable to make a judgement because he creates no hypothesis with this notion. The third possibility is that he is aware of the colonisation of inhospitable environments by simple plants but he is unaware of this particular instance. Ideally he considers all possible hypotheses and rejects all but the last. In actual life this third situation occurs infrequently but it forms a significant part of much teaching in science. In the exercises which were written to test this skill there is a considerable element of cueing by inclusion of relevant ideas in the suggested hypotheses. Thus the hypotheses included in the item on the plant might have included ideas on the importance of daylight and on the gradual increase in the mass of the material in the reagent bottle. These ideas together with the information that the material is 'soft and green' was thought to lead the respondent to the correct cause. The information about the green solid is thus transformed into a statement about an unexpected plant and it is subsequently absorbed into cognitive framework as an instance of plant behaviour, as a facet of manganese (II) chemistry and as an example of natural and harmless contamination.

In the context of the study which is reported here it was postulated, therefore, that judgement is required only if the learner is faced with' items of information which he has not previously encountered or which he

---

Footnote: these suggestions were all submitted by students in an open-ended test.

cannot accommodate readily into an already established cognitive structure. Further, the item of information is such that for its "assimilation" by the student his cognitive structure has to be modified. The process which is hypothesised for this involves:- (a) hypothesis generation and (b) scanning of the new information with respect to the hypothesis generated, leading (in a positive instance) to some kind of "accommodation" of the new information in a modified cognitive framework. Failure to accommodate the information would lead to its rejection. Thus the act of judgement is seen essentially as an assessment of the "reasonableness" of the various hypotheses vis-a-vis the new information. In the assessment an attempt is made to match the item of information against the various hypotheses on the grounds of the student's *personal* logic. The combination of item and hypothesis would be deemed to be the best because it appeared to be the most logical, it does not transgress any accepted principle or established fact and it minimises cognitive conflict.

Finally the transformation of information is the outcome of the process whereby new information becomes integrated (i.e. accommodated) into cognitive structure which has itself to undergo change in order to allow this integration. Judgement in the sense which has been described above is the, or a, means whereby the transformation of information is achieved. It is evident, therefore, that the key issue is the effective transformation of new information by a learner in his judgement ability, in terms of both "hypothesis generation" and the matching of these against the new information. Hence an investigation of judgement ability seems fully warranted.

Section 1.3 : The Research Problems

The one central issue which this research work explored concerned the learning sequence and its implications in the classroom. It was undertaken in full awareness that the sequence is an incomplete theory of learning. In essence there are two aspects which pervade the work and each aspect has an origin in the sequence. The information perception stage of the sequence led into the cognitive preference aspects and information transformation stage initiated the study of judgement. The decision was taken to examine each of these two aspects from three angles viz:- (a) the psychological nature; (b) the measurement; and (c) the value of each in terms of assessment of pupils. In the closing section of the work the two aspects were drawn together.

The cognitive preference testing was originally introduced by Heath in order to assess curricular influences in behavioural terms. The psychological issues which were examined are concerned with the behavioural traits and the relationship between cognitive preference and cognitive styles. It is the nature of cognitive styles that they should be relatively stable with time and broadly apparent in people's thinking over a wide range of situations. Thus hypotheses were formulated to examine the stability across both time and different disciplines and the differentiation of cognitive preference behaviour. The character of cognitive preference concerned the structure and the significance of the subject matter within the items. In addition, as will be seen, several unresolved aspects concerning the measurement of cognitive preferences exist which, it was felt, required attention. They are essentially technical in nature and they embrace the alternative scoring procedures, the validity, item design and test length. This analytical work is readily justified on the grounds that it would lead to greater confidence in subsequent cognitive preference data. The specific aspects which were examined will be found in Chapter 4.

Judgement has previously been described as an intellectual skill. The psychological examination made in Chapter 3 explores the relationship between judgement, critical thinking and understanding. In broad terms the character of judgement is considered to relate to the assessment of "reasonableness". This must be viewed in the light of both broad scientific principles as well as the issues germane to the particular situation in question. Thus the topics which arose from this and which warranted the study concerned the relationship between the knowledge of those scientific principles and "reasonableness" and between the experience of scientific things in a broader context and "reasonableness". A number of technical questions about the measurement of judgement are also examined in Chapter 8 and consideration is given to the value of the instrument in classroom situations in Chapters 9 and 10.

The detailed structure of the thesis is as follows.

Chapter 2 of this thesis includes a detailed description of cognitive preference and of the evaluation and interpretation of cognitive preference scores. It also includes a review of the relevant literature and a detailed statement of problems for investigation. Chapter 3 contains a detailed account of the nature of judgement and its measurement as well as a review of the literature on critical thinking. This is followed in Chapter 4 by an account of the research problems. The hypotheses which were researched are stated and a general description of the sample populations and ancillary tests is included. Chapter 5, 6 and 7 form Part I of the research work and are exclusively concerned with cognitive preference. Chapter 5 contains the discussion of the technical aspects, Chapter 6 is concerned with structure while those aspects which relate to stability with time and to cognitive style are examined in Chapter 7. Part II contains Chapter 8 where all aspects of the judgement exercise are discussed, Chapter 9 which contains the work which draws cognitive preference and judgement together and the new work is concluded in Chapter 10 by an examination of the relationship between these two with a number of important variables so that the importance of information perception and transformation can be related to work in the classroom. The findings are summarised in Chapter 11.

CHAPTER 2 : THE NATURE OF COGNITIVE PREFERENCE

Section 2.0 : Introduction

The first section of this chapter contains an account of the principles of cognitive preference testing; an example is given and the task which confronts the respondent is explained. Section 2 presents a consideration of the methods of responding and scoring cognitive preference instruments. The problems of evaluation and interpretation of the scores in terms of both validity and statistical treatment are discussed in Section 3. This is followed in Section 4 by a review of earlier studies and, in the final section, more specific statements on the problems for investigation are given.

Section 2.1 : The Principles of Cognitive Preference Testing

The cognitive preference testing in physics first introduced by Heath in 1964, was seen by him as a procedure to supplement traditional tests and measures employed in order to assess the avowed goals of new science curricula. Heath argued that the different approaches and aspirations of modern courses might lead to differences in the modes in which the learner attends to the knowledge portrayed by the curricular materials. Thus, he suggested that our interest should focus upon what a pupil does with information intellectually and not merely on whether he can identify the information as correct or incorrect. Thus his cognitive preference test was designed to permit the students 'to exhibit some preference in cognition'. Heath analysed the objectives of a modern syllabus and selected four modes of attending to the material in the syllabus; a preference expressed by a student for a mode was termed by Heath a 'cognitive preference'. The four modes were labelled 1) *recall* of specific facts or terms, 2) practical *application*, 3) critical *questioning* of information and 4) fundamental *principles*.

Heath's test comprised 20 items in multiple-choice format. Each item contained an introductory statement (or stem) followed by four alternative statements or options from which the respondent was required to select the one statement that 'most appealed to him'. All the options were, of course, factually correct and the choice of a particular option was a matter of personal preference only. Each option related to the stem and to one of the four modes. The cognitive preference behaviour of the respondent was quantified by counting the preferences cast in favour of each mode. One of the items which was used in a recent cognitive preference test is given as an illustration.

A Cognitive Preference Item

Lime Water

(1) Carbon dioxide reacts in solution with calcium ions to form insoluble carbonate of calcium, this is precipitated;

(2) A solution of lime water is turned milky by carbon dioxide; may go clear again if lots of gas is used;

(3) A solution of barium ions would serve just as effectively as a solution of calcium ions;

(4) Provides convenient method for distinguishing carbon dioxide from other gases.

The brief title serves only to locate the syllabus area of the material which is embodied in the item. The first option in this example refers to the chemical *principle* which underlies the simple test for carbon (IV) oxide in the laboratory. The phrasing of the option is slightly unfamiliar in order to minimise mere recall of chemical information. It was considered that both the appearance and the disappearance of the precipitate were suitable obvervations for inclusion as *recall* responses. The rôle played by barium ions as substitutes for the calcium ions was thought to be both sufficiently unfamiliar and a warranted extension of data that might initiate quizzical thought about the property that these ions have in common.

hence this option is the *questioning* option. The fourth option signals a possible *application* for this reagent.

Subsequent workers with cognitive preference tests include Marks (1967), Atwood (1966, 1969), Kempa and Dubé (1973), Mackay (1971, 1972), Tamir (1975, 1976, 1977a, 1977b), Tamir and Kempa (1976, 1978), Williams (1975), Rogel (1974) and Brown (1974). It will be apparent from the detailed survey of the literature in a later section that although differences in administration have occurred most basic principles of cognitive preference testing have remained unaltered from the original proposed by Heath. For example, it has been standard practice to prepare items of the type described by Heath using factual material which is already in the cognitive structure of the respondents and taking care that the linguistics are well within their intellectual range. It has also been customary to submit items for rigorous scrutiny by a panel of adults competent to assess whether the options correctly reflect the intended mode and to guarantee that the content will be familiar to respondents and then to pretest the items. The number of items in the tests has varied but each test constructor has endeavoured to sample the syllabus which had been used with the respondents prior to the moment of testing, as widely as possible.

It was tentatively suggested in <u>Section 1.1</u> that the act of making a preference for one option (or casting votes between all the four options) indicates the nature of the existing cognitive structure of the respondent. If the material in the option generates some conflict or tension with the respondent the option will be neglected or relegated to a low preference position in the rank ordering of the responses. If, however, the information in the option is in sympathy with the material already embedded in the student's cognitive structure there will be an unconscious attachment for the option and it will receive high preference. If similar attachment is felt for all the options of that mode over a test as a whole, that mode score will be high. Thus a high mode score may be taken to indicate both maximum attachment and minimum cognitive conflict with the mode in question.

It must be acknowledged that different intensities of attachment may be felt for the options representing a particular mode; for example the degree of preference for the selected option of one item may be different from that for the selected option of another item. It has not been customary to make allowances for this difference and, instead, preferences expressed for particular options have generally been accepted as being quantitatively and qualitatively the same. This matter is discussed in more detail in the following section.

It has already been stated in Section 1.1 that it is hypothesised that the response to a cognitive preference item involves a minimisation of cognitive conflict. This raises the question about the extent to which such conflict relates to the content of cognitive preference items on the one hand and to the mode implicit within each option on the other. Williams (1975) raises the question whether the selection of an option, say an A-type option, is inspired by the A character of that option or by the subject matter in the option without reference to whether it is A type or not. Tamir (1975) extended this to a division of the content element into a 'subject matter' aspect within a discipline and a 'discipline' aspect. Tamir's tentative hypothesis is that there are, therefore, three facets to cognitive preference response behaviour of which 'subject matter' and 'discipline' are two. The third is a general and stable characteristic of the individual student which is, in part, inherited and due, in part, to experience. Williams conducted an analysis of variance in order to determine the common variance attributable to preference as distinct from content and method. He found that preference contributed less to a score than either content or method but they do contribute sufficient variance to make them worthy of examination (Williams 1975, p.73-75). One part of this research was to seek evidence to clarify our understanding of the position. This is reported in Section 5.3.

Section 2.2 : Methods of Responding and Scoring

For his test of cognitive preferences, Heath (1964) required the respondent to select only the one most preferred option from each item. Thus, with a twenty item test, a maximum of 20 preferences could be expressed and these could be distributed over the four response modes adopted by Heath. The four mode scores thus have a high degree of inter-dependence and, in consequence, should be treated and evaluated with procedures applicable to ipsative scores only. This severely limits, as it did in Heath's case, the range of statistical procedures employable. In particular all, or most, of the conventional normative statistical techniques are, in the strict sense, inapplicable to ipsative data. Marks (1967) in an attempt to overcome this, prepared four parallel subtests in order to arrive at four mode scores which were not inter-dependent, i.e. were normative. Because of the large number of items required Marks' instrument was rather too lengthy for comfort and also appeared wasteful in terms of the limited use which was made of each item. Kempa and Dubé (1971) and most subsequent workers have required their respondents to vote on all four options presented in each item in the form that four votes are given to the 'most preferred' option, three votes to the next, and two, and one to the 'least preferred'.* Mode scores are then obtained by summing the votes cast on all the options of that mode. These mode scores are still ipsative but Kempa and Dubé felt that their scoring procedures ensure that better use is made of each item without loss of discrimination. Subsequent analysis of the data was done with normative procedures but the results were treated with considerable caution.

Footnote: This voting procedure is identical with rank ordering, except in that the numerical values are reversed, and with a system which requires two ticks for the most preferred option, a cross for the least preferred and one tick for the more preferred of the two remaining options (✓✓,✓,o,x). It was suggested that more reliable voting is obtained if the most and least preferred options are identified first.

King (1971) showed that Kempa and Dubé's scoring procedure is identical with pair comparison between six pairs. On the strength of this evidence a cognitive preference test by pair comparison was used in this research.

Mackay (1971) conducted an exhaustive study of scoring procedures and in the light of this he required his respondents to indicate only the most and the least preferred options of each item. Brown (1974) did likewise. Williams (1975) omitted the Questioning mode and worked only with the remaining three modes. He utilized a normative scoring procedure by instructing respondents to rate each option on a 6-point scale. He accepted the loss of a considerable degree of discrimination but he was able to place confidence in correlation analyses.

It is important to note that it the <u>total</u> score for each mode that is taken as the quantified orientation of cognitive preference. The contribution made by each item separately can be analysed for the purpose of identifying faulty items. Careful sampling of the entire syllabus which is available and the inclusion of the maximum number of items is also important. The actual number of items is governed by the time available for the administration of the test and by the degree of maturation of the respondents.

Section 2.3 : Evaluation of Cognitive Preference Data.

Part (A) : Validity and Reliability

The four constructs which Heath first identified have generally been accepted and they are a pre-requisite that,while other constructs may also be valid, it is not possible to validate the constructs with an instrument which is already formulated in terms of those constructs. Brown (1975) considered, also, that there is insufficient evidence of reliable relationship between the options and the modes which they purport to represent. Further, she pointed out that the preference orientation may have little bearing on how respondents <u>do</u> behave though it may be indicative of how they can behave.

It is a fact that cognitive preference tests have not been validated by independent measurements and it has therefore been possible only to establish 'face validity'. This has invariably been done by "judges", usually experienced teachers or academics who were requested to examine the items of cognitive preference tests to ensure that the construct of the mode was accurately reflected in the option to the exclusion of other constructs and that both the language and terminology and the syllabus content were familiar and readily intelligible to all members of the target populations. Although this procedure is generally satisfactory, it does not guarantee completely that pupils are fully familiar with the syllabus content reflected in the test items.

Many workers have assessed the reliability of cognitive preference tests; recently a summary of test-retest and α-Cronbach coefficients has been published. (These have been gathered together into a useful table by Tamir (Table 3, p.115, 1977a) and this shows that the values of these coefficients differ quite widely and it is possible that students may respond to linguistic features in the options. The table shows that the subject matter for the tests has been drawn from science subjects, from mathematics and, even, from social studies. The number of items has been as low as 10 and as high as 60 with consequent differences in reliability. Brown (1974) reported low coefficients for α-Cronback reliability, due, no doubt, to the brevity of her test and to the immaturity of her sample. In her consideration of the higher reliabilities obtained by other workers, she wondered whether verbal cues implicit in the wording of the item were responsible but Tamir has effectively disputed this.

Part (B) : Statistical Evaluation

It has already been pointed out that the usual voting procedure for responding to cognitive preference items leads to ipsative data. This in itself may be seen as a disadvantage, but it offers the important advantage of providing a discrimination between the response modes and thus enhancing the distinction between students in relation to their cognitive preferences. The consequence of ipsative scoring is that all conventional analysis and allied statistical procedures give rise to problems of interpretation, basically because the scores are inter-dependent. For example, Hicks (1970) demonstrated that random ipsative scores on a four item test automatically give correlation coefficients of -0.33; if the data are non-random it is found that this base is flexible. This lays severe restraints on the use of normative statistical procedures with ipsative data and the need for restraint has been fully acknowledged by several workers using normative procedures: for example by Kempa and Dubé (1973), and subsequently by Tamir, who employed normative procedures for analysis of their data.

Mackay (1972) endeavoured to interpret his data in terms of a different form of data evaluation known as multidimensional unfolding. The essence of this procedure is to locate the stimulus point associated with each response mode in a three-dimensional data space, thereby 'unfolding' the structure of the answered data. The general theory underlying this procedure was developed by Coombs (1964), although other workers had previously considered special cases, e.g. the location of three stimulus points and of four stimulus points (Wish, 1964; Bennett and Hays, 1960; Hays and Bennett, 1961). From a data-theoretical point of view, it is apparent that a stimulus can always be accommodated in an (n - 1) - dimensional data space. However, an attempt is invariably made by users of the multi-dimensional unfolding procedure to accommodate data in an (n - 2) - dimensional space, or a space containing even fewer dimensions, thereby 'simplifying' the data structure (this is essentially analogous to the process of factor analysis as used for normative data).

An interesting analysis of the data reduction of four stimulus points into two dimensions was made by McElerain and Keats (1961) on the basis of geometrical configurations. They developed a set of different geometric configurations into which four stimulus points could be located in a two-dimensional way, and provided guidelines for the choice of the 'best' configuration for a given set of data. Invariably, such reduction in the number of dimensions entails the rejection of some of the original data as 'not fitting' the model. The problem of the McElerain and Keats method is that it does not provide a reliable means of determining how much data which failed to fit the model could be tolerated without invalidating the model.

Section 2.4 : A Review of the Studies of Cognitive Preference Tests

It will be recalled from Section 1.4 that one of the specific aims of this work is to examine cognitive preference testing from three angles, namely validity, scoring procedures and item design, and test length. No previous workers in this field have been concerned with the psychological nature of cognitive preferences and there are no reports or discussion of the intellectual processes which are involved. One of the principal aims of this research is to examine the extent to which cognitive preferences possess the attributes of cognitive styles: stability of behaviour is an important aspect of the argument and in consequence a review of the literature on *stability* is given in Part A below.

Hitherto no mention has been made of the observation first reported by Kempa and Dubé (1973) and subsequently confirmed by numerous other workers that a relationship appears to exist between cognitive preference modes, such that those respondents with a marked preference for the Recall Mode tend to be associated with a low preference for the Questioning Mode, and also that a high preference for the Application Mode tend to be accompanied by a low preference for the Principles Mode, and vice versa. This is an important aspect of the second aim of this research, namely to investigate character and measurement of cognitive preference and the review of the relevant literature is given below in Part B. The literature on the aspects of cognitive preference which relate to pupil assessment and treatments (the third principal aim) is extensive. A summary of relevant facets is given in Part C of this review under four headings viz:- i) academic ability, ii) the origin of cognitive preference testing (curricula and teacher influences), iii) interests and occupational choice and iv) sex.

Part A : Stability

Cognitive styles have been shown to be fairly stable with time and under different conditions of intellectual functioning. Therefore, these are the characteristics that have become associated with the notion of cognitive style. In consequence, in an exploration of the question whether cognitive preferences have the quality of cognitive style, one must inevitably pay attention to the stability approach.

The stability of cognitive preference within each individual over a period of time has not been explored. Witkin, Goodenough and Karp (1967) in a study of cognitive styles (e.g. field dependence - field independence and the rod and frame test) found that the evidence showed that groups clearly maintained their positions on the score continuum over a long period of time (age ten to twenty four) even though there were marked changes in the scores of the groups. They observe that the evidence suggests a high degree of continuity during an individual's development in relative level of differentiation. Kagan, Moss and Siegel (1963) have shown that an individual's style of categorisation progresses from a relatively global approach toward one that is more analytical but the development is a steady and continuous one, a student generally retains his stylistic position relative to the other students. The cognitive style seems to be somewhat resistant to changes of a wholesale nature. It seems reasonable to anticipate, therefore, that if cognitive preference behaviour is to have the characteristics of a cognitive style, it should also be resistant to change both during the duration of the administration of the test instrument and between one administration and the next.

The stability of cognitive preference behaviour has four dimensions. Stability of response behaviour within a single administration is clearly one dimension. The values of the four internal consistency coefficients are widely reported and the values quoted are generally amply sufficient to inspire confidence. Thus, in five tests reported by Tamir (1977a, p.115) with 20 items or more, the mean values of the coefficients are 0.70 (.06), 0.66 (.07), 0.54 (.16) and 0.56 (.21) for the four modes in the order R, A, Q and P. (Standard deviation in brackets). For five tests with 30 items or more these values rise to 0.81 (.05), 0.71 (.08), 0.83 (.05) and 0.74 (.11) respectively. The measures which were used to generate these coefficients register the students' behaviour in putting the mode in question first but they afford him no credit for consistency in other placings. The scoring procedures which permit expression of preference for all items generate important information which is not utilised in computing the normal coefficients of consistency.

Stability within a discipline or academic subject but across the individual topics which constitute that discipline has been examined by Tamir (1975, p.254-257). He administered a 30-item cognitive preference test in biology to 989 students. The 30 items were classified into six subject areas: one area (botany) received 15 items and two others (human biology and micro-organisms) received only 4 items. Some items were allocated to two or more subject areas so that the students' behaviour was actually examined on the test as though it comprised 41 items. He comments that correlated t values derived from the scores "point at the existence of remarkable dependence of cognitive preference styles on specific subject matter areas" (p.256).

Part B : The Structure of Cognitive Preferences and their Measurement

i) Structure of Cognitive Preference

Kempa and Dubé (1973) were the first to report on the structure of cognitive preference scores. Principal components from factor analysis of the inter-mode correlation coefficients followed by varimax rotation gave two factors which accounted for all but 15% of the variance. On one of the factors the R mode was loaded heavily with -Q mode (and a small contribution from -P mode), on the second factor the A mode was loaded with -P mode. From this information they identified two bipolar, orthogonal scales. The R/Q scale was described as a 'curiosity' scale, and it distinguishes those who are satisfied by recall of the chemical information that is already familiar from those who have a preference for information that appears to raise queries about their knowledge and to present it with a different slant. The A/P scale distinguishes those with pure science learnings from those with more applied interests and it has been identified as a 'utility' scale. They found that high achievers scored high P and high Q scores, while low achievers scored high R and high A scores.

Very similar results were obtained by King (1972) using the same test with a fresh body of post O-level students and by Rogel (1975) who used the test after translation into Hebrew and administered it to an Israeli population. Tamir and Kempa (1976) and Tamir (1977a) used a 60 item test divided equally between chemistry, biology and medicine and their findings afford unequivocal support from Kempa and Dubé's dimensions. Tamir (1975, 1977b) also reports the results of the analysis of a cognitive preference test which was administered both to teachers and to students; there is further support for two scales from the analysis of the teachers' scores (N = 37) but not from the scores obtained by the students (N = 748). Tamir (1975) interprets the students' scores on slightly modified and more flexible scales whereby a high score on the Q mode might expect to be associated with a low score for either the R or P modes, and students with a low R mode score were generally found to have a high A mode score and/or a high Q mode.

There is no evidence of structure in Williams (1975)'s work and it would be reasonable to attribute this to his use of normative data. Mackay (1972a) subscribes to this view and on the strength of his reanalysis of King (1972)'s data he suggests that the orthogonal bipolar scales are merely an artefact of the statistical derivation. It has been shown subsequently that principal components factor analysis of the correlation coefficients on the three tests which he administered himself gave unambiguous support for Kempa and Dubé's structure before he modified the coefficients but it was the factor analysis *after his modification* which was confused. Brown (1975) contributed no fresh evidence on the subject but she registered deep misgivings on this form of analysis (p.58). She is equally unwilling to accept the 'unfolding analysis' as employed by Mackay (1971, Chapter 5) and first described by Coombs (1964). Unfolding analysis has the advantage of using rank ordered data. Mackay gathered data on large bodies of students and he analysed them both by unfolding analysis and by Q-type factor analysis. He claimed that the structure which was revealed showed orthogonal axes with the R mode opposite the P mode (not Q as Kempa and Dubé found) and with the A mode opposite Q. (Evidence will be produced later to show that Mackay's work was incomplete and that Kempa and Dubé's structure is at least an equally acceptable interpretation of his data). (See Section 7.1, Part B). Brown suggested that a Q-type factor analysis treated rather differently from the method used by Mackay might reveal the structure of cognitive preferences but she cautioned that loadings must exceed 0.917 for $p < 0.01$.

ii) Measurement

A brief account of some of the earlier work on this topic has been given in Section 2.2 and there are two aspects of measurement which are well worth reviewing. The point that ipsative data should only be subjected to the statistical treatments reserved for normative data if much caution is exercised in the evaluation of the evidence, has been made. The suggestion that cognitive preferences should therefore be measured in such a way as to yield normative data has the merit of permitting normative statistical evaluation but the method may lose discrimination. In a recent paper an account is given of the analysis of ipsative and normative data which has been gathered from the same population with the expressed intention of comparison of these two data types. Closs (1976) collected ipsative data on career interests from a large sample of undergraduates (N = 2200) by pair comparison between eight categories by career-type. The normative data was intended to gauge the intensity of that interest in careers and it was gathered by asking the respondents whether the options of the pairs were 'liked', 'disliked', or 'don't know'; a score of +1 was given to 'liked' options, -1 to 'disliked' and 0 to 'don't know' and a 'like-disliked' total was computed for each respondent. Closs used this normative data to identify three patterns of behaviour which he showed to be unresolved in the analysis of his ipsative data. He demonstrated that the distortion which occurs when ipsative scores are normalised or when 'percentile equivalents' are computed can lead to erroneous careers advice. Attention is paid to the difficulty experienced by respondents which arises from personal differences in response styles. Idiosyncratic differences in breadth of feeling generated by terms like 'like' and 'dislike' lead to differences that have only slight relevance to demonstrable differences in career selection. Closs uses normative data to identify respondents with differing intensities of feeling but he gives no indication of the distribution of the different response types within the sample. The general conclusion drawn by Closs is that there is little to choose, for a number of reasons, between the two types of data.

The second aspect of measurement which is of interest concerns the bipolarity since it has been suggested that there are artefacts of the pseudo-normative treatment of ipsative data (Mackay 1972). Tamir and Lunetta (1977) sought to clarify the point. They gathered cognitive preference data from a sample of high school students (N = 177). An established instrument was divided into two equal parts and scored differently to obtain the two types of data. Normative data was obtained by scoring on a four point scale. (It is presumed that only integer responses to the scale were provided for). Half the sample made ipsative responses to the first half of the instrument and normative to the second, while the other half scored the instrument in the reverse order. Moderately strong correlation coefficients were obtained between ipsative and normative mode totals. Inter-mode scores were correlated and it was found that coefficients with ipsative scores were in general less than those obtained with normative scores. Factor analysis of the intermode correlation coefficients of the four ipsative mode total scores and the four normative mode total scores showed loadings of all normative scores positively loaded on the first factor with no significant ipsative loadings, the second factor loaded R/Q with both ipsative and normative scores and the third factor showed significant loadings of ipsative A/P only. Their evidence is, therefore, that the bipolarities do exist with normative data and that normative data treatment does provide support for the ipsative results of many workers in this field, e.g. Kempa and Dubé.

Part C : Pupil Assessments and Treatments

i). Academic Ability

Brown (1975) examined the findings of Kempa and Dubé that on the basis of the two bipolarities claimed by them (R/Q and A/P) a major differentiation of academic ability categories appeared and on the strength of the findings she postulated a strong relationship between cognitive preference modes and the levels of academic ability in the hierarchies of Bloom (1956) and of Krathwohl et al. (1964). She carefully related the individual modes to appropriate levels on the scale and she used the relationship to predict that reliability would be highest for the R mode since, in her view, this is the least demanding academically, that the reliability of the P mode would follow, then A and finally Q since this mode was the most challenging and fewest pupils would respond to it with reliability. Her work with young students followed this pattern and supported her contention that students respond to preferred cognitive levels rather than preferred cognitive modes. Tamir (1977a, p.115) has gathered together the coefficients of internal consistency of previous cognitive preference studies. Examination of the twelve studies in which all four modes were used furnishes no other instance of coefficients in the order predicted by Brown and affords little support for her contention. Tamir examined the relationship further by a factor analytical study of cognitive preference data with cognitive abilities scores related to the three relevant Bloomian levels (Tamir, 1977a, p.116-118). He concluded that 'cognitive preferences are *not* expressions of cognitive abilities of the student'.

ii). <u>The Origin of Cognitive Preferences</u>

Heath developed the cognitive preference test for the evaluation of curricula. This use was taken up by later investigators, e.g. Marks, Atwood, Kempa and Dubé, King and Tamir. Implicit in the selection of a cognitive preference test for this purpose is the assumption that cognitive preferences are influenced by curricula and teaching approaches. Two aspects emerge: (a) curricular influence through exposure to different subject treatments, e.g. enquiry modes; (b) influence of the teacher through his/her emphases on particular aspects and facets, which may reflect the teacher's personality more than subject matter influence. Both aspects have been the subject of some research.

(a) <u>Curricular influence</u>

Tamir (1975) summarised the findings of six previous studies on the relationships between cognitive preference measurements and curricula (Table 1, p.239). In each of the six studies the cognitive preference of behaviour of a sample of students engaged in a modern curriculum is compared with the behaviour of a sample engaged in a traditional course; in two of the studies high achievers are compared with low achievers. Behaviour in the four modes is therefore compared eight times making 32 comparisons in all and significant differences are found in 23 of them. Thus for the R mode, the traditional sample scored higher preferences in four studies, in two the low achievers on a modern course scored higher preferences than the low achievers on a traditional course and in two there was no significant difference. The modern course favoured Q mode in 4 studies, the traditional favoured Q in one and there was no significance in the remaining 3 studies; for P the figures are 5, 1 and 2 and for the A mode they are 2, 4 and 2. There is therefore some contradiction but the evidence of a relationship inwhich modern curricula favour Q and P modes is quite strong. Elsewhere he provides evidence that confirms early work of Kempa and Dubé (1973) in which higher achievers were found to have high P and Q mode scores (Tamir and Kempa, 1976). He also found that a modern curriculum (BSSC) promotes curiosity even among low achieving students and that a high A mode score is associated with high achievement on a traditional course. Brown (1974) found no evidence of a difference between students who had studied an integrated science course and those who had been taught the three science subjects separately. King (1972) compared the behaviour of students who had received a Nuffield Chemistry course with those whose training to O level had been more traditional and he found that the differences in scores of cognitive preference (and critical thinking) were significant in directions which suggested that the Nuffield course, or some inherent property of it, did lead to the genuine and estimable differences which were anticipated by the designers of the course. Mackay (1971) gave sufficient information in Table 5.4.1 (pages 147-150) to demonstrate that the 640 students who had studied a PSSC course in the 11 and 12 grades had higher mean A mode and Q mode scores and lower mean R mode and P mode scores than the 239 students who were similar in all respects except in that they had studied a traditional course of physics but only the difference in A mode score was found to be significant ($t = 2.35$, $p = 0.02$). (The examination was done on Groups 5 and 6 in Mackay's data).

The conclusion is that while there is widespread evidence of relationships in the literature, the extent to which a modern treatment favours one mode at the expense of another is unclear but it is possible to generalise and say that modern courses do tend to promote the Q and P modes.

(b) Influence of Teacher

A number of studies have examined the relationship between cognitive preference scores and teacher influence. Schedemann and La Shier (1967-68) unsuccessfully attempted to detect relationships between pupil cognitive preference behaviour in physics and the pupils' perception of their teachers and between pupil behaviour and teachers' cognitive preferences. King (1972) was no more successful in demonstrating a post-instruction correlation between teachers' cognitive preferences and the scores of their pupils. Mackay (1971) found no evidence that the change of cognitive preference after a one year period could be related to the cognitive preferences of the physics teachers. Barnett (1974) found that non-directive teaching promotes academic achievement and he also found that differences in mode scores in the R, A and Q modes in groups distinguished by student perception of teacher directiveness were not significant. Tamir (1976) separated teachers into two groups by reference to their attitude to the use of new syllabuses in biology and he reported a number of interesting and significant differences in mode scores between pupils of those who supported the new programmes and of those who did not. The contention by Brown (1975, p.55) that cognitive preference patterns adopted by pupils merely reflect extraneous influences from the past, particularly those of the teacher, inspired Tamir to investigate the extent to which the claim was valid (Tamir, 1977b). He reported that he found little evidence that pupils conform in patterns designed to gain teacher approval. Vinton (1972) revealed that the greater the match between a pupil's preferred learning process (not cognitive process in this instance) and learning structure, the greater will be the learning style 'actualisation' and the more the individual will learn. Similarly Crawley and Shrum (1977) reported that physics students who experienced learning conditions compatible with their preferred learning structure register greater gains in preference scores (again, not *cognitive* preference scores) for physics than did those students who were in incompatible learning situations. (Students of chemistry, biology and geology also registered gains but they were not significant). Tamir postulated that the "high Q, high P, low R" scores of teachers promote similar behaviour in the patterns of those of their students whose patterns match theirs. The research evidence which has been cited would serve to encourage support for Tamir's hypothesis but there is, of course, considerable difficulty in "correlating" the response of a teacher with the average response of his class.

iii) Interests and Occupational Choice

Tamir and Kempa (1976) reported significant differences in cognitive preference behaviour between students who selected training in physics and biology from those who selected medicine and engineering. The former were found to be more curious and to have less preference for the R mode while the latter students were found to have higher preference for the Application mode. Williams (1975) found that 'vocational technical' groups selected the Application mode more strongly than other groups and he attributed this to the influence of the practically orientated course which they were following.

iv) Sex

King (1972), Williams (1975) and Tamir and Kempa (1976) examined the cognitive preference scores of males and females and found no significant difference. Kempa and Dubé (1971) used an all male sample, Barnett (1974) and Brown (1974) did not investigate this parameter. Tamir (1976) showed significant relationships between sex, academic ability and cognitive preference mode scores by two-way analysis of variance: for the Q mode, for instance, there were two significant differences between girls of high ability and girls of low ability, between boys of high ability and boys of low ability and between boys and girls both of low ability. Another two-way analysis of variance found significant relationships between type of school, mode scores and academic ability and extension of this analysis to include sex showed further complex relationships. Tamir did not offer any explanations for these findings. In the same study he reported that he found no relationship between mode scores and career expectation of Canadian college students but he found some distinction between Israeli students with a medical career in mind and those with an interest in pure sciences.

Section 2.5 : Cognitive Styles, a brief review of their nature and of the Literature

An account of cognitive style theory is included because the suggestion has been made that cognitive preference behaviour resembles in some important respects, behaviour which is more broadly classified as stylistic. The extent of the conformity of cognitive preference behaviour to cognitive style behaviour is an important part of this study. The method of study of human personality in terms of styles involves the separate study of each of the psychological levels at which a person functions in each of the many domains of human behaviour. The rationale for this approach is that the analysis of the cluster of characteristics subsumed under the differentiation process will, on ultimate completion, aggregate into a valid description of the whole individual. This is as valid for learning styles as it is for styles of behaviour from wider psychological fields.

This approach is justified only if experimental evidence of differentiation in one aspect of behaviour is repeated in cogent form in other areas as well. The concept that perceptual and intellectual tasks serve to assess broader and more salient dimensions of behaviour is deep rooted in the history of psychological measurement. Tasks of an essentially cognitive nature have been used in the measurement of facets of behaviour; many of the facets which the tasks illuminate have been designated 'styles of behaviour' or 'styles'. Cognitive style theory insists that complex behaviour can be partialed into a myriad of inter-related, self-consistent modes of functioning and that the sum of these modes will fully describe the perceptual and intellectual activities of the individual. The method approaches the individual from the perspective of the actual person engaged in the activities; experiment has shown that essentially the same ways of functioning pervade both perceptual and intellectual behaviour and learning and that these are characteristic of the individual.

Psychological differentiation is conceptualised as a high order construct; lower orders include perceptual and intellectual functioning, the articulated body concept, the sense of separate identity and structure and specialised defences (Wilkin, Ottman, Raskin and Karp, 1971). Each cognitive style is a subset of perceptual and intellectual functioning and is therefore a still lower order concept. Embodied in each cognitive style are two process, one is concerned with the perception of information and the individual's response to external stimuli and the second involves the processing of that information. The steps tentatively proposed follow this pattern.

Warr (1970) described "the habitual ways or modes of dealing with information about onself and one's environment which are to a large degree independent of the content of the information being handled" as cognitive styles. Messick (1970) described them as "information processing habits".

He claimed that they would "promise to provide a more complete and effective characterisation of the student than could be obtained from intellectual tasks alone" and he foresaw future research into the course of individual learning and of didactic interactions through the examination of cognitive styles. Leach (1967) in reviewing work on creativity commented that the term cognitive style, though variously defined, always embraced a flexible approach to life and marked response to environmental stimuli. Scott (1973) defined them as "the patterns that an individual may use in coping with environmental stimuli" and Field (1972) simply described them as "the pattern of thinking".

Ausubel (1968) in a more exhaustive study refers to "self consistent and enduring individual differences in cognitive organisation and functioning" as cognitive styles.

> "The term refers both to individual differences in general principles of cognitive organisation and to various self consistent idiosyncratic tendencies that are not reflective of human cognitive functioning in general. It reflects differences in personality organisation as well as in genetically and experimentally determined differences in cognitive capacity and functioning; and in a very real sense, it mediates between motivation and emotion, on the one hand, and cognition, on the other". (p.170).

He goes on to comment that the inter- and intra-task generality of function of the measures has not been adequately established. This was true at the time when it was written and much subsequent research principally by Witkin, Gardner, Kagan and co-workers has shown that cognitive styles are indicative of stable and generalised cognitive traits.

Finally, in a lengthy review of the educational implications of cognitive styles Kagan (1971) examined the rôle played in education by each of the nine cognitive styles which he describes. He observed that the research evidence of interaction between the variables of classroom instruction and cognitive styles is limited and he found little evidence of shared variance with cognitive styles and traditional indices of ability.

He drew an interesting distinction between styles which possess the quality of a *capacity* (e.g. analystic - global dimension) and are found to be resistant to change by didactic means, whereas those having the property of a *strategy* (e.g. conceptual groupings in sorting tasks) are more readily modified. Scott (1973) has since reported that exposure to two or three years of enquiry strategy in science classes did enhance the analyticalscores of elementary and junior high school students in tests of categorization style. The improvement was significantly greater than that of a control group which was denied the training.

Kagan found strong evidence that the traditional measures of verbal and mathematical ability have little predictive value beyond their immediate classroom context while measures of associative thinking are related to a wider range of accomplishments. He concludes that

> "though the practical pay-offs of cognitive-style research are still quite meagre for the teacher, the promise of future rewards is strongly indicated if energy and imagination are applied to the task". (p.292).

Section 2.6 : Problems for Investigation

The aims of this research have been stated broadly in Section 1.3 and the research problems will be outlined in detail in Section 4.1. It will suffice here to indicate the nature of the problems in so far as they relate to cognitive preference. The first group of problems, called measurement problems, are essentially technical in that they relate to issues such as the nature of cognitive preference and the design of the items which are used to measure it, to the scoring procedures and to the statistical analysis of the data which is obtained and also to the validity of the instruments. All these problems are similar in that they do not relate to the dimension of time and they are therefore grouped together in Chapter 5. The problems for research which will be found in Chapter 6 are all concerned with the structure of cognitive preference and their use for the evaluation of pupils. Earlier workers have found conflicting evidence and as the work is of considerable importance in determining the psychological nature of cognitive preference much attention will be paid to resolving these conflicts. Further work on the actual nature of cognitive preference has been assigned to Chapter 7 but in this chapter the interest centres on the change, if any, in behaviour over a span of time. The suggestion has been made (King 1972) that cognitive preference behaviour has the characteristics of cognitive styles.

The use which cognitive preference tests might have for teachers in assessing the behaviour of individuals and treatments has received scant attention. Much interest rests in their potential for classification of students into meaningful categories beyond the divisions which are already possible through traditional tests. The study which will be found in Chapter 9 concerns the broad inter-relations that were thought to exist between the main variables which are examined in this research (i.e. cognitive preference variables and judgement variables). The study concerns their relationships to one another and, in Chapter 10, to other variables which have important connections with classroom behaviour and achievement.

CHAPTER 3 : JUDGEMENT

Section 3.0 : Introduction

The aim of this chapter is to scale down the broader aspects of judgement to the narrow aspects which were researched. In Section 1.2 a concise account was presented of the relationship seen to exist between the third stage of the learning sequence, namely information transformation, and that part of the judgement process which was referred to as the assessment of reasonableness. The intention in that section was to postulate how judgement might form an important part of the learning sequence and to show it might be relevant to the alteration of an existing cognitive structure. The rationale was that learning is synonymous with modification to the existing cogntive structure. In this chapter the issue of 'reasonableness' is tentatively related to the hypothesis generation aspect of judgement since the two aspects were studied separately and in relation to each other.

The first section (3.1) develops the ideas of judgement so that the inter-dependence of 'assessment of reasonableness' and 'hypothesis generation' can be appreciated. The literature review in Section 3.2 concerns first the literature which deals with philosophical aspects of the study of judgement and it embraces reports of work by Kelly (1955) and by Ennis (1962). These reviews are given so that the broad context in which work of this type is undertaken may be appreciated. The work of Peel on the Judgement of Adolescents (1971) is quoted extensively in the practical section of the review because of its strong relevance to the work which was undertaken here. The last two reviews are from the work of Abercrombie (1960) and elsewhere from Peel (1971) and they were included because both were concerned with the processes and strategy that judgement situations invoke.

The report in Section 3.3 is concerned with earlier work done by King in which he attempted to explore judgement and critical thinking, since these qualities had been advocated as deliberate outcomes of modern curricula. The work was valuable in that it led to a refinement of the process for the assessment of reasonableness and revealed the new horizons which were examined here. These are listed in the final section.

Section 3.1 : The term 'Judgement' as it is understood in this study

It is important to establish the meaning of the term judgement. For the limited purpose of this study, judgement is defined as a logical operation of thought which embodies some measure of creative process and possibly some degree of value or improvement to the individual engaged in the act of judgement. Valuing , assessing, forming an opinion, critical thinking and evaluating are allied thought processes: in judgement alone, a strand of thought is generated from a body of information which seeks to unify the evidence. The thought itself is not in the evidence and the act of thinking itself is an act producing new insights. It is a highly egocentric process and it demands the highest level of intellectual activity (as judged in Bloomian terms). The condition of self improvement requires that a judgement must be made and a decision must then be taken in the light of that judgement.

Judgement is the facet of human thought in which the new and unique features of a particular situation are assessed and evaluated in the light of previous experience. Novelty of the information to be evaluated is essential to judgement. If the situation has no aspect that is in some way novel, the activity may be more accurately described as evaluation: an act of evaluation by a person with wide experience may well be judgement for another with limited or restricted experience.

The point in the foregoing paragraph may be amplified in relation to the formulation and evaluation of hypotheses. A hypothesis is a generalisation; the generation of a hypothesis involves a strand of novelty in that it represents some kind of intuitive leap from accepted fact to generalisation. Testing the extent to which a hypothesis is tenable requires the exercise of evaluation and if inconsistencies between the reliable data and the novel generalisation are revealed, the hypothesis is rejected. There are two elements in the process in which judgement is embodied. The first is the intuitive leap, and the second is the evaluation of that leap. If there are already sufficient instances of the fact in the reliable data at the person's disposal the unifying generalisation is essentially called a conclusion. The creative act is then the generation of the conclusion, and the evaluation decrees the tenability of the conclusion. In practice one hypothesis may embrace a succession of minor hypotheses which are all "chained" to one another by the circumstances of the problem.

An attempt may now be made to apply this to science learning. It may be argued that the first step in the operation of scientific method is the aggregation of observations. There follows the acknowledgement of a problem which is related to those observations, whereupon a hypothesis (see Footnote) is generated and subsequent experiment is devised to test their validity. It is reasonable to suggest that both the number and quality of the hypotheses vary greatly from person to person when confronted with the same data. (Indeed, evidence in support of this will be produced later). It seems that both facets are personal to the individual who is engaged in the act of making a judgement although they are not normally articulated, nor differentiated.

The second element in the overall act of judgement is, as has been indicated, an evaluative act. In the context of this study, this has been restricted into an 'assessment of reasonableness' but even this has a number

---

Footnote: Although the term 'hypothesis' is used here, it should be understood that, under certain circumstances, the term 'conclusion' might be more appropriate. This would be the case if the range of observations is so extensive that the process of generalisation becomes essentially 'deductive' in nature, rather than predominantly 'inductive' as would be the case in true hypothesis formation.

of dimensions. It will be evident that a hypothesis cannot be regarded as reasonable if it is in any way in conflict with established scientific facts or principles in so far as these are known to the person making the judgement and in so far as he understands and interprets them correctly. The judgement may be false but if no conflict is apparent to the judge he will be unaware of the error. Similarly there must be no conflict with the information which is available on the specific problem and again the judge may make a false judgement without creating conflict. Besides being true to scientific fact and principle in both broad and narrow contexts, a judgement can only be deemed to be reasonable if it is strictly relevant to the problem in hand and it must also make complete use of the data; a judgement which makes anything less than full use of the data is inadequate. Again it is possible that the judge may be either unaware of the irrelevancy or inadequacy in which case there is no cognitive conflict,or,if he is aware of it, he may select the hypothesis which serves to produce a minimum of congitive conflict.

Section 3.2 : A survey of the literature on Judgement and Critical Thinking

The literature on judgement is extensive but the problem is that the word judgement is used in several different meanings. Much of the literature is concerned with judgement of personality, with consistency and with value judgements of intrinsic qualities. The importance of judgement in the statistical sense (like levels of probability and adequacy of sample size) has also received much attention in psychological work. However, the treatment of the cognitive aspects of judgement is less extensive.

Judgement was seen by Bloom (1956) as a part of his category of "evaluation" which he placed at the highest level of his Taxonomy of Educational Objectives. To quote Bloom "it (i.e. judgement) is regarded as being at a relatively late stage in a complex process which involves some combination of all the other behaviours". Bloom subdivided his level into two; judgement in terms of internal evidence like accuracy and logical fallacy, and judgement in terms of external criteria. The latter embraces a wide range of activities including the judgement of aesthetic attributes and of the merits of alternative courses of action. Cognitive judgement as such is mentioned only as "the comparison of generalisations" (p.192).

Kelly (1955) considers that the process of judgement follows a cycle of circumspection, pre-emption and control. In his view, circumspection entails mastering the environment and then generating a series of propositional constructs to establish the diversity of the possible decisions. Pre-emption results in the selection of the most relevant axis on which to construe this situation and control leads to the selection of the one alternative. He devised a test of cognitive complexity in which the constructs for the assessment of character were provided. Kelly assumed (and Bieri et al. (1966) subsequently verified) that the use which is made of the constructs is representative of the constructs available to the assessor, i.e. of the cognitive complexity of the assessor.

Judgement and critical thinking are terms often used synonymously in research reports. . Therefore, consideration must also be given to the work concerned with critical thinking. In his careful study of the abilities which constitute critical thinking, Ennis (1962) used three dimensions: a logical dimension, a criterial dimension and a pragmatic dimension. He claims that these "if appropriately weighted and empirically supported" furnish a framework upon which the whole study of critical thinking can be based. The logical dimension covers alleged relationships between meanings of words and statements. Thus, a person who is competent in this dimension knows the meaning of the basic terms in the field in which the statements are made. The criterial dimension is rigorously defined into twelve subdivisions, called aspects of critical thinking, and further subdivision of many of the aspects is given. The pragmatic dimension covers the background purpose of the judgement and it covers the decision whether the statement is good *enough* for the purpose.

Part of the 8th aspect is quoted here as an example, not only because it explains the approach to critical thinking which Ennis used, but also because it is relevant to the exercises developed for this study.

> *Judging whether an inductive conclusion is warranted*
>
> Inductive conclusions are of three types, simple generalisations, explanatory hypotheses, and theoretic systems. Though similar in many aspects, they are different enough to warrant separate treatment.
>
> 1. A simple generalisation about experience. Such a generalisation is warranted:
>
>   .11 to the extent that there is a bulk of reliable instances of it
>   .12 to the extent that it fits into a larger structure of knowledge
>   .13 to the extent that the selecting of instances is unbiased. (Four methods of obtaining unbiased samples are given)
>   .14 to the extent that there are no counter instances
>
> For example, the generalisation that red-headed people tend to have hot tempers would be warranted to the extent that there is a large number of reliable instances of red-heads with hot tempers, to the extent that we are able to account for red-heads being hot tempered, to the extent that our instances of red-heads are picked without bias and to the extent that there is a lack of reliable instances of red-heads with even tempers.
>
> The *criterial dimension* is involved here through knowledge of the above criteria. There is a vast literature on judging the adequacy of samples, which very rarely are purely random.
>
> The *logical dimension* is involved in the recognition of instances and the application of the above principles. And the *pragmatic dimension* is invoked in deciding that there is or is not enough evidence for the purpose of the inquiry.

The paper continues with discussion of explanatory hypothesis and theoretic systems in similar vein. The section on hypothesis concludes with this paragraph:

> But most important of all, the so called 'inductive leap' utilizes this (the pragmatic) dimension. Deciding that the evidence at hand is *enough* to establish the hypothesis requires consideration of the degree of satisfaction of the criteria, the purpose, and how important it is to be right, then the leap may or may not be made. (p.95).

Later in the paper, Ennis goes on to say that he would expect a sizeable correlation with IQ even though all three dimensions are learnable rather as academic subject matter is learnable: he expects the logical dimension to have the highest correlation with IQ and the pragmatic dimension to have the lowest. He predicts that age will correlate well with the logical discussion with children in elementary and secondary schools and ,as experience might be expected to strengthen pragmatic and criterial thinking, these might be expected to correlate more highly with age after the age of sixteen.

A different approach to the analysis of thinking (and judgement) is made by Peel (1960) in his book "The Pupil's Thinking." The first chapter of his book is devoted to an analysis of the different kinds of thinking. Peel identifies three kinds of thinking which one might expect to encounter in the classroom. (There is a fourth which is primarily the preserve of only the most lucid thinkers at the sublime moment when man's thinking makes a significant surge.) Thinking which is free from constraint is described by Peel as *thematic*. If the thinking is constrained by having to describe and explain events and things, it is classified as *explanatory*. Finally *productive* thinking is said to have occurred at the point when explanatory thought has provided a jumping-off point for control over or manipulation of the environment. This kind of thinking gives rise to material changes, inventions and new products. It is invoked in school when a student is told to apply his knowledge or to make use of his explanations in a new situation. Many problems in science require this type of thinking; for example, the analysis of an unknown compound in the laboratory may require it to a marked degree.

The conditions which, according to Peel, must be fulfilled by thought when it is to be classified as productive, are five in number.

(i) It must contain an element of forward thinking beyond mere explanatory thought.

(ii) It becomes effective by changing the problem situation *materially* in order to achieve solution.

(iii) Some old problems may be solved by restating them in the light of the established explanation.

(iv) It always appears when a new problem situation is met.

(v) These situations may be material, social or personal.

It has already been argued that the learner's cognitive structure is modified in the process of making a judgement. If this does not happen, the information is merely accommodated and incorporated into an already existing cognitive framework. Thus existing concepts and notions provide adequate explanation of or for the "new" information. Peel's first point is that productive thinking "must contain an element of forward thinking beyond mere explanatory thought". Explanatory thinking clearly represents the accommodation of information within an existing cognitive framework, without the latter being materially modified or altered. "Going beyond mere explanatory thought" must therefore mean that in the process of absorbing new information and applying existing knowledge to problem-solving, this knowledge itself is modified in its conceptualisation: hence, the cognitive structure is altered. This is essentially the interpretation of judgement which is used in this research.

In his later book, 'The Nature of Adolescent Judgment', Peel (1971) addresses himself directly to the issue of judgement, but in so doing makes only a passing mention of his own earlier work (Peel, 1960) and that of other workers in this field. Of particular interest, in relation to the present study, is Peel's definition of judgement. He sees judgement as a form of thinking which is invoked in those situations where there is no ready-made solution, nor is there necessarily a single correct response "but rather a spectrum of responses satisfying different number of different criteria". Again, this view is fully compatible with that adopted in this study and proclaimed earlier.

It is also of interest to make brief reference to the strategy adopted by Peel and co-workers (mainly higher degree students in Peel's department at Birmingham University) in their investigations into adolescent judgement. Their technique was to present short passages to students. The questions which followed each passage could not be answered adequately on the basis solely of the information which was supplied and the student was obliged to draw on independent experience in answering the questions. The analysis of the answers led them to propose three basic response categories, each representing a different form of judgement:

Type of Judgement	Characteristics
Restricted	- tautological, premise-delaying, irrelevant
Circumstantial	- bound solely by the content of the passage, often taking account at first of only one element
Imaginative-comprehensive	- involving the invocation of independent ideas and the consideration of the problem in their term

The results obtained by Peel and others indicated that there is a noticeable increase in imaginative judgements in the year between the 13th and 14th birthday and that the most mature judgement is not established until a chronological age of 15 (a mental age of 16) years has been reached. Furthermore he finds that maturity of judgement is associated with the socio-economic level of the home. Also, the provision of more information in a passage tends to spark off more acts of judgement than passages in which slightly less information is given, and identical passages with responses in multiple choice format evoke many more imaginative responses than do open-ended questions starting with 'why?'

These are findings of considerable interest, both in the psychological and measurement sense. A most important finding, of particular relevance to this study, resulting from Peel's work is that each subject responded with a fair measure of consistency when working in passages of similar quality. This is clearly indicative of a "stability" aspect in the exercise of judgement ability, which suggests that judgement may have a 'styles' quality. In other words this finding of Peel's points to the possibility of judgement styles which, like other styles (and especially cognitive styles), appear fairly stable. Peel also reports that a high degree of consistency of the category of judgement exists across different test situations, and this again lends support to the 'styles' argument.

Passing reference may be made here into the work of Bruner, Goodrow and Austin (1956) into concept acquisition. They found that subjects, in trying to identify concepts from given information, tended to use particular strategies (of different types) with considerable consistency and stability. The particular strategies identified by Bruner et al. were (a) simultaneous scanning, (b) successive scanning, (c) conservative focussing and (d) focus gambling, with the subject being free to use any one of them for the purpose of concept identification. Although Bruner et al. did not see this work in the context of judgement as such, it is evident from the process of concept acquisition that this involves judgemental behaviour. Therefore, their finding of a relatively consistent and stable use by individuals of particular strategies again supports the 'styles' argument.

In Chapter 4 of his book on adolescent judgement, Peel considers processes involved in making a judgement. It is beyond the scope of this review to give a full discussion of his arguments. However, it is of interest to point to his view that the most mature thought requires imagination and the generation of hypothesis and that the number of imaginings is some measure of creativity. Selection of the most appropriate hypothesis and rejection of the remainder is then required to resolve the conflict between the student, as a processor of information, and a problem situation containing previously unexperienced information. He also makes the most interesting point that "it may be inadequate and misleading to talk of observations as objectively unasssailable data, since they may be partial and selected as a result of direction by the hypothesis held" and asserts that reasoning from hypotheses is a most sophisticated thought form calling for experience and intellectual maturity in the thinker.

He concludes the chapter with a summary of the capabilities required for intellectual judgement:

This work of Peel's is essentially in agreement with the definition of judgement which is given in Chapter 1 with the stress on hypothesis generation and the requirement to assess the "reasonableness" of the hypotheses.

Section 3.3 : A Measurement of Judgement.

The exercise in the measure of judgement which was used by King (1972) had its origins in the various notions of judgement in the works which have been quoted. The main focus of this work by King was to examine critical thinking and judgement as it developed in the context of the chemical education of O level students. New chemistry curricula stressed the desirability of pupils acquiring critical thinking and judgement abilities, but left unanswered the question whether these qualities were actually acquired. Against this background, the study by King was conceived and carried out. It involved, among other things, the clarification of what might be meant by "judgement" in the chemical education context, and the developments of appropriate judgement tests.

The first exercise of judgement (called "Judgement Exercises") was written at the University of E. Anglia in the Spring of 1972: The first part of the exercise had ten problem situations each of which was followed by six statements representing conclusions or hypotheses. The student was required to judge and decide whether a statement was

(a) sound and satisfactory in that it might serve either as a partial explanation of the information or it might help to resolve the problem,

(b) sound and satisfactory but irrelevant because it either fails to explain the information or it fails to help in resolving the problem,

(c) unjustified either for chemical or other reasons or because use was made of information that should have been given.

Each student was awarded a 'conclusion score' and a 'hypothesis score' but early analysis of the data showed that the distinction between a conclusion and a hypothesis to be an artificial distinction to the student, and therefore this distinction was dropped for all subsequent analyses.

The second part of the exercise, called 'Next Line of Action', had eight chemical problem situations. Four possible actions were appended to each situation and they were rank ordered by each respondent in the order in which they were thought to be most likely to elucidate the problem which was inherent in the description of the situation. In each situation one of the four options proferred the opportunity to obtain more information by a fresh experiment on a relevant parameter, an experiment on an irrelevant parameter was contained in another option, recourse to another source of information was offered in one of the four options and spurious information was furnished by the fourth.

The findings from the study of chemical judgement behaviour have been described elsewhere (King's M.Sc. thesis, UEA, 1972), and no details need to be given here. However, it is probably justified to point out that, notwithstanding minor shortcomings in the test materials developed for the study, much new ground was broken in that for the first time judgement in a chemical education context was examined and measured, whereas previously generalised judgement tests (e.g. the Watson-Glaser Critical

Thinking Appraisal) had uncritically been accepted as valid measures.

Section 3.4 : Problems for Investigation

The problems for the present study have been touched upon in Section 1.3 and they will be described in greater detail in Section 4.2.

The act of judgement is exercised by giving the student some information about a chemical situation which is beyond his previous experience. A question is posed and six, or so, statements are made. It is axiomatic to this research, for the reasons given in Section 3.1, that the student makes an act of judgement by (a) creating a hypothesis (or more than one hypothesis), to "explain" the chemical information and (b) evaluating the statements in the exercise singly in the light of the hypothesis that he has made. If he considers that more than one hypothesis is tenable he must decide whether to proceed with both or reject one. He may find that clues embodied in later statements help him to reject one and he may have to re-evaluate earlier statements if he changes his hypothesis during an exercise.

It is also axiomatic to this research that judgement can be classified in terms of the four attributes, labelled sense, relevance, adequacy and justification. Two of these attributes (namely sense and relevance) are concerned with the *scope* of the judgement and the other two are concerned with *quality*. In the broad sense the judgement will be considered to be good if it conforms to the body of established scientific experience, i.e. it must be *sensible*. In the limited sense of the particular judgemental situation, the judgement must also be *relevant* in that it relates the information available to the problem which has been posed. The actual quality of the judgement is measured by the extent to which the judgement makes full use of the information and if it does so it is said to be *adequate*. A timid judgement fails to make full use of data but one which requires excess information beyond both the body of knowledge which might be expected in cognitive structure and the information in the item is said to be *unjustified.*

The measurement of judgement for the purpose of this study is the sum of all the acts of judgement in the test document. This contains some test situations which are theoretical and others which are practical since the study of chemistry in the classroom and laboratory requires both these activities. Sub-test scores which arise from the attributes described above will be prepared and analysed in order to explore the psychological differences that there may be in the attributes. The validity of the test is difficult to establish because there is no recognised test of chemical judgement but comparison with a similar test is made in Section 8.1. The number and quality of the hypotheses are examined in Section 8.2 and the measurement of judgement both alone and with other variables is reported in Section 8.3. The differences between 'scope' and 'quality' are also considered in this section.

The act of judgement is also seen to have wider implications. In particular it was proposed, most tentatively, in Chapter 1 that judgement ability constitutes that part of the learning process called information transformation. If this proposal is to be sustained some bearing of judgement upon academic achievement must be expected, and, also, the relationship between information transformation and the other part of the learning process which is incorporated into this work, namely information perception, was also expected to be a fruitful avenue for research. Judgement in its restricted sense is confined to Chapter 8 but these more global aspects are treated in Chapters 9 and 10.

CHAPTER 4 : THE RESEARCH PROBLEMS AND STRATEGY

Section 4.0 : Introduction

The learning sequence was introduced as the principal theme of this study in Chapter 1 and two of the three stages of the sequence were described at some length in Chapters 2 and 3. At the end of these chapters some problems for investigation were posed in outline form. The purpose of this chapter is to describe these problems in more specific terms and to identify additional problems which may be resolved by unifying the evidence from the two separate strands of the study, dealing respectively with cognitive preferences and judgement. It will be recalled from the earlier chapters that the first stage of the learning sequence, information reception, could not be distinguished from the second stage, information perception, in operational terms that would have permitted investigation of the two stages separately. For this reason no work has been done on the first stage. In Chapter 2 the precise nature of the second stage was described and it was related to cognitive preference behaviour; the relationship between the third stage and judgement was explained in Chapter 3.

It will be recalled from the brief summary of the problems for research given in Section 2.6 that the research problems which relate to cognitive preference are of three broad types. The first of these are technical in nature. It relatesto the actual measurement of cognitive preferences and include statistical evaluation and item design, and it will be identified in Section 4.2 and the results will be reported in Chapter 5. The second type is concerned with the structure of cognitive preference behaviour and will be reported in Chapter 6. The third part of this work is contained in Chapter 7 and embraces the dimension of time. It concerns the change, if any, of cognitive preference behaviour in the same sample population in order to explore the possibility that cognitive preference behaviour may be interpreted as a cognitive style.

It will be evident from this brief survey of the work which will be reported that two different strategies were required. The first required the use of populations of students with sufficient experience in chemistry to ensure that the examination of preference was a valid exercise and to furnish sufficient chemical background for the exercises in judgement and also with sufficient time available for testing in depth on both the major and the minor variables that were considered to be relevant. (These latter variables are reported in Section 4.5). The populations were subjected to 'one shot' testing and the data which was obtained furnishes evidence in Chapters 5, 6 and 8.

The second strategy required that the same population of students should be available over a period of years for 'longitudinal' testing. It was considered essential that all the students should have some experience of chemistry before measurement by the first test of the series. It was also deemed to be important that the series of tests should coincide with a period of proven change to strengthen the case for the cognitive style issue. The tests which were used for the 'longitudinal' study with the same students were reused for comparison with three separate bodies of students of equivalent ages and class levels.

The sample populations were drawn from several secondary schools. The test series for the 'longitudinal' study started when the population had completed the IIIrd Form year, the second test battery was administered to them at the end of the IVth Form year and the final battery coincided with the 'O' level examinations at the end of the Vth Form year. All the students in the samples were engaged in an 'O' level course in chemistry during this time. (There was a considerable decrease in the size of the sample after the IIIrd Form because some students discontinued chemistry). The use of chemistry for the test medium was fortuitous: other factual material could have been used. Both the 'one shot' and 'longitudinal' studies required information from a number of ancillary tests. These are described with the justification for their inclusion in Section 4.4 and a full description of all test instruments is included here. A brief account of the test samples and the programme which was used in order to resolve the research problems is given in Section 4.3. A summary of the investigations using evidence from all measurements together with special intent to seek out aspects with direct relevance to the classroom is contained in Section 4.5

Section 4.1 : Information Perception

The purpose of this section is to amplify the three broad areas of research relating to cognitive preferences which were described in Section 2.5, and develop them into more specific areas and then into individual research problems. The three problem areas were identified as (a) measurement of cognitive preferences, (b) the structure of cognitive preference and (c) their stability and possible character as a cognitive style. The evidence which relates to these area will be reported in Chapters 5, 6 and 7 respectively.

Part (A) : Measurement Aspects of Cognitive Preference

It was pointed out in Section 2.3 (A) that cognitive preference instruments cannot be fully validated by independent measurements and that it has been the practice invariably to establish 'face' validity by employing experienced teachers to scrutinise the tests. The scrutineers were asked specifically to ensure that the construct of the mode was accurately reflected in the option to the total exclusion of the other three constructs.

Research Problem (A.1)

It may reasonably be argued that (a) the modes in a cognitive preference instrument are not perceived in the manner intended by the designer of the instrument and by the scrutineers and (b) the four individual options of the items are sufficiently 'mode specific' to ensure the validity of the test instrument.

The restraints which are imposed on the analysis and evaluation of cognitive preference data by its ipsative nature have cast doubt on the strength of the evidence. Confidence in the research evidence requires that the procedures used with cognitive preference data have secure statitistical foundation. It may be recalled from the discussion in Section 2.2 that cognitive preference tests may be scored in a number of different ways. Several workers including Heath (1964) required the respondents to select the one most preferred option. Kempa and Dubé (1973) specified that all the four options of each item should be rank ordered, Mackay (1972) and Brown (1974) asked for only the most preferred option and the least preferred option of each item.

Research Problem (A.2)

To explore the degree of confidence that resides in normative treatment of ipsative data. This is to be done by first measuring cognitive preferences on a normative scale, then to rank order this data in order to obtain both normative and ipsative data for the same preferences, to subject both sets of data to the statistical procedures which are normally employed for normative data only and finally to compare the evidence from both.

Clearly the length of the test instrument is relevant to this problem. Investigators who used moderately sized tests have invariably reported adequate internal consistency of the answers, suggesting adequate reliability. Brown queried the reliability aspect on the basis of her own findings, but she paid inadequate attention to the fact that her data was derived from a test with only ten items. Potentially, the question of how test length affects test reliability in the case of cognitive preference instruments might be thoughtby some to require attention. In the author's view, however, this is not so: the general issue is adequately covered by the Spearman-Brown prophecy relationship (the longer a test the more reliable it tends to be), and no a priori reason exists to assume that this relationship would not also be fulfilled in the case of cognitive preference tests.

The (pragmatic) decision taken for the purpose of the present research was to use cognitive preference tests with a minimum of 24 items.

Attention has also been drawn to Tamir's claim, made in 1975, that cognitive preferences have the following elements: general, disciplinary, and subject specific. The first is characteristic of the individual student and Tamir postulated that it may have two components, one inherited and the other acquired through experience. Tamir went further to postulate that it is the discipline dependent element which is measured by a cognitive preference instrument. It is the intention of the administrator that the instrument which is used should sample the discipline sufficiently broadly to ensure that the third element that Tamir proposed, the subject specific element, should have no specific bearing.

Tamir's claim was made more in the nature of a prediction than a conclusion reached from direct research enquiry. In the context of the present study, the opportunity was seen to exist (and taken) to examine one particular aspect of the Tamir claim, namely the subject-specifity of cognitive preferences. It may be argued that, if cognitive preferences do indeed contain the subject-specific element suggest by Tamir, items with common subject matter should give rise to preference scores which correlate more strongly with one another than with preference scores from items with no common subject matter.

Research Problem (A.3)

Cognitive Preference items will be examined to determine whether subject-specificity can be demonstrated.

Part (B) : The Measurement of Cognitive Preference and the Evaluation of Pupils

Part I : Structure (Research Problem (B.1)).

The structure of cognitive preference behaviour has already been established by earlier workers and their findings are reported in Section 2.4 Part B. The structure described by Kempa and Dubé (1973) involves two bipolarities which were identified as an R/Q scale (curiosity scale) and an A/P scale (utility scale). Confirmation of the evidence using normative data was sought and the alternative structure reported by Mackay (involving an R/P scale) when a different analytical procedure was employed, was thought to merit investigation.

Part II : The Evaluation of Pupils (Research Problem (B.2)-(B.7))

(Research Problem (B.2)

Kempa and Dubé (1973) found a relationship between cognitive preference orientation and academic achievement as measured by 'O' level grades. They reported that students with high achievement preferred the Q and P modes while those with low achievement preferred the R and A modes. Kempa and Dubé's measure of academic achievement (O-level grades), though convenient to use, can only be regarded as a coarse measure for it is not known what qualities such grades reflect. A further point, made by Brown (1975) for example, concerns the question whether the expression by a person of his cognitive preference might not simply be an expression of his academic ability. This type of question is no doubt prompted by the use of two terms in connection with cognitive preference testing which are also used by Bloom in his Taxonomy of Educational Objectives (the two terms are :- Recall and Application). In consequence evidence was sought to establish whether the relationship was as strong when academic achievement was measured by scores in a test of recall of chemical information and when measured in a test of understanding of chemical principles.

Kempa and Dubé's finding about the relationship between academic achievement and cognitive preference orientation applied to students subsequent to their completion of the 'O' level chemistry course. In view of the fact that for the present study a somewhat younger population was used,(viz IIIrd formers) it cannot automatically be assumed that the same or a similar relationship would apply here also. This, of course, provides an additional justification for the pursuit of this part of the investigation.

(Research Problem (B.3))

One difficulty which is inherent in the measurement of chemical achievement in a situation in which the student population originates from several different schools, stems from the likelihood of different curricular structures existing in the different schools. Although tests of chemical/science achievement can be designed which, at least superficially, take into consideration such differences in so far as they are discoverable by interviews with teachers, full certainty about the 'fairness' of these tests cannot be gained since curricular differences may exist in different schools. For this reason, it was thought essential in the context of this study to obtain scores on a more global measure of ability/aptitude as a reference measure. The tests which were selected for this purpose were the Bristol measures (Achievement Test, Study Skills, Level 5) since these are broadly 'science' related.

Attention has already been drawn to the point that the stability of cognitive preference can be conceived of in relative terms, i.e. where the position of an individual in relationship to others remains fairly constant with time, despite there being a gradual overall shift in such orientations with time. If, indeed, such a gradual overall shift takes place (evidence for or against this was sought as part of the study), the question arises what causes or might cause such shift. A number of obvious variables and factors present themselves and these included i) pupil maturation as measured by class level, ii) syllabus content and curricular influences, iii) school and/or class environment, iv) teacher attributes, v) the science interest and bias towards or away from science of each respondent and vi) sex. The psychometric problems and the need for substantial, or even vast, sample populations together limited the issues that could be examined to the following three research problems.

Research Problem (B.4)

To investigate the relationship between pupil maturation (as measured by class level) and cognitive preference behaviour.

Research Problem (B.5)

To investigate the relationship between science interest and bias towards science and science studies and cognitive preference behaviour.

Attempts were made to examine these relationships in so far as they exist by appropriate correlational analysis.

The final problem which was attended to in this section was to examine whether there is any difference in cognitive behaviour between the sexes (Problem (B.6)). The particular intention was to repeat the work by Tamir on older students (Tamir 1976) to see whether the differences in cognitive preference behaviour between the sexes is also apparent with pre 'O' level students. At the same time an investigation was undertaken to see whether there is any difference in cross relationships between cognitive preference and academic achievement for the two sexes.

Part (C) : Stability and Style

(Research Problem (C.1))

The interest in this section of the research stems from the belief that genuine style characteristics may be associated with cognitive preferences. The most important evidence which could be gained in support of this concerns the stability of cognitive preference ratings for an individual since that would point to a consistent deep-rooted response pattern. Stability in this context has two dimensions: (a) stability across subject boundaries and discipline divisions. This has been explored and, by and large, been substantiated although the claim by Tamir (1967) to there being in existence a distinct subject-specific element to cognitive preference ratings must inevitably weaken the usefulness of stability in this respect (i.e. the across-subject stability) as providing firm evidence in support of the styles notion. (b) the more important line of enquiry concerns stability with time since this has been used by workers in cognitive styles areas as major evidence.

Two propositions are worthy of consideration:

(i) retention of *absolute* position in a styles network (i.e. no shift or adjustment is occurring with the passage of time)

(ii) retention of a relative position (i.e. relative to *others* in a styles network.)

Previous findings by workers in cognitive styles seem to suggest that (i) is hardly ever found. There is indeed some shift, but (importantly) an individual's position in the distribution pattern obtained for a largeish population has been found repeatedly to be and to remain fairly constant. It is this that tends to be taken as evidence in favour of styles and their existence. In the present situation, this type of stability will be looked for, i.e. the stability of an individual's position in a relative sense, although the absolute position will also be explored.

The notion of stability in cognitive preference orientation differs from the stability which is normally associated with styles in that four, usually ipsative, measures are involved. It is therefore necessary that the stability of the four measurements together is examined as a single relationship (hereafter labelled relationship stability). It is also necessary that the examination of relationship stability is preceded by evidence to show that the respondent's behaviour is stable during the administration of each test instrument and that each of the four mode preferences from which the relationship is built is separately stable (i.e. that there is an acceptable measure of mode stability). An extension of the study of stability takes the cognitive preference relationship and

and examines it in the light of the structure of cognitive preference behaviour.

Section 4.2 : Information Transformation

The term judgement and the relationship between judgement and information transformation has been discussed in Section 1.3 and in Chapter 3. In brief, the principle is that judgement is a two stage process in which the learner, as the "judge", has first to create hypotheses and then to assess their reasonableness. Subsequent 'adjustment' of the learner's cognitive structure as it exists prior to the information transformation permits the new information to be embedded in new or modified cognitive structure. This part of the study was designed to examine each of these stages separately and, also, together. The assessment of reasonableness is the principal objective of the Judgement Exercises. Each of the items in Part I (the theoretical part) provides the respondent with information about a chemical situation, a problem relating to that situation and some statements. The respondent is required to create one or more hypotheses which relate the information to the problem and then classify the statements into categories of reasonableness in the light of that hypothesis. (If the respondent has more than one acceptable hypothesis he must assess reasonableness with each and arrive at a category that he deems most likely). The items in Part II (the practical part) are similar except in that an experiment is proposed and its outcome, if positive, has to be assessed for reasonableness ; an experiment is reasonanable if it explores a parameter that is both relevant and useful in resolving the problem.

The quality of the hypotheses is examined by Part III of the Judgement Exercise. The items are similar to those which are used previously but only the information and the problem are given. The student has to devise experiments which he thinks will give relevant and useful outcomes. The quality of the experiment is then assessed by an adult chemist.

It will be recalled from Chapter 3 (Section 3.4) that the categories were chosen so as to embrace both scope and quality. The classifications which offer the parameters of scope are (a) 'sensible' in that the statement is in harmony with the whole body of science and (b) 'relevant' in that the statement has bearing on that part of the body of science which is germane to the problem situation. Quality is a notion that is specific to the actual problem; if the statement makes full use of all the information that is given it is classified as being 'adequate', if it requires no more information than might reasonably be expected to be present in the cognitive structure of the respondent it is classified as being 'justified'. Copies of the Judgement Exercises are given in Appendix D, E and G.

It is therefore deemed to be axiomatic for the purpose of this research into judgement behaviour (a) that the act of judgement is initiated by an introductory statement of genuine chemical fact and a question which relates to that fact, (b) that the respondent creates a hypothesis (or more than one) and (c) that he evaluates the statements which are incorporated in each item (using a classification procedure with multiple choice format). The hypothesis that is employed by the respondent is not revealed but the extent to which it accords with accepted scientific fact is revealed by the responses which are given. This procedure was adopted by King (1972) and the test instruments used in this study were developed in the light of that research.

The research problems which arise in the new study are of three broad types. The first concerns the conceptualisation of what constitutes judgement in the chemistry context, and how such judgement might be measured. Essentially this concerns the development of suitable 'instruments' and it is discussed in Section 3.3. The second type of problem concerns the evaluation of the instrument for the measurement of judgement ability, in terms of its validity and reliability. The third type is the relationship, if any, between judgement ability and general chemical competence and achievement, and curricular influences upon the student. Research problems (D.1), (D.2), and (D.3) relate to the second problem type and Research Problems (E.1) - (E.4) to the third.

Research Problem (D.1)

Reliable measurement of judgement in the manner prescribed requires that the instrument should have 'face' validity and that the individual statements should respond well to item analysis.

The 'face' validity of the instrument was examined by chemistry teachers.

Research Problem (D.2)

It was considered reasonable to ask respondents to write their responses to the test ; in a supplementary test respondents were asked to state the experiment(s) that they would perform in a number of further situations similar to those encountered in the test in which problems of similar type had been given. It was theorised that respondents who were consistently able to suggest experiments which were considered by the investigator to be relevant to the problem, had consistently hypothesised correctly in the supplementary test and it was anticipated that they would score better on the main test instrument.

Research Problem (D.3)

Instruction in scientific method is not a specified requirement of the syllabus of most examination boards. While it is likely that most teachers of science dwell upon this topic at some time during the O level course it is unlikely that it would be in the forefront of the minds of the respondents. It was deemed to be important to establish whether or not brief training in the rôle of the hypothesis in scientific method would improve the performance on a test of judgement. A sample population was divided at random into groups which received different training and all were then required to respond to the instrument.

Research Problem (D.4)

It was postulated that the facility of an item might be related to the number of 'cause-and-effect' links between the observations in the opening statement and the parameter for investigation proposed in the ensuing statement.

It will be evident from the discussion on the curricular differences between schools (see under Research Problem B.2 above) that similar differences might be relevant in the study of judgement ability and general chemical competence (the third problem mentioned above).

Research Problem (E.1)

It was considered to be important that the scores in a test of judgement ability should be related to the scores in other and more global measures of achievement in science, i.e. O Level grade and other scores in chemical tests.

It may be argued or at least thought possible that judgement ability should in some form be linked to what may be described broadly as "intellectual maturity". Since the subjects involved in this study as a whole were students in the critical phase of intellectual development during which, according to Piaget, formal reasoning ability develops, it was thought advisable to explore whether students' developmental levels do in any significant way influence or correlate with their judgement behaviour. For this reason a suitable test of Piagetian tasks was administered as part of the investigation. It should be pointed out that this aspect of the study was decided upon for exploratory reasons and not in pursuit of a definite hypothesis perceived in an *a priori* way.

Research Problem (E.2)

To explore the relationship, if any, between developement level and judgement behaviour.

'Intellectual maturity' cannot be seen simply in terms of some Piagetian or other developmental process, but must also encompass broad intellectual skills, e.g. those customarily measured in terms of I.Q. or similar tests. It was thought desirable therefore to examine the extent to which judgement ability relates, in a correlational sense, to an accepted measure of intellectual skills which, at least in part, reflects on science orientation. The particular measure adopted for the study, was the Bristol Achievement Test, Study Skills, Level 5. Details of this test are given in Section 4.5, Part E and a copy will be found in Appendix F.

Research Problem (E.3)

To explore the relationship, if any, between intellectual awareness for science (as measured by the Bristol Achievement Test) and judgement behaviour.

The Bristol Achievement Test is a standardised test and does not cover chemical situations as such. There is thus no information to be gained from the comparison of the Bristol Test scores and judgement scores about the effect which chemical achievement has on judgemental competence. King's (1972) study indicated that such effect, if it exists, is of limited magnitude only, but this result was considered by him to require confirmation. In consequence this facet was re-examined in the context of the present study, by comparing judgement behaviour of students with their chemistry achievement grades obtained in O Level examinations as well as in end-of-year tests in chemistry.

Section 4.3 : Multivariate Analysis

In the previous sections cognitive preference behaviour and judgement ability have been dealt with as unrelated (i.e. separate) variables. It was suggested in Chapter 1 that they are in fact linked through the learning model. In consequence an attempt will be made in Chapter 9 to explore whether any such link is borne out by the data which was obtained in this study. Questions such as whether judgement abilities can be related to particular types of cognitive preference behaviour are of interest here. It should be stressed that no particular hypotheses are advanced beforehand because no adequate theoretical link has been established between the processes of judgement and those of cognitive preference choicesto allow such hypotheses to be formulated. Instead the exploration which is to be made here is of a tentative nature, in pursuit of hypothesis-generation, rather than hypothesis-testing. In the course of this work a population of students will be subdivided into groups by scores on all four cognitive preference modes together and the within-group similarities and differences on other variables will be considered, as will the inter-group comparisons.

In Chapter 10 this learning model will be studied in the light of conventional classroom variables in order to explore its authenticity and its relevance and value in the process of learning

Section 4.4 : The Test Samples and Research Programme

All the students in all the samples which were used were actively involved in a course of chemistry at the time of the administration of the test. Students from grammar and independent secondary schools were represented in approximately equal numbers in the 'longitudinal study' but there were more boys in the sample than girls. The students in the 'one-shot study' at third form level were all boys and they were all from the same independent school. All the students were being prepared for 'O' Level examinations; some intended to take Chemistry as a separate subject while others planned to take it in combination with another science subject a 'O' Level. The sample which was studied at third form level in 1974, at fourth form level in 1975 and at fifth level in 1976 numbered 830 at the beginning and 439 at the end. The loss occurred because a number of students, particularly girls, decided to abandon Chemistry in favour of the other sciences two years before 'O' Level and because the administration of the tests at the end of the fifth form level conflicted with the 'O' Level

Table 4.1 : The Test Samples

Year	Test Instruments	no. of schools	no. of classes	no. of respondents	class level
	'Longitudinal' Study				
1974	Cognitive Preference Test-I and Bristol Achievement Test. Study Skills, Level 5	9	40	717	IIIrd
1975	Cognitive Preference Test-II and Judgement Ex-II	9	31	606	IVth
1976	Cognitive Preference Test-III and Judgement Ex-III	9	39	729	Vth

Note: (a) Judgement Ex-I was used and reported by King (1972).

(b) The total number of respondents who completed all tests was 439.

(c) A schedule of tests is given in Appendix

(d) CPI-III and JE-III were also administered to 123 respondents in 1976.

'One-Shot' Study

Tests Administered

1) Cognitive Preference Test-III
2) Judgement Ex-III
3) Differential Aptitude Test - Number*
4) - Space*
5) English Comprehension Test*
6) Piagetian Test (combinatorial)
7) I.Q. (Otis α & β)
8) End-of-year examination of achievement in chemistry

Year : 1976

No. of Classes : 7

Class Level : IIIrd Form

No. of respondents : 161

* Note: the data from the tests marked (*) were gathered and processed but they were not used extensively.

examinations which were going on at the same time.

Tests were administered by class teachers. Before each period of administration of test, each school was visited, the invigilators were instructed on test procedure and whenever possible the students were given an outline of the project in sufficient depth to permit them to appreciate the importance of the work. The care which the students took with their tests was generally most satisfactory; the 8 scripts blighted by post 'O' Level euphoria were jettisoned. Each test or combination of tests was designed to last for the duration of one class period of 40 minutes.

Three different preference tests were administered to the same students in June and July 1974, 1975 and 1976 and the data used for the research which required 'longitudinal' treatment. The same students were tested on the Bristol Achievement Tests, Study Skills, Level 5 in 1974, a test of judgement in 1975 and a second edition of the edition of the same judgement test in 1976. Most of the 'one shot' data were gathered from 1976 IIIrd Form students but a IVth Form sample was also used for a cognitive preference test and a test of judgement in 1976 and another IIIrd Form sample was used for a judgement test in 1977. All the 'O' Level grades which were used were awarded in the Summer of 1976.

Section 4.5 : The Test Instruments

The six different types of tests which were used are grouped as follows: (a) cognitive preference test, (b) judgement test, (c) achievement, (d) science orientation, (e) intellectual development and (f) other tests.

Part (A) : Cognitive Preference Tests

Three instruments were used; the first two are similar and both follow the model used by Kempa and Dubé and used some of their items.

Seven items are common to the two instruments. The third was written after the publication of Brown's paper (1975) and it contained only original items.

CPI-I

The first cognitive preference instrument, called 'Thinking in Chemistry I', contained concepts suitable for IInd and IIIrd form students of chemistry. The pretest version of the test was issued to 55 IInd form students in June 1973 and subsequently amended in the light of the pretest findings. In its operational format, it was used with 717 IIIrd form students in 9 schools during June and July 1974. It contained 24 items each with 4 options and responses were given by allocating 3, 2, 1 and 0 votes to the options in order of preference. (This scoring procedure is identical to the 4, 3, 2, 1 procedure used by other workers. It leads to smaller mode totals and facilitates the mental arithmetic involved in processing response sheets but it may seem misleading in that it accentuates proportional differences between mean mode scores.) The pretest data is given in Table 4.2.

The pretest was achieved by sending the material to one, two or three students in each of 27 schools; in each of these schools the science teacher was asked to scrutinise the language and content and to say whether each of the items had been covered during routine school work. The items were further examined by three experienced teachers of chemistry who had received instruction in cognitive preference test construction and had critically examined three previous tests. All the items of this instrument were new and the only innovation was an attempt to minimise student anxiety by stressing that each person's way of approaching data is unique and not necessarily inferior or superior to that of others. The anxiety arises because the administration of a cognitive preference instrument places the student in a novel situation. He is accustomed to expressing preferences in Arts subjects but he is not normally invited to do so in a Science subject. A copy of this instrument will be found at Appendix A.

Table 4.2 : The Pretest Data and Operational Test Data of Cognitive Preference Test-I, II and III

Pretest Data: No. of students : 55, Class Level - IInd Mean Age : 12.11 (S.D. 3.7 month)

No. of items : 24 Scoring Procedure : (✓, 0, 0, 0)

Mean Mode Scores :

R	7.5	(S.D. 2.3)
A	5.9	(1.3)
Q	5.1	(1.2)
P	5.4	(0.9)

Comments from the respondents and their teachers were invited; 8 items were altered and 4 were replaced.

CPI-I Thinking in Chemistry - I (Appendix A)

No. of students : 606 Class Level - IIIrd

No. of items : 24 Scoring Procedure : (3, 2, 1, 0)

Mean Mode Scores :

R	38.42	(S.D. 8.09)
A	32.93	(10.64)
Q	33.90	(7.45)
P	38.40	(10.01)

CPI-II Thinking in Chemistry - II (Appendix B)

No. of students : 426 Class Level - IVth

No. of items : 28 Scoring Procedure : (3, 2, 1, 0)

Mean Mode Scores :

R	41.81	(S.D. 11.09)
A	41.13	(9.97)
Q	37.70	(8.66)
P	47.27	(10.85)

Seven items in CPI-II were also used in CPI-I (Nos. 5, 8, 16, 29, 23, 27 and 28). These were the so-called 'link items'.

CPI-III Teaching Notes (Appendix C)

No. of students : 439 Class Level - Vth

No. of items : 28 Scoring Procedure : (✓✓, ✓, o, x)

Mean Mode Scores :

R	13.39	(S.D. 9.07)
A	10.88	(9.84)
Q	9.98	(8.86)
P	11.82	(9.44)

CPI-II

The second cognitive preference instrument, called 'Thinking in Chemistry II', contained 28 items each with 4 options. 17 of the 28 items were written by Kempa and Dubé (1971), 4 were new and 7 were from CPI-I. The intention of these 7 items was to repeat them a third time in CPI-III and then examine the stability of students' cognitive preference behaviour on these items in each of the three years. The need for extensive modification of CPI-III frustrated this plan. CPI-II was administered to 606 students in 1975; 477 of these had also worked CPI-I in the previous year. A copy of the instrument appears as Appendix B.

CPI-III

The papers by Williams (1975) and Brown (1975) stressed the need to minimize the factors such as verbal cueing which interfere with the examination of preference. Williams measured the contribution made by preference to a cognitive preference score and found it to be small in comparison both to method variance and content variance but he concluded that "preferences do contribute sufficient variance to make them worthy of examination" (p.74). Brown (p.53) had raised the possibility that the tests used by Heath and by others could have provided extensive cueing with linguistic flags and that this might have been responsible for the high internal consistency reported by these investigators. In the light of this criticism by Brown a drastic revision was undertaken of the instrument which had already been written, scrutinised and pretested. The new document was written during the Spring of 1976; it was scrutinised and discussed, but was not pretested through lack of time before administration. No 'item-total score' analysis was done and thus some imperfections survived in the final document.

A particular innovation made for the third cognitive preference instrument (CPI-III) concerned the introductory statement of each item. The rôle of the statement is to identify a small part of subject area and as a rule it is interchangeable with the R (= Recall) mode option. It is not clear whether respondents perform a triangular exercise when declaring a preference between every pair of options by relating each to the introductory statement separately, or whether they simply ignore the introductory statement after identifying the subject area which entails a straight comparison between options, either on a pair-wise basis or in some other pattern. If in the preference rating, reference is consistently made to the introductory statement, this might possibly bias the response towards the mode implicit in the introductory statement (for instance, the introductory statement of item 13 of CPI-II has a P mode implicit in it, of 15 has an A mode and of 28 has a Q mode). The decision was made, therefore, to prepare four items on each of six topics, to assign the four items to the same page of the instrument and to give each page and each item brief titles. The reduction of the introductory statement to a title was thought to minimise any cueing or bias in the act of expressing preferences, it also eased the construction of test items by releasing the customary introductory statement for utilization as an option. Analysis of the draft instrument showed that most Q and P options were at a higher level of abstraction though not necessarily more complex than the R and A options. Uniformity seemed to be easy to achieve and small alterations were made to the few items which required modification.

CPI-III was administered to 439 Vth Form students, 123 IVth Form students and 161 IIIrd Form students in June and July, 1976. It was thought that '✓✓, ✓, o, x' scoring procedure might facilitate an easier expression of preferences, and hence this mode of responding was chosen. A copy of CPI-III is given as Appendix C.

As an alternative exercise the four items on each of the first three pages of the CPI were broken into 24 pairs and used with 45 IIIrd Form students for the Pair Comparison Exercise (Section 5.1) and the scoring procedure was altered for 45 IVth Form students to permit them to rate each response on a 100 point scale. The data from the former was used to gauge commitment and construct validity of the four modes and the data from the latter furnished an opportunity to view each item as an independent test unit with normative score, to assess "item-total score" correlation, to see how various scoring procedures and methods of analysis behave after the alteration to ipsative scoring has been superimposed on them, to examine fluctuations in intensity of response preferences and to illuminate the concepts of internal consistency and reliability when applied to cognitive preference measurements.

Part (B) : Judgement Tests

The Exercises of Judgement set out to present genuine judgement situations. They were intended to be genuine in the sense that the respondent had sufficient knowledge of the general situations but no specific knowledge of the problems. They were also intended to be well within the intellectual capacity of target samples of respondents. It will be evident that the severity of these restraints ensures that no document will be ideal for any one student and that the test exercise must represent some compromise for both ability and experience.

A sound scientific judgement was theorised as having four attributes labelled sense, relevancy, adequacy and justification. These are described in Section 3.4. The classification procedure which was used by the respondents was related to these attributes.

It was also thought necessary to accommodate within the set of Judgement Exercises both the theoretical and the practical dimensions which are characteristic of the sciences. For this reason the Judgement Exercises were divided into two parts. In the first part, judgements of each of the 4 attributes were demanded in relation to theoretical situations. In the second part, further judgements were required concerning the practical problems. For the latter purpose judgements of these attributes only were required. The fourth attribute, 'justified', was not used here because it is a necessary condition of the items in this part. (Judgement Exercise - II, which served as the pretest for the ensuing document, was structured slightly differently).

A third part was added to the new edition of the test in order to explore the respondents' ability to devise useful experiments for themselves. It was open-ended and asked only for a written record of experiments which were thought to be helpful and worth further study to elucidate four further practical problem situations. The proposed experiments were scrutinised to see whether they examined even remotely relevant parameters.

The exercises were scrutinised by adult chemists to ensure face validity and the score which was allocated to each respondent was simply the sum of those of his judgements which matched those judgements of the adults.

A test (called JE-II) which purported to measure judgement was prepared in 1973. (JE-I was the test which was used and reported by King (1972).) It contained five theoretical situations and five practical ones. Each situation was described briefly and in each a problem was posed. The theoretical situations were followed by some statements and the practical ones by some experiments. The statements and experiments were intended to stimulate the student into making hypotheses which would unify the problem and the statement/experiment. The test rubric required the statement/experiment to be classified in terms of the four attributes of judgement. The sum of correct classification of responses was taken as a measure of ability of judgement. Response was made in multiple-choice format and no guidance was given on the number of exemplars of each attribute. A copy of the test booklet appears as Appendix D.

The responses of JE-II were subjected to an 'item' analysis. This was done by hand with a top and bottom group of 82 scripts (27% of the total sample). Only 25 statements were found to have both a discrimination coefficient greater than 0.30 and a facility coefficient greater than 0.50, 23 had coefficients greater than 0.10 and 0.25 and the remainder were even less satisfactory. These coefficients would not have been acceptable for the usual normative tests of cognitive ability but in the present case it was thought justifiable to retain items with lowish statistics on the grounds that the test is operating at the highest level of intellectual ability. It was particularly noticeable that respondents had considerable difficulty in discriminating between items keyed as *inadequate* and those keyed as *justified* and also in Part II between *inadequate* practical operations in Part II and those which were keyed as *useful*.

The restriction on the choice of material which was imposed by the need to select situations with some chemical flavour, which were readily comprehensible to students at 'O' Level and yet which were sufficiently novel to ensure that they would not have featured in normal classroom teaching, made the task of item construction rather difficult. Even so further items were added to bolster categories and situations which appeared to have a dearth of suitable items with adequate statistics and the expanded version of JE-II was administered to 23 teachers of secondary school chemistry.* The body of information gleaned from the students and teachers was carefully examined and a new version of the test was produced. It was called "Chemical Judgement" and is abbreviated JE-III.

The instructions to the new edition of the test were amplified. A copy appears as Table 4.3(a) for Part I of the test (relating to theoretical situations) and Table 4.3(b) for Part II (dealing with practical situations). A copy of the test booklet appears as Appendix E. A sample situation is shown with the postulated thoughts of a respondent at Table 4.4(a) and 4.4(b) respectively and an example of a situation from Part II is given at Table 4.5(a) and 4.5(b). The terms which are used in the instructions to respondents are defined as follows:

---

* 11 at the University of E Anglia, 2 at the University of Keele and 10 in schools in Kent.

sensible	:	This means that the statement is in harmony with the whole body of scientific experience in its widest aspects. All statements in the JEs were intended to be sensible.
justified	:	This means that the statement is in harmony with the facts and observations which are described in the situation in the test booklet. It is an essentially local concept.
adequate	:	A statement is adequate if it makes full use of the information which has been given in the description. This is both local and pragmatic. The decision at issue here is the point at which the use may be judged sufficient.
relevant	:	The statement has bearing on the problem; not merely on the situation alone.

There are five situations in Part I each with 8 statements (two for each of the 4 attributes) and there are a further five situations in Part II each with 6 proposed actions (two for each of 3 attributes: as noted before, the attribute 'justified' was omitted). Some simple illustrations were introduced to generate interest and to enhance comprehension.

The third part which was incorporated into JE-III was treated separately from the other two parts. Four new problem situations were presented. The respondent was required to devise experiments which he thought would serve to elucidate the problem; he was told that the experiment should be suitable for an 'E' classification (lead to relevant information beyond that already given). He was invited to devise ten experiments and he was free to distribute the ten between the four situations in any combination. The purpose of this part was two-fold: it was designed to examine the nature of the experiments which were suggested and it was designed to furnish evidence on whether students were able to devise experiments to explore

(to page 81)

Table 4.3 (a) : Instructions to Respondents for Part I of JE-III

THIS IS A TEST OF CHEMICAL JUDGEMENT. It is designed for those who do not necessarily know a great deal of chemistry but have an understanding of the subject. You are given information and you are asked to say whether some further observations and experiments make sense. There are no catches, no mistakes, no false statements. The test is in THREE parts. Work carefully and sufficiently fast to finish the test in the time allowed (40 minutes).

INSTRUCTIONS for PART ONE.

In Part One there are five situations which are typical of the many situations which a chemist encounters. Each situation is described briefly, a question is posed and some statements are made. You are asked to put an 'X' in one (and one only) of the four boxes for that statement.

Put an 'X' in the A Box if the statement has a SENSIBLE bearing on the problem and if it is also JUSTIFIED and HELPFUL in relation to the facts and observations which are given.

Put an 'X' in the B Box if the statement has a SENSIBLE bearing on the problem but it is NOT JUSTIFIED because it makes use of important information about the situation which has not been given in the description.

Put an 'X' in the C Box if the statement has some bearing on the problem but contains information which is INADEQUATE AND UNHELPFUL because it merely restates the information which has been given already or fails to throw any light on the problem.

Put an 'X' in the D Box if the statement, though sensible, is IRRELEVANT to the problem given and so has no bearing on it.

REMEMBER Choose	A if the statement is	JUSTIFIED
	B	NOT JUSTIFIED
	C	INADEQUATE
	D	IRRELEVANT

THINK CAREFULLY about the DIFFERENCES between these choices BEFORE you start. When you have started, think about each statement carefully and quite independently of all the other statements.

Put ONE 'X' ONLY in EACH GROUP OF FOUR BOXES.

Table 4.3 (b) : Instructions to Respondents for Part II of JE-III

In Part Two ....

.... of this test you will meet five further situations. This time experiments are suggested which might be undertaken to shed light on the problem which is posed.

Put an 'X' in the E column if the experiment which is suggested will lead to information BEYOND that already given, AND this information is RELEVANT to the problem stated.

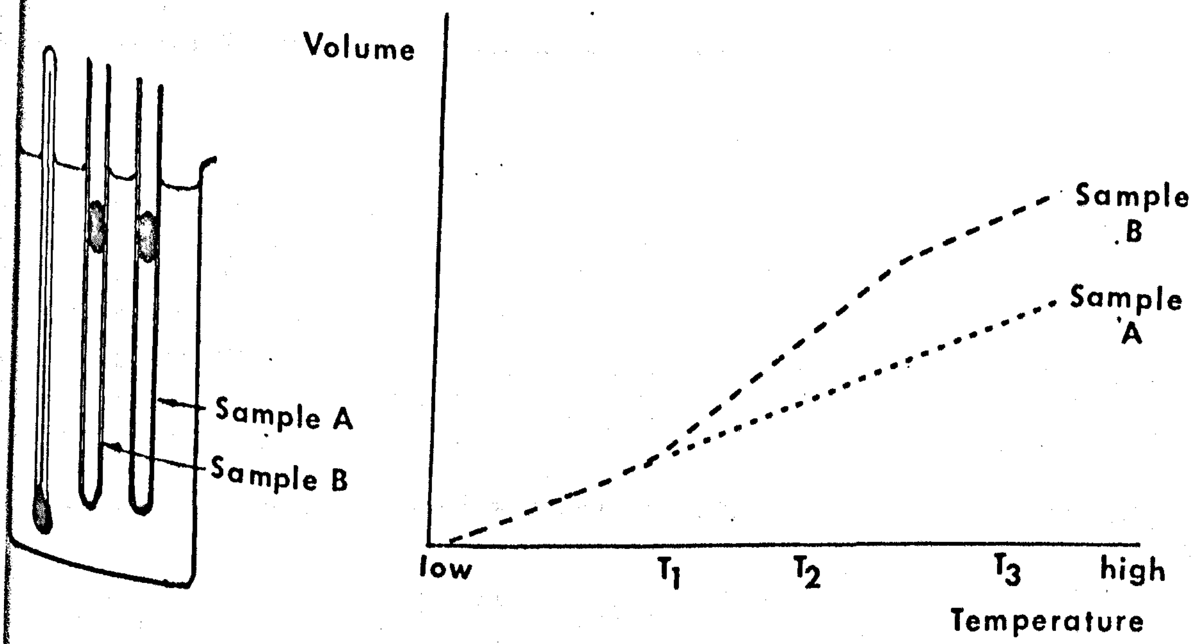
Put an 'X' in the F column if the experiment which is suggested will lead to information BEYOND that already given, BUT this information is IRRELEVANT to the problem stated.

Put an 'X' in the G column if the experiment which is suggested will merely give information which is ALREADY AVAILABLE.

Put ONE 'X' ONLY in EACH GROUP of THREE boxes.

Do not assume that information from other experiments is available; consider each experiment according to its own merit.

Table 4.4 (a) : Situation 4 from JE-III



The diagram shows two tubes containing samples of gas trapped by a drop of mercury. The tubes are held in an oil bath which can be heated. The volume of the two gases at different temperatures can be measured. The graph shows the volumes of the samples over a range of temperatures.

The two gases behave differently. Why?

25. Sample B may be turned into a liquid by cooling.
26. At temperature T_3, there are more molecules of gas in Sample B than there are in Sample A.
27. Sample B undergoes a reaction when heated.
28. The graph shows that at temperatures higher than T_1, the two gases behave differently.
29. Sample B was dinitrogen tetroxide (N_2O_4).
30. Both samples are gaseous at temperatures below T_2.
31. The gas hydrogen iodide would behave like Sample A at temperatures near 100°C and like Sample B at temperatures near 500°C.
32. The experiment was conducted at constant pressure.

Table 4.4 (b) : Thoughts postulated for a respondent while he is engaged on Situation 4 of JE-III

	Postulated thoughts of student (possible hypotheses)	Decision and Classification
The two gases behave differently Why?	Interpret the graph/decrease in Vol B/T_2 could be b.pt. of B.	Accept - D data says tubes filled with gas.
	more molecules/greater volume/ change in no. of molecules.	Accept - A
	change in no. of molecules	Accept - A
The two gases behave differently Why?	kink in curve/more molecules of B/behave differently	Accept - C Does not explain cause of difference.
	N_2O_4 must react when heated	Accept - B Uses extraneous information
	Interpret the graph/both are gases.	Accept - C Does not explain cause of difference.
	ammonia not react at lower temp., does at higher	Accept - B Uses extraneous information.
	interpret graph/a straight line therefore no pressure change	Accept - D irrelevant : same for both gases.

Note: (i) Postulated hypotheses/inferences underlined.

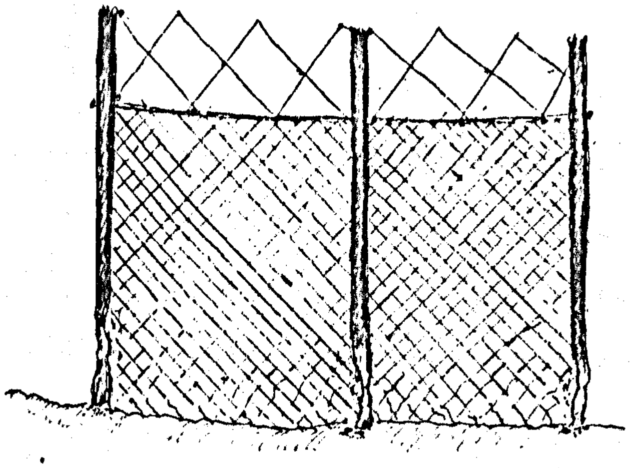
(ii) The following must be in the cognitive structure of the student:

- ability to interpret graphs
- comprehension of phase change/temperature relationship
- kinetic theory of matter
- Arogadro's hypothesis
- Charles' Law

(iii) The following must not be in the cognitive structure of the student:

- knowledge of thermal dissociation of gases.

Table 4.5 (a) : Situation Z from JE-III



The painted iron posts which are used to support the netting round playgrounds and tennis courts rust far more rapidly near the ground (i.e. within 25 cm. of the ground) than they rust elsewhere.

The rusting is uneven. Why? Assume that one or more poles are treated in each of the following ways and they will be carefully examined for rust at suitable intervals of time thereafter.

25. Provide a shield to prevent rain water from splashing up from the ground on to the post.

26. Raise a pole and replant it with the buried section encased in polythene.

27. Replace an iron pole with another which is identical in all respects except in that it is made of aluminium.

28. Raise a pole and replant it the other way up.

29. Replant the poles with rust resistant paint.

30. Shorten a pole by cutting it through at 25 cm. above ground level and reset it in the ground.

Table 4.5 (b) : Thoughts postulated for a respondent while he is engaged on Situation Z of JE-III

Item	Postulated thoughts of student	Decision and Classification
25	rusting is due to chemicals leached from the ground.	accept - E
26	rusting is due to chemical interaction with soil water and/or solutes/polythene will serve as barrier.	accept - E
27. The rusting is uneven Why?	aluminium may or may not behave similarly.	accept - F new info. but irrelevant
28	rusting will recur	accept - G info. already available
29	rusting may be delayed.	accept - F new info. but irrelevant
30	rusting will recur	accept - G info. already available.

Note : the galvanic nature of rusting must not be in the cognitive structure of the student.

alternative parameters (It will be recalled that Mealings (1961, 1963) found that this was an area of difficulty for students). A record was made of all the E-Category (E = beyond and relevant) experiments which were submitted, each student was awarded a point for each of the experiments which he devised if a relevant parameter was implicit in it and his total score was recorded. If an experiment involved more than one apparently relevant variable two points were awarded; a mere explanation (which embraced no relevant parameter) did not score even if it was the correct explanation.

The final exercise (designated JE-IV) comprised seven problem situations (drawn from all three parts of JE-III). Each of the seven problems had proposed experiments and the student was required to say in which of the experiments he thought that the effort necessary for the completion of the experiment was justified on the grounds that the experiment explored a parameter which was both useful and relevant. Extensive use was made of experiments which had previously been proffered by other students in their responses to Part III of JE-III. The rationale for the test was that the student in responding to it would have first to generate one or more hypotheses and then review the experiment in their light. The instructions merely asked that the student should indicate whether the experiment was worth doing. He was not asked to give a reason nor to provide any classification. 34 of the 70 items were keyed as category E (= beyond and relevant) items and the remainder were equally divided between category F (= beyond and irrelevant) and G (= information already available). A copy of the booklet and of the response key appears at Appendix G.

Validity of Judgement Exercises

The construct validity of these must be a matter of definition. It was assessed by 23 judges who found the items representative of what was intended. Concurrent validity cannot be established because there is no recognised other measure with which to compare the judgements as they are defined here and assessed by the judgement exercises.

Table 4.6 : Analysis of Judgement Data (JE-III; N = 439/Vth/1976)

(a) Means and Standard Deviations

	Total	Sub tests						
		Theoretical Part				Practical Part		
		A	B	C	D	E	F	G
Mean	32.70	4.00	4.61	2.71	4.12	6.03	5.60	5.51
S.D.	9.29	1.78	2.07	1.72	1.82	1.98	2.03	2.39

Note: Maximum possible sub-test score : 10; maximum possible total score : 70

(b) Correlation Coefficients between 'total' score and sub test score ($x\ 10^2$) ('total' score = sum of other sub test scores)

A	B	C	D	E	F	G
47	49	30	48	48	59	57

(c) Correlation Coefficients between sub test scores ($x\ 10^2$) and factor analysis

B	23					
C	00	06				
D	16	15	10			
E	40	20	03	17		
F	23	22	14	24	39	
G	27	25	16	29	33	42
	A	B	C	D	E	F

($r < .123$, $p < 0.01$)

Rotated Factor Analysis (factor loadings $x\ 10^2$)

Factor	I	II	III	IV
A	70	-16	-01	23
B	15	05	09	97
C	01	94	01	03
D	08	-02	95	06
E	83	-01	04	00
F	61	30	29	05
G	51	29	42	16
	(loadings > .163, p < 0.01)			
Eigenvalues	2.371	1.071	0.852	0.843
Cumul. %age	33.9	49.2	61.3	73.4

Inspection of Table 4.6(a) shows that Sub tests E, F and G have a higher facility (larger mean values) than sub tests A, B, C and D, i.e. that the practical situations appeared to be easier than the theoretical ones. This is due, in part, to the fact that the practical statements required classification into only three categories while the theoretical statements required classification into four. There is no obvious difference in standard of the situations in the two parts and the conclusion that respondents are more confident in practical situations is justified. The mean scores and correlation coefficients may be regarded as satisfactory but the low mean score in Sub test C is interesting. This category required an assessment of whether the statement made *adequate* use of the information which was supplied. The concept of adequacy was evidently too demanding for O Level students (the mean score is scarcely greater than the random value of 2.5) and the lower correlation of this sub test in parts (b) and (c) of the table suggests that it did not measure 'judgement' as defined previously. It could be argued that at the level of the sample under inspection it was no more than an exercise in semantics.

No tests of reliability were undertaken. The very nature of the exercise precluded 'test-retest' and 'split-half' tests would have been equally inappropriate because no 'parallel' halves can be defined. The correlation coefficients between total score and sub-test scores were determined to provide some evidence of internal consistency as was the rotated factor analysis. The latter reveals that sub-tests E, F and G (the practical sub tests) load together with Sub test A (= sensible and justified) and that Sub tests B, C and D are three independent measures. The distinction between the 'global' sub tests (A and E, D and F) and 'quality'-of-data sub tests (B, C and G) which was drawn in Chapter 3 is not obvious in the factor analysis. ('Global' concerns a judgement which invokes the whole body of science while 'quality' is immediate to the usage that is being made of the data within the item in question). D is not loaded with the 'global' sub tests but it does load with G (Factor III); B and C load separately and even when fewer factors than four were used they appeared to share no properties in common. Sub test C proved to be much the most difficult for the respondents (and for the test constructors) and later it will be shown (Chapter 8, Table 8.7) to have greater facility for students with less academic achievement.

Part (C) : Achievement Tests

As was mentioned in Section 2.4 Part C some relationship has previously been reported to exist between cognitive preference behaviour and chemical ability/achievement. In view of this appropriate measures of such ability were required in relation to the three age levels involved in the study as a whole, i.e. Vth Form, IVth Form and IIIrd Form levels.

(i) Vth Form Level

Rather than develop an independent achievement test for this level it was decided to accept O Level grades in Chemistry as an adequate and reliable measure of a pupil's achievement in chemistry. This was not merely a matter of convenience: it also eliminated the considerable problem that would have arisen if a single achievement test covering the content of different examination syllabuses had had to be designed.

The basic assumption that has to be made in accepting O Level grades from different examining boards as measures of chemical attainment is that a B-grade awarded by one board is equivalent to the same grade awarded by another board. It is generally accepted that this assumption is justified in terms of the percentile ranges of candidates obtaining the different grades.

ii) IVth Form Level

The end-of-year examination scores were available but no use was made of this information.

iii) IIIrd Form Level

The test instrument which appears at Appendix H allocates seventy per cent of the available score to Knowledge and thirty per cent to Understanding.

Table 4.7 : A Breakdown of the O Level grades awarded to the students in the 'longitudinal' study. (N = 439/Vth/1976)

	Grade						Total
	A	B	C	D	E	U	
Non Nuffield Chemistry	62	99	73	12	9	6	261
Chemistry in Combination (see note)	1	15	27	11	5	2	61
Nuffield Chemistry	7	15	42	16	10	27	117
Total	70	129	142	39	24	35	439

Note: Chemistry in Combination embraced Physics-with-Chemistry, Chemistry-with-Biology and Combined Science.

Table 4.8 : Test of Achievement in Chemistry at IIIrd Form Level (N = 161/IIIrd/1976)

	Mean Score	S.D.	Max. Score
Ability 1 (Knowledge)	27.84	10.64	70
2 (Understanding)	9.40	4.62	30
Total	30.19	5.94	100

No. of structured items : = 25

No. of longer questions : = 10

See Appendix H.

The test was scrutinised by the six teachers of the classes involved to ensure that the classification of content to Knowledge and Understanding was correct and that the test was of reasonable standard for their students. Both the multiple choice and structured items were pretested during the previous years and subsequently modified.

Part (D) : Tests of Intellectual Abilities

The homogeneity of the sample populations which were used in this study caused concern. The 'longitudinal' research was done with respondents from nine different schools and the 'one shot' population, even though it was all from one school, included a wide range of young students who had only fairly recently moved schools. It was deemed necessary, therefore, to adopt instruments for standardising the behaviour of the students in the samples and these had, of necessity, to be as close to the area in question as possible. Normally IQ tests are employed for this but these are generally inadequately differentiated in relation to different intellectual ability traits. The Differential Aptitude Tests are exceptional in that they do distinguish between such traits (e.g. numerical ability, mechanical reasoning, space relations, spelling and grammar, abstract reasoning etc.). Of these, three were thought to be particularly relevant to this study as giving information which might broadly relate to science namely numerical ability, mechanical reasoning and space relations but it was decided not to adopt the mechanical reasoning test because it is strongly sex dependent.

Part (E) : Bristol Achievement Test, Study Skills, Level 5

The development of achievement tests which have validity at the IIIrd Form Level poses major problems, in view of the enormous (and now readily quantifiable)curricular differences that exist in different schools especially during the first two or three years. There is relatively little content that may be relied upon as 'common' to all schools; thus, the notion of a content-free test (at least in relation to direct curricular content) must be considered.

Table 4.9 : Tests of Intellectual Abilities (N = 161/IIIrd/1976)

	Score Range	Mean	Stan. Dev.	
Differential Aptitude Tests:				
Numerical Ability (Form M)	0 - 40	22.32	5.62	Note (i)
Space Relations (Form M)	0 - 60	34.98	8.67	Note (ii)

Note (i) : Standard score for Grade 8 students is 16.3 and for Grade 9 is 18.7. Third Form students are comparable in age to Grade 8 students but the sample is from the upper range of the total population of that age.

(ii): Standard Score for Grade 8 : 25.6 and Grade 9 : 25.7.

Table 4.10 : Analysis of Bristol Achievement Test, Study Skills, Level 5 with IIIrd Form Sample (N = 326/III/1974)

Part		No.of Items	Mean	S.D.	Correlation Coefficients ($x 10^2$)			
I	Properties	15	8.49	2.17				
II	Structures	15	11.39	1.96	35			
III	Processes	10	7.25	1.63	42	31		
IV	Explanations	10	6.65	1.59	32	24	35	
V	Interpretations	10*	7.10	1.84	28	20	38	30
		65	40.97	6.18				

$r > 0.14, p < 0.01$

* Note: Mean chronological age of sample 13.8 months

In the present case, the Bristol Achievement Test, Study Skills, Level 5 was considered to be a suitable instrument whereby general scientific achievement (and aptitude) could be adequately measured for the group of IIIrd formers. The test which is intended for use with 12 to 13.11 year olds attempts to test and measure in 5 sub-tests the skills and competences that are required for the study of science without prescribing the form and contents of the curriculum. Part I of the test strives to expose properties of materials and situations that enable them to be characterised and assessed. Part II examines structures in which there is interdependence of parts. Part III enables respondents to display their ability to deal with processes and sequences and so carry out extrapolations and interpolations. Part IV focusses on explanations and on the level of concept-making through which explanation is generalised and Part 5 is concerned with interpretation of information that is symbolic or graphical. A copy of the test is appended (Appendix F).

The administrative manual specifies that the students should be allowed 50 minutes. One school period of 40 minutes was allotted to the test and the students who had failed to finish were encouraged to use the next ten minutes; few incomplete booklets were submitted. The test was administered to 864 students. Mean part scores, standard deviations, and inter-part correlation coefficients are given at Table 4.10.

The mean age of the sample of students was rather greater than the age intended by the test designers. Therefore the test statistics provided by the designers of the test were thought not to be strictly applicable. In consequence the multiple choice items were subjected to item analysis (top and bottom groups = 164). A satisfactory dispersion of facilities among the 40 items analysed was observed but no item had a discrimination greater than 0.50 and only 15 had a discrimination greater than 0.30. In only 9 of the items were all the distractors acceptable to at least 2% of the sample; each of the remaining 56 items had one, or more than one, unsatisfactory distractor.

A modified form of the Bristol Achievement Test was prepared from 30 multiple choice items selected on the evidence of the item analysis. The significance of cartography in the original was diminished but in other respects the balance was preserved. The new test was administered as a power test to the third form sample (N = 161) and the time allotted to it was 20 minutes. The mean score for this test was 16.21 (S.D. : 3.52; max. score 30).

Part (F) : Science Orientation

It has previously been argued that cognitive preferences reveal aspects of cognitive structure and that they must therefore be held to belong to the cognitive part of human behaviour. Further, it is not inconceivable that at least one of the scales of cognitive behaviour, the R-Q scale might reveal a 'leaning towards science' in that the question behaviour at one extreme may be thought of as being more scientific than mere acceptance behaviour at the other extreme. If this is accepted it is not unreasonable to enquire whether there is a relationship between science and scientific activities and the respondent's position on the R-Q axis. For this reason a brief questionnaire was compiled to assess orientation to seven scientific activities (namely science classes, scientific pursuits and hobbies, science reading in books, magazines and journals, work in the laboratory, the use of scientific tools, science programmes on television and scientific careers). Each of these was covered by one question which was answered on a 5 point Likert scale thus giving a score range of +14 to -14. This was not a major measure and no undue importance should be attached to it. It was only validated by inspection and no estimate of reliability was taken. The test is appended (Appendix J).

Another measure with intent to measure orientation was also made. This endeavoured to assess attitude to chemistry in particualr rather than science in general. Students were asked to list in order their three most preferred subjects for study at A Level and to name in any order their three least preferred. A five point scale of chemistry bias was established using points 1, 2 and 3 if chemistry featured in the most preferred list, point 4 if no mention was made of chemistry and point 5 if chemistry was listed among the least preferred subjects.

Table 4.11 : Chemistry Bias of Vth Form Students (N = 439/Vth/1976)

		Number of respondents	
1.	Chemistry most preferred A Level	30	)
2.	second most preferred A Level	72	) 173
3.	third " " " "	71	)
4.	unplaced	118	
5.	among the three most disliked A Levels	148	
		439	

Part (G) : Test of Intellectual Development

It will be recalled from Section 4.3 that one of the problems which was to be researched concerned the relationship, if any, between intellectual development and judgement (Research Problem (F.1)). The intellectual development was assessed by the Piagetian combinational test and the procedure which was used was the one which Rowell and Hoffman (1975) found to be satisfactory for use with a class of students. In this test the students are shown that a brown colour is observed when two solutions from a group of five numbered but otherwise unidentified solutions are mixed. The student is given the five liquids and a worksheet on which he is required to keep a log of his experiments to determine the behaviour of each in relation to the chemical change which produces the brown colour. The five solutions which were provided were: (I) dilute mineral acid, (II) water, (III) hydrogen peroxide, (IV) sodium thiosulphate and (V) potassium iodide.

It was possible for him to make one or more of the following findings:

(a) III and V are essential but slow

(b) I assists

(c) II makes little difference

(d) IV hinders

(e) III and V and I fast

The students were assessed at one of the seven levels as follows:

Level 1. (a) to (e) found.

2. (a) or (e) and one from (b), (c) or (d) omitted.

3. (a) and (b), (a) and (c) or (a) and (d).

4. III essential or V essential and either (b) or (c) or (d).

5. III essential or V essential.

6. obvious plan but no deductions drawn.

7. no plan and no deductions.

It was felt that the restriction imposed on the accuracy of the measurement by the use of only a single test and by the group administration procedure should limit the use made of this data. The administration was satisfactory but the circumstances precluded the use of the prompts and cues which are valued by some authors for improving the precision of the test. The levels which were assigned to students should be treated with caution but the marking procedure allows a reliable distinction between Piaget's Level IIA and IIIA at Level 4 on the scale above.

CHAPTER 5 : COGNITIVE PREFERENCE : MEASUREMENT ASPECTS

Section 5.0 : Introduction

Much of the work in this chapter stems from the fact that it is impossible to validate the instruments which are used to measure cognitive preference orientation by any other comparable instrument. The nature of cognitive preference and the literature which covers the research on it are discussed in Chapter 2. Section 4.1 deals with the division of the new research into three areas for study. The measurement of cognitive preference is the first of these areas and it is treated in this chapter. The work is divided into two parts and it is assigned to Sections 5.1 and 5.2: the two research problems which are concerned with the instrument (Research Problems (A.1a) and (A.1b) are covered in the former section and the aspects which are related to scoring procedures will be found in Section 5.2. Tamir (1975) postulated that cognitive preference measurements comprise three elements. The third of these he called the subject-specific element and it is that part of cognitive preference behaviour which can be specifically assigned to respond to the fragment of syllabus within an item. He suggested that response to different fragments may vary. The opportunity to investigate the differences, if any, between such fragments, was taken and the research is reported in Section 5.3. The evidence for new information on measurement of cognitive preference is summarised in the final section

Section 5.1 : Cognitive Preference Modes and Items (Research Problems (A.1a) and (A.1b)

The nature and development of cognitive preference studies has been outlined in Chapter 1. It will be recalled that Heath (1964) culled the four cognitive preference modes from the literature which described the new courses in science but he gave no interpretation of the nature of these modes of psychological theory. Subsequent research has either revolved round these four modes or three only have been employed and little attention has been paid to the manner in which they are perceived by the respondent. The criteria which respondents employ have not been reported and no attempt has previously been made to explore differences in intensity of preference. The single aim of the evidence which follows is to demonstrate that respondents do perceive the four modes in the manner described by Heath and the four items in each option accurately reflect the mode intended by the designer of the instrument.

The evidence which is described was gathered by administering a pair comparison exercise and followed immediately by an interview. It was postulated that the evidence would furnish information on the extent to which the respondents were actively involved and intellectually committed to the preferences which they were expressing.

The rationale for this piece of work stems from the belief that if a student is presented with an option which embodies R mode information, for instance, and if he first signifies that he prefers it to the other three options and subsequently justifies the preference with a statement which also reflects the R mode, his behaviour supports the thesis that he is exercising a genuine cognitive preference. (It is obviously essential that he should select the option before he is told that he will be required to justify the selection since the knowledge might modify his selection). Consistent behaviour towards this mode and towards the other three modes also confers validity on the existence of the four cognitive preferences; if the behaviour is further repeated across a population it has the hallmarks of a cognitive style.

If the student gives as reason for selecting the R mode option a statement containing extraneous information about his feeling with no R mode thought implicit in it (or worse, with one of the other three modes implicit in it), his decision to choose the R mode option is suspect. If the student is unable to choose consistently it must be either that he is unconscious of the perceptual process, and so is unable to justify his choice, or that the perceptual process has no existence. It is also possible that analysis of the extraneous material might reveal modes of attending to information beyond the four described by Heath.

The simplicity and the well-established credentials of the pair-comparison procedure (see Guildford (1954)) commended its use for this purpose. It is recognised that an instrument which is operating in the affective domain might prove to be unreliable when the constraint of justification is imposed on the respondent. Possible interference was forestalled by gathering the information in two parts. It has previously been shown by King (1971) that scores which were obtained in a cognitive preference exercise with the traditional multiple choice format of introductory statement and four options correlated perfectly with scores obtained from an analysis which treated the rank orderings as decisions between six pairs of options (i.e. (a) with (b), with (c) and with (d), (b) with (c) and (d) and (c) with (d). Pair comparison was thereby shown to be a valid procedure for exploring cognitive preference hence its adoption for Part I of this experiment. Part II involved a private interview with each respondent.

Procedure

Part I : A pair comparison instrument was prepared in which 48 stimulus statements were arranged in 24 pairs. Each statement concerned a small fragment of the O Level syllabus and each was familiar to all respondents. The respondents were told that all the information was correct science and they were asked to pay special attention to the selection of the preferred statement in each pair because the information would be used for subsequent refinement and improvement of syllabus. The statements were structured to pair each mode with each of the other three modes twice and the pairs were distributed randomly within the test. The statements which were used were the options of the four items on pages I, II and III of CPI-III (Appendix C). The maximum score for each student for each mode was 12 and the scores were ipsative. The two classes of students were selected because together they embraced students with the full range of abilities which one might expect to find in any normal group of 'O' Level candidates and because they had the same chemical background and had been instructed by the same teacher.

Part II : When all the respondents had completed the pair comparison exercise, they were asked to consider their decisions again, and without altering any, to select ten or so and explain why the favoured statement was preferred. This was done initially in note form and then more fully in discussion with the tester in private to ensure correctness of respondent intention. Great care was taken to ensure absence of cueing and of any indication of 'tester preferred' responses.

The data which were gathered in Part I and analysed in Table 5.1 show that the R mode was strongly preferred to the other three modes and particularly to the P mode. It should be noted that the respondents were IVth Form students. No attempt will be made to rationalise the difference but it will be discussed further in Chapter 6 in relation to maturation (Research Problem (B.4)). The standard deviations are predictably close because the scoring was ipsative; the correlation coefficients and the factor loadings are familiar in that they closely resemble, in both size and direction, the figures obtained by Kempa and Dubé (1973), by King (1972) and by others. Nothing new arises from them but they do confirm that gathering cognitive preference data by pair comparison is acceptable and they lend credibility to the data which were gathered in Part II (the interviews).

The 45 respondents each selected ten or so responses and explained the reason why one option in each pair was preferred to the other option. The reasons which they gave were then assigned to categories; the principal division which was used separated responses which confirmed the mode implicit in the selected option from those responses in which the mode in the option was contradicted by the response. Example of four confirmatory responses are

Table 5.1 : Analysis of Cognitive Preference Data gathered by Pair Comparison
(N = 45/IV/1976)

Means, Standard Deviations and Correlation Coefficients (x 10^3)

R	8.13	2.11			
A	6.38	1.74	-104		
Q	5.16	1.87	-572	-100	
P	4.31	2.19	-374	-590	-229
			R	A	Q

($p < 0.01 = .372$)

Rotated Factor Analysis (loadings x 10^2)

	Factor I	II
R	88	-22
A	-04	-75
Q	-82	-20
P	-09	96
Eigenvalue:	3.84	3.56
Cumulative %age of variance:	42.0	81.3

Table 5.2 : Classification of responses to Pair Comparison and Interviews
(N = 45/IV/1976)

	Top (N = 25)	Bottom (N = 20)	Total (N = 45)
No. and %age of confirmatory responses	135 (57%)	114 (61%)	249 (59%)
No. and %age of contradictory responses	21 (8%)	5 (3%)	26 (06%)
No. and %age of unclassifiable responses	82 (34%)	68 (36%)	150 (35%)
	238	187	425

given below with the pairs to which they refer:

Pair No. 8

(i) The bleaching action of chlorine on wet litmus serves as a good test for the gas.

(ii) Chlorine bleaches by removing electrons, sulphur dioxide bleaches by adding electrons.

Pair No. 24

(i) Coloured chemicals are oxidised by chlorine because their structure is changed to different (and colourless) chemicals by altering their electron structure.

(ii) Dilute solutions of chlorine in water or alkali bleach discoloured fabrics and ink stains.

(These are the A, Q, P and R mode options respectively of Item 3 from JE-III).

	Student Statements	Classification
Sample responses to Pair No. 8	"gives use of bleaching effect not what bleach it is".	A preferred to Q
	"I would like to have this explained to me. It is interesting because bleaching can be done by oxidising and reduction".	Q preferred to A
Sample responses to Pair No. 24	"description of work done more interesting than more complicated one.	R preferred to P
	"to find why chlorine bleaches chemicals"	P preferred to R

This piece of work shows that in about 60% of responses respondents were able to select one option of a pair and give, as justification for the choice, a reason which embodied the mode which was implicit in the option. The respondents as a whole had a bias towards R in Part I but the top group selected options for all four modes more nearly equally when asked for reasons in Part II, while the bias towards R was intensified by the bottom group. Careful scrutiny of the responses which were given show that the four modes do have real identity in the perception of the respondents and that in making preferences between the modes they are making conscious decisions without being aware of the structure of the test instrument. They are not responding merely to the cues in the options as Brown suggested. Unexpected features which emerged include the stress which was placed on the laboratory and on practical skills by both groups and by the bottom group in particular. The importance of precision of language and the generalisability of information were featured by the top group only but little stress was placed on teaching style, on relevance, on importance and on familiarity with the material by either group.

Section 5.2 : The Evaluation of Cognitive Preference Data

Cognitive preference data may be collected for two broad purposes; to illustrate the inherent properties of the response behaviour itself and to relate this behaviour to other variables and examines implications. Underlying both of these avenues of research lies the ultimate objective of demonstrating cognitive preference behaviour as a cognitive style. Unfortunately this work is hampered by the fact that cognitive preference data is ipsative. Examination of ipsative data with procedures which are insensitive to data type (e.g. means, standard deviations and analysis of variance) severely restricts the scope of the analysis and so resort has been made to normative analysis procedures followed by cautious interpretation. The work which is described in this section was done in order to estimate the confidence which can be placed in this practice.

Normative cognitive preference data was gathered by requiring the respondents to use linear scales instead of the more usual voting or 'most preferred option' procedures. They were told to use the whole of the scale and they were asked not to give equal preference to any two options in the same item. This data was then examined both in its raw state and after it had been made ipsative by rank ordering and by subtraction of the mean preference score for each item from each of the four option scores of that item. It was postulated that the greater the similarity of the properties of these two types of data, the greater is the confidence which accrues to the practice of normative analysis of ipsative cognitive preference data. If, in addition, other procedures which are tolerant of ipsative data confirm the practice and none are found to refute it, confidence in the practice is further enhanced.

A sample of 45 students at IVth Form level were asked to respond to Cognitive Preference Instrument - III (CPI-III) on a 100 point linear scale. One such scale was provided on the response sheet for each of the 24 items of the instrument. Indication of intensity of preference for each option was achieved by marking the scale at the appropriate point with a thin line and identifying the line with the number of the option. The ends of the scale were marked 'satisfying' and 'unsatisfying'. The items on pages 3 and 4 of the instrument (i.e. 8 items in all) were readministered 35 days later; no warning of the retest was given and a period of school holiday accounted for the entire interval between the two administrations.

The position of each line on each scale was measured and recorded as a number between 0 and 99. It was not possible to analyse the 96 x 96 matrix (i.e. 24 items each with 4 options) and recourse was made to analysis of the 16 options on each of the 6 pages of the instrument separately and to <u>scale analysis</u>* on the 24 option scores of the four modes separately.

This method of scoring has not been in general use because it is laborious and because of misgivings over differences in intensity between students and between items by individual students.

Scrutiny of the data was conducted in order to see whether it had defects. Four qualities were examined, namely 1) intensity of preference for each of the four modes, 2) item intensity, 3) option-total correlation, 4) reliability.

---

* S.P.S.S. Reliability Analysis for Scales : Version 6.

1. <u>Intensity of Preference for each of the four Modes</u>

It was postulated that the sum of the 24 options for each of the four modes should be the same. In fact it was found that the mean R preference mode score was significantly greater than the mean Q preference mode score and the mean P preference mode score and that all other differences were not significant. The relevant figures are given in Table 5.3. It is evident that the stimuli of the four preference modes do differ and that the scoring procedure or the instrument or both are defective at $p < 0.01$ in 2 of the six t-tests. It will be shown later (Section 6.2, Part C) that both the ipsative data generated from this data and the ipsative data collected with the same instrument on another sample of IVth year students using the normal voting procedure have similar patterns. Every effort was made during test construction to minimise differences (and when expressed as percentages (Line (c), Table 5.3) they are seen to be small) but they are a feature of cognitive preference measurements. The differences are therefore a feature of the instrument and not of the scoring procedure. They do not invalidate the findings reported later in this chapter and they are also relevant to the study of change in cognitive preference behaviour with age in Chapter 8 for instance.

Table 5.3 : Total and Option Preference Mode Scores and t-tests
(N = 45/IV/1976)

	R Mode	A Mode	Q Mode	P Mode
(a) Mean Total Preference Mode Score	1427.56	1342.27	1231.69	1233.69
(b) Standard Deviation of (a)	309.46	280.62	238.87	254.27
(c) (a) as a %age of grand total	27.26	25.63	23.52	23.56
(d) Mean Option Score	59.48	55.93	51.32	51.40

t-tests on Mean Total Preference Mode Scores:

R - A : 1.35
R - Q : 3.32** A - Q : 1.99
R - P : 3.21** A - P : 1.90 Q - P : 0.04

($p > 0.01$ $t = 2.69$, shown **)

2. Item Intensity

It is inevitable in a forced choice situation that each of the items contribute equally to the final total. It was postulated that in a well balanced instrument with scale scoring it should still be possible to achieve this obviously desirable quality. The data in Table 5.4 shows that the total mean score of each item differ but the standard deviation is only 8.2% of these totals and that 17 of the 24 items are within one standard deviation of the mean. The item on fireworks (page 4, item 3) was the most preferred item and the second that on distillation of alcohol (page 2, item 1). High scores for these items might have been foreseen; the low scores for the item on solubility (page 5, item 1) and sublimation (page 1, item 1)were less easy to rationalise. It is difficult to substantiate the argument that familiarity with item content is a relevant factor; moreover, utility and intellectual strain do not appear to have obvious bearing on item totals either.

Table 5.4 : Total mean scores on the 24 items of CPI-III with normative scoring (N = 45/IV/1976)

Page	1	2	3	4	5	6
Item 1	196.76	246.08	214.20	222.31	182.54	231.00
2	220.98	209.62	225.42	213.45	209.43	244.06
3	226.18	209.94	229.98	254.74	200.35	196.91
4	204.02	191.05	220.04	237.42	215.81	222.93
Page Totals	847.94	856.69	889.64	927.92	808.13	894.90

Mean of the item totals = 217.72 (S.D. = 17.84)

The topics on each of the six pages of CPI-III made fairly equal contribution apart from the solubility topic on page 5 which was noticeably disdained. It is considered that this examination of items has not revealed any gross failing and that all the items could be retained in a new edition of the instrument.

3. Option-Total Correlation

Table 5.5 : Option-Mode Total Score Product Moment Correlation Coefficients (N = 45/IV/1976)

	Correlation Coefficients (x 10^2)																							
Page	1				2				3				4				5				6			
Item	1.	2.	3.	4.	1.	2.	3.	4.	1.	2.	3.	4.	1.	2.	3.	4.	1.	2.	3.	4.	1.	2.	3.	4.
Option R	53	62	61	52	72	60	45	67	74	54	31	62	61	69	13	26	49	37	30	66	72	49	36	56
A	24	14	16	17	28	61	58	40	74	46	72	10	62	74	31	57	14	61	57	48	55	59	32	44
Q	36	12	09	05	52	37	67	24	54	62	62	58	72	30	35	57	04	53	59	52	52	35	41	56
P	30	16	43	35	29	47	47	42	51	54	66	21	37	41	33	33	37	57	49	58	54	32	60	69

($r > 0.288$, $p < 0.05$; $r > 0.372$, $p < 0.01$)

The interest in the correlation coefficients in Table 5.5 lies in the extent to which they generate confidence in the quality of the instruments. Only 16 of the 96 options failed to correlate with significance less than 0.05 and 8 of these were on Page 1 while the figures would be of value to the author of a new edition of the CPI-III; there does not appear to be any obvious reason why the A and Q options of the first item on page 5 have failed. There are a number of instances of options which correlate more strongly with a mode total other than their own; thus option 12 on page 3 which is an R option correlates more strongly with the A total and option 9 which is an A option correlates more strongly with the Q total. This evidence points to some weaknesses in test construction but it is valuable in enhancing construct validity because reexamination of these options reveals impurities which escaped notice when the instrument was scrutinised.

It is interesting to speculate why so many poor options appear on the first page. It is reasonable to postulate that the four items are unsatisfactory but as 7 of the 8 options which failed to reach the 0.05 significance level are on the first page this seems unlikely. It is a fact that students of science are not generally asked for information on preferences and they may therefore take time to adjust to the new task. If this is so, and there is some evidence to sustain it in the analysis of reliability, it suggests that further examination of the phenomenon is called for and also that it is advisable to discard the scores which were registered on the first four items on cognitive preference tests in the future.

4. <u>Reliability</u>

reliability coefficient	R	A	Q	P
Cronback α	0.833	0.693	0.520	0.796
1st half: 2nd half	0.617	0.632	0.758	0.626

These figures, with the possible exception of the Cronbach α figure for the Q mode, are similar to the figures quoted by Tamir from 9 different sources (Table 3, p.115, 1977a).

The eight items on pages 3 and 4 of CPI-III were retested 35 days after the first administration. No warning of the retest was given at the first test and only school holidays occurred between the two events. Test-retest correlation coefficients were calculated and these are shown in <u>Table 5.6</u>. The significance level of 28 of the 32 options exceeded 0.01 and only one option failed to exceed significance level 0.05. This is considered to be evidence of the satisfactory quality of the normative data.

All possible measures of data quality have been examined and reported in this section. Important information about some inadequacies of the

Table 5.6 : Test-Retest Correlation Coefficients (x 10^2) (N = 45/IV/1976)

	Page 3				Page 4			
Item	1.	2.	3.	4.	1.	2.	3.	4.
Option								
R	66	59	49	69	50	75	45	63
A	59	52	70	46	39	63	48	52
Q	50	49	37	52	56	40	<u>35</u>	59
P	34	67	31	47	<u>24</u>	39	57	44

Note (1) $r > 0.288$, $p < 0.05$; $r > 0.372$, $p < 0.01$

(2) coefficients underlined are not significant.

instrument has been revealed but the evidence supports the contention that this procedure for gathering normative data on cognitive preferences is reliable and that statistical comparisons between normative and ipsative data are justified.

The next stage in this study starts from the premise that the measurement of cognitive preference on a normative scale is the most satisfactory method of measurement. Confidence in the quality of the data which is gathered in this way rests on the findings reported above. It is believed to be the best method because it reflects preference with greatest accuracy and because it exposes weaknesses in the test instrument with precision. (It is, however, unlikely that the method will be adopted for purposes other than instrument evaluation and improvement because it is laborious and data analysis requires extensive computer memory).

The interest moves on to the relationships between normative data and both the ipsative data which may be generated from it as well as ipsative data which is gathered independently of it. Special attention will be paid to features which appear during analysis of ipsative data which might be described as mere artefacts of the procedure which was used rather than as properties inherent in the data.

Ipsative data can be generated from normative cognitive preference data in a number of ways. Firstly each item can be treated as a subtest which is distinct from all other subtests, the mean score for the four options of that item can be calculated and then subtracted from each of the options of that item. This method yields ipsative scores which preserve the intensity of preference. Another method was to subject the normative scores to the scoring procedures which have been used in the past. Thus as Heath (1964) required his respondents to select only the most preferred option 'Heath data' was prepared from the normative scores by counting the number of items in which each mode was most preferred. Kempa and Dubé (1974) and later Tamir in several studies required all options to be voted upon, (by scoring the four options 4, 3, 2, 1) and then summed the item scores to yield a total score for each mode. Kempa and Dubé data were prepared by assigning votes in the order of preference. Mackay (1971) and later Brown (1964) asked their respondents to give +1 to the most preferred otpion and -1 to the least preferred option, i.e. +1, 0, 0, -1, and then they summed to obtain the total given to each mode. In like manner, Mackay data can be obtained from normative data. It was postulated that the students might be more aware of dislike than preference and so a 0, 0, 0, -1 procedure was also examined. The analysis of all four methods of scoring, i.e. Heath, Kempa and Dubé, Mackay and dislike, will be reported. In addition the normative data which were gathered on 45 IVth year students is compared with normal (i.e. ipsative) cognitive preference data which were gathered on 78 IV year students. The data are given in Table 5.7.

It is necessary to postulate that if the respondents had been asked to score their preferences in one of the four methods described above they would have done so in a manner which could be predicted exactly from the normative method which they did in fact use. The loss of reliability incurred by each of the methods is examined here and examination of the mode total scores by analysis was undertaken to see which portrayed normative behaviour with least damage. Cronbach α coefficients and test-retest coefficients were calculated for each mode to assess the reliability of each method of scoring.

The coefficients in Table 5.8 show that with all three statistics the 4, 3, 2, 1 method of scoring is superior to 1, 0, 0, 0 and 1, 0, 0, -1 and that the indication of least preferred options is actually detrimental to the reliability. The figures given in the note at the bottom of Table 5.8 are particularly interesting because they were obtained with ipsative scores which were gathered in the normal manner. The Cronbach α reliability coefficients with scores made ipsative by (4, 3, 2, 1) scoring (i.e. Kempa and Dubé's method of scoring) compare very closely with those in the note to the table. The general conclusion to this part of the study is that there is no difference between the reliability of normative data and either the ipsative data which may be prepared from it or traditional ipsative data. Thus, it is reasonable in so far as reliability is concerned, to use ipsative data.

Table 5.7 : Means and Standard Deviations of Normative scores and of Ipsative data prepared from them by different scoring procedures (N = 45/IV/1976)

	R	A	Q	P
Normative:	141.09 (33.41)	132.53 (29.35)	123.44 (23.52)	122.73 (25.25)
Ipsative:				
1, 0, 0, 0 (Heath)	2.96 (1.70)	1.96 (1.33)	1.49 (1.47)	1.60 (1.45)
4, 3, 2, 1 (Kempa & Dubé)	22.62 (4.33)	19.96 (4.79)	17.98 (4.68)	19.44 (4.18)
1, 0, 0, -1 (Mackay)	13.11 (6.23)	9.71 (4.39)	8.69 (5.08)	8.49 (5.53)

Table 5.8: Product Movement Correlation Coefficients between Normative Data and Ipsative Data prepared from them by different scoring procedures. (N = 45/IV/1976)

Coefficients (x 10^3)

Cronbach α Scoring Procedure	R.	A.	Q.	P.
1, 0, 0, 0	566	478	467	712
4, 3, 2, 1	750	635	541	771
1, 0, 0, -1	721	566	447	731
0, 0, 0, -1	732	464	249	600
(Normative	833	693	520	796)

Test-Retest Correlation Coefficients				
1, 0, 0, 0	401	354	393	347
4, 3, 2, 1	366	651	583	399
1, 0, 0, -1	362	197	340	341
(Normative	585	689	567	501)

Product Movement Correlation between Normative Scores and Ipsative Scores				
1, 0, 0, 0	715	617	572	436
4, 3, 2, 1	719	720	736	557
1, 0, 0, -1	677	486	438	366

Note: the Cronbach α coefficients obtained with CPI-I and a population of 717 students (scoring procedure 4, 3, 2, 1) were:-

633	776	548	669

These compare favourably with the figures given for 4, 3, 2, 1 scoring above and they are interesting because they were obtained with traditional ipsative scores.

It is not adequate, however, to rely only upon reliability before employing ipsative scores. It is also important that the two types of data should furnish the same information on inter mode correlation and on factor analysis. The correlation coefficients which are reported in Table 5.8 are strong and particularly so for the (4, 3, 2, 1) scoring procedure. The rotated factor analysis data for this data aregiven in Table 5.9 and arenot quite so straight forward. It may be recalled that Kempa and Dubé (1973) and King (1972) among other workers found loadings on one factor of a two factor analysis with +R and -Q and on the second factor they found +A and -P. These loadings are found on Factors II and III with (4, 3, 2, 1) scoring when a three factor analysis is performed but under these circumstances 50% of the variance is on Factor I with loading on R and -P. The other scoring procedures with three factors are similar. When two factors are examined the loadings with ipsative scores are not as previously reported by any earlier workers other than Mackay (1971). Inspection shows that the loadings are +R and -P on Factor I and +A and -Q on Factor II. The original normative data from which the ipsative data was derived loads +A and +R on Factor I and -Q and -P on Factor II but when three factors are used Q alone is heavily loaded on Factor III and +A and -P loadings are found on Factor II. The loading pattern for scores made ipsative by subtraction of item mean in which score intensity is preserved resembles the patterns of other ipsative scores more closely than it resembles the factor loading on normative scores.

The general conclusion which may be drawn is that the act of obliging the respondent to 'quantise' his scores by a voting or rank order procedure using whole intervals is not responsible for differences in factor loadings but rather it is the act of turning normative scores into ipsative ones that is responsible. Ipsative scores in which intensity is preserved resemble other ipsative scores more closely than normative scores and one must conclude, very tentatively, that there is a difference in behaviour that requires further examination. This conclusion must be tentative because the sample size was small (N = 44).

Section 5.3 : The 'Subject-Specific' Element of Cognitive Preference Behaviour

Tamir's suggestion that cognitive preference behaviour has three facets

Table 5.9 Rotated Factor Analyses of Normative Data and of the same data after being made Ipsative. (N = 44/IV/1976).

loadings ($\times 10^2$)

		Three Factors			Two Factors	
	Factor	I	II	III	I	II
Raw Scores (normative)	R	70	25	-20	72	25
	A	50	55	-15	50	49
	Q	12	-27	72	01	-64
	P	-18	-49	36	-23	-59
Scores made Ipsative by subtraction of Item Mean	R	77	04	-51	85	26
	A	20	84	-38	25	89
	Q	-31	-37	81	-44	-74
	P	-69	-60	01	-67	-44
1, 0, 0, 0 Scores (Ipsative)	R	86	-12	-45	93	08
	A	06	92	-26	09	89
	Q	-23	-26	87	-38	-68
	P	-83	-44	-12	-79	-22
4, 3, 2, 1 Scores (Ipsative)	R	79	03	-52	87	25
	A	19	89	-29	23	88
	Q	-29	-42	81	-42	-77
	P	-72	-59	03	-69	-40
1, 0, 0, -1 Scores (Ipsative)	R	85	-15	-39	91	02
	A	01	83	-25	04	82
	Q	18	-15	81	-33	-56
	P	-79	-36	-11	-75	-16
Eigenvalues		10.308	3.116	2.583	10.308	3.116
Cumul. %age of variance		51.5	67.1	80.0	51.5	67.1

Note: Loadings $> .39$, $p < 0.01$ loadings $> .38$, $p < 0.01$

has been reported in Chapter 2 (Section 2.3). He described these three as general, disciplinary and subject-specific. The subject-specific element has not been confirmed so far by experimental data, but the opportunity for checking on it was seized. The 24-item cognitive preference test was administered to 44 students at IVth Form Level. The administration permitted the respondents to give each option any preference score between 1 and 99 (as described in Section 5.2). The test covered several distinct areas of chemistry, including physical properties and purification techniques, preparative aspects of chemistry, specific chemical features in relation to the halogens and to ammonia and aspects of rate of reaction phenomena. It would be expected, from Tamir's hypothesis, that cognitive preferences would show pronounced consistencies within such subject areas and this was examined.

The strategy for this examination was as follows. Preference ratings received for each of the four response categories were factor-analysed, using the usual technique involving principal component analysis followed by Varimax rotation. In each case, the Varimax rotation was restricted to four factors which ensured that at least 50% of the total variance was considered. (Listing of variance percentages - see Table 5.10). Loadings of items on one and the same factor were taken as evidence that such items were "undifferentiated" in terms of their subject-matter content. Where items loaded on different factors, the possibility was considered of this being due to subject-matter influences, as proposed by Tamir. Separate analyses were carried out for R, A, Q and P responses. (See Table 5.10).

It should be noted that the data used was normative rather than ipsative and there was consequent loss of some discrimination. (This problem is discussed in Section 5.2). Further, the total sample size of 44 would not normally be regarded as adequate for a fully valid factor-analytic data evaluation but it is thought to be acceptable for the purpose of this provisional analysis. A third point concerning the format of the

Table 5.11 : Item Description and Qualitative Factor Loadings (Summary of Actual Loadings shown in Appendix K) N = 45/IV/1976

Item Description	Loading (Qualitative on Factors) R	A	Q	P
1. Sublimation	I	III	IV	I
2. Lime Water	I	II	(-ve)III	III
3. Oxides of Carbon	II	IV	IV	I
4. Prep. of CO_2	III	I		III
5. Distillation	I	III	I	III
6. Chromatography	I	II	I	I
7. Crystallisation	I	I	I	III
8. Melting pt.-Purity Test	I	II		I
9. Prep of Cl_2	I	I	I	III
10. Reaction of Fe with Cl_2	IV	I	III	I
11. Uses of chlorine	IV	I	III	I
12. Halogens (Periodic Table Group)	I	II		IV
13. Catalysts and Prep of Oxygen	I	III	III	IV
14. Slow Reaction (Rusting)	I	I	IV	I
15. Fast Reaction (Fireworks)	II	III	III	II
16. Nature's Catalysts-Enzymes	I	III	III	
17. Solutions - Solubility	I	II	II	I
18. Water of Crystallisation	II	I	II	I
19. Solution of Gases	II	I	I	(IV)
20. Solution of Acids	I	IV	I	I (II)
21. Physical Properties	I	I	III	IV
22. Manufacture of Ammonia	II	III	IV	II
23. Behaviour of Ammonia in H_2O	III	I	I	II
24. Tests for Ammonia	IV	IV	III	I

items is also worth noting. The items used in this cognitive preference test were constructed with a substantially diminished stem. This was done to conserve information for possible use in an option and to minimise confusion. It was not thought to change the character of the items nor the students' modes of responding to them.

Results and Discussion

R - Responses

The distribution of items, according to R - responses which they elicited, among the four factors is summarized thus:

	No. of Items
Factor I :	1, 2, 5, 6, 7, 8, 9, 12, 13, 14, 16, 17, 20, 21
Factor II :	3, 15, 18, 19, 22
Factor III :	4, 23
Factor IV :	10, 11, 24

Table 5.10 : Factor-Analysis of R, A, Q, P responses

Response Category	Percentage of Variance accounted for by 4 Factors	Percentage of Variance accounted for by Factors after Varimax rotation			
		FI	FII	FIII	FIV
R	59.7	64.0	15.7	11.8	8.5
A	56.0	55.3	17.7	16.5	10.5
Q	50.3	50.8	18.0	16.2	15.1
P	51.7	50.4	22.6	15.8	11.2

The table showing actual loadings on the factors, subsequent to Varimax rotation, is given in Appendix K.

Fourteen of the items load on Factor I, a further ten are found associated with Factor II, whilst the remaining five items have their main loading on Factors III and IV, respectively. Analysis of the themes brought together by items loading on the various factors does not reveal any immediately obvious patterns. Factor I, for example, brings together diverse themes such as

i. Sublimation of CO_2

ii. Lime water in CO_2 testing

iii. Distillation

iv. Chromatography

v. Crystallisation

vi. Melting point testing

vii. Preparation of Chlorine

viii. Halogens (as a group in the Periodic Table)

ix. Catalysis in the preparation of oxygen

x. Rusting (as a slow reaction)

xi. Enzymes as "natural" catalysts

xii. Solutions and solubility

xiii. Solutions of acids

xiv. Physical properties of ammonia

Within this array of themes, certain clusters may be identified, e.g. one comprising themes iii, iv, v and vi all of which deal with separating and purifying materials, or one comprising themes ix, x and xi all of which concern aspects of reaction rates. However, such clustering is at best tenuous and, whilst it may indicate some measure of "subject-specificity" of cognitive preferences, it is outweighed by the substantial loading of so many items on one factor. This would suggest that the similarities among items are more prominent than the differences.

The second factor identified in the factor-analysis of the R - responses does not contradict this; on the contrary, the themes which are brought together in this factor (oxides of carbon, fast reactions, water of crystallisation, solution of gases) show little thematic coherence, and this provides little support for the Tamir hypothesis.

A - Responses

The distribution of the items, in terms of their A - responses, among the four factors is as follows:-

	No. of Items
Factor I :	4, 7, 9, 11, 14, 18, 19, 21, 23
Factor II :	2, 6, 8, 12, 17
Factor III :	1, 5, 13, 15, 16, 22
Factor IV :	3, 10, 20, 24

As in the case of the R - responses, none of the factors brings together items of a distinct thematic connection. Some tenuous links could be suggested as before, but one cannot have much confidence that these are really significant groupings.

It is also noticeable that for the A -responses the distribution over the four factors appears more even (and random?) than for other response types. This may have its cause in the nature of the A - responses itself: although these were constructed to reflect genuine "applications", it was not always possible to eliminate elements of 'recall' or "questioning" from such responses. This is a reflection of the difficulty associated with the design of good cognitive preference items.

Q - Responses

These distributed over the four factors as follows:-

	No. of Items
Factor I :	5, 6, 7, 9, 10, 12, 19, 20, 23
Factor II :	17, 18
Factor III :	2 (-ve loading), 11, 13, 15, 16, 21, 24
Factor IV :	1, 3, 14, 22

Items 4 and 8 do not load strongly on any of the four factors and are thus not included in the listing.

Factor I brings together items which, in terms of the item analysis of the Q - responses, are generally deemed to be satisfactory. There are certain thematic connections here, especially the following two:

(i) Items 5, 6 and 7 (all dealing with techniques for the separation and purification of materials).

(ii) Items 19, 20 and 23 (all relating to solutions in the broadest sense).

Whilst this may marginally support the Tamir hypothesis, there is no obvious reason for the co-appearance of the two groups of items on the same factor other than that all the items have genuine Q - responses.

Factors II and IV are 'minor' factors, in terms of the number of items accounted for, and so do not merit discussion. Factor III, which brings together seven different items with respect to their Q - responses; inspection of these items shows no obvious thematic connections of groupings. What is noteworthy, though, is that in these Q - responses, issues are raised/touched upon which involve aspects of chemical principles. Perhaps this explains the communality of these items.

P - Responses

The following distribution pattern was found:-

	No. of Items
Factor I :	1, 3, 6, 8, 11, 14, 17, 18, 20, 24
Factor II :	15, (20), 22, 23 (Item 20 also loads on Factor I)
Factor III :	2, 4, 5, 7, 9, 10
Factor IV :	12, 13, 19 (weakish loading), 21

The general pattern observed for P - responses is not different from that for the other responses. Thematic groupings, where they exist, appear submerged in the broader clustering of items on the basis of their overall character.

General Conclusion

The present data do not offer any real support for the Tamir suggestion that cognitive preference responses should incorporate a strongly subject-dependent element, in addition to the differentiation according to response-type which has been the foundation of all cognitive preference work so far.

It must of course be recognised that the themes of the cognitive preference items used for this part of the study may not be sufficiently highly differentiated to reveal any subject-dependent variations in students' response pattern: all items in the test fell within the framework of O-level chemistry and may thus, from the students' point of view, have appeared to be exemplars of the same subject. Until the Tamir suggestion is tested on more differentiated subject matter, it cannot be regarded as "experimentally established".

Section 5.4 : Conclusions (Reference to Research Problems in brackets)

It will be observed from the new work in Section 5.1 that in future confidence can rest in the relationship between the modes of a cognitive preference test and the options which are used to assess those same modes (A.1a and A.1b). Evidence was given which shows that students do find the options to be sufficiently mode specific for abundant confidence to rest in the validity of the cognitive preference instrument. Extensive analysis of scoring procedures shows that scores which are gathered in an ipsative form give similar results regardless of the precise scoring procedures used to gather the data (A.2). If normative score is first gathered and then made ipsative there is difference in the analyses of the two bodies of data but this analysis is incomplete because there is no evidence from comparison of two identical samples, one using ipsative scoring and the other using normative. The present findings are not in direct conflict with those of Tamir and Lunetta and of Closs because the opportunity to drive this research to its logical conclusion was not taken. Finally the subject-specific element of cognitive preference behaviour was not established.

CHAPTER 6 : THE STRUCTURE OF COGNITIVE PREFERENCE AND THE EVALUATION OF PUPILS

Section 6.0 : Introduction

The data structure of the cognitive preference test is three dimensional by definition since the four modes are four stimulus points and as such they are locatable in three dimensional space. (This is because n stimulus points can be located in (n - 1) dimensional space). Investigations of particular data structures seek to establish the relative location of stimulus points in the data space and/or to find a suitable reduction of data space that is needed if the stimuli are to be accommodated in less than (n - 1) dimensions.

The procedures which were used by Kempa and Dubé and by Mackay had these two objectives in mind. Kempa and Dubé employed factor analysis and obtained two major factors after rotation which resulted in two bipolarities. They found that the R mode scores loaded with -Q mode scores on one factor of an analysis with two factors and the A mode score loaded with -P mode scores on the other. The disadvantage of their procedure was, at least in principle, that it employed statistical techniques, strictly applicable to normative data only, in relation to ipsative data.

Mackay (1971) in contrast employed the unfolding procedure described by Coombs (1964). His reduction of the three dimensional data structure into two dimensional data space apparently gave a format of data arrangement which, after the introduction of axes, may be represented by R/P and A/Q. Mackay's work is therefore in conflict with the findings of Kempa and Dubé and some questions may be raised as to the real nature of the structure of cognitive preference data. In order to shed some light on this, two additional analyses were performed on the data obtained in this study and the data reported by Mackay was re-examined. The additional analyses were an R- and Q- type factor analysis on the same data and a cluster analysis.

Both of these are also strictly applicable to normative data so the findings must be viewed with some caution.

The second part of Chapter 6 focuses on the relationships between the cognitive preference measurements and the more routine measurements of classroom behaviour and operation. These have been outlined in Section 4.1 (B) Part II and will be described in greater detail in the Introduction to Section 6.2. It was anticipated that the information on these relationships would be of considerable value in demonstrating the value of cognitive preference data in student assessment for predictive purposes and in establishing suitability for particular careers. Certainly it was felt that the relationships needed to be explored fully before the inclusion of a cognitive preference test in a battery of assessment tests could be recommended.

Section 6.1 : The alternative methods of demonstrating the structure of cognitive preference behaviour (Research Problem (B.1))

Part A : Mackay's method

Mackay (1971) sought to expose the structure which underlies cognitive preference by the method of unfolding analysis. His method is fully explained in Section 2.4 Part (B). This was first developed by Bennett and Hays (1960) and Hays and Bennett (1961) and described by Coombs (1964). McElwain and Keats (1961) described a simple method of determining configuration from a given set of orderings. They showed that there are only a limited number of distinguishable geometrical configurations of four stimuli intwo dimensional space which each require that 6 out of the possible 24 orderings of the four stimuli are missing: the orderings which are missing are used to identify the precise configuration and the confidence which can be placed on the chosen configuration rests on the percentage of rank orderings which are accommodated in the 18 orderings which are acceptable to that configuration. Wish (1964) and Brown (1975) used this method with cognitive preference data; Wish accepted a configuration on the evidence of 88.6% and 88.9% accommodation but Brown rejected the method since at best it only accommodated 88% of her data. The percentage of data that must be accommodated within the 18 orderings out of the 24 before confidence can be said to reside in the structure which is then exposed, has not been determined.

Mackay considered 3 possible two-dimensional structures (1971, p.170) after rejecting all possible one-dimensional structures. He used orderings obtained from 13 samples of respondents. He concluded that 'cognitive preference stimuli can best be represented in two dimensions ... The ordering of the four stimuli on the first dimension is A, R, P, Q ... and on the second dimension is R, A, Q, P.' All 13 groups exceeded 88.4% accommodation and Sample 1, the best sample, had 96.3% accommodation. This configuration is derived from a structure which is described as the (3, 6, 6, 3) structure since the R mode score is least preferred in 3 acceptable orderings, as is the P mode score, and the A and Q mode scores are acceptable in all the orderings in which they are least preferred. The two-dimensions reported by Mackay lead to a structure of R/P and A/Q. (It will be recalled that Kempa and Dubé found R/Q and A/P).

In Table 5.4.1 (p.147-149) Mackay gave frequencies of all orderings for all 13 groups of data and these are repeated (in condensed form) in Table 6.1 with rank orderings from four other sets of data. For instance, analysis of Mackay's data shows that (3, 6, 3, 6) accounts for 92.05% of the data for Sample I. This is actually greater than the percentage accommodations for all except 4 of the 13 samples of data with (3, 6, 6, 3). (Sample I is one of the four exceptions). Thus it is evident that Mackay did not examine all possible structures with sufficient care; this was therefore remedied and it is reported in Table 6.2. Inspection of the figures show that both (3, 6, 3, 6) and (2, 6, 4, 6) are at least equally acceptable (or unacceptable) as the (3, 6, 6, 3) structure which Mackay selected.

Four further points are worth making in this context:-

(i) The instrument used by Mackay had only 11 items.

(ii) It is possible to recreate scores from the rank orderings given by Mackay; correlation and subsequent factor analysis of these gives the figures shown in Table 6.3. It is evident that the structure which they reveal conforms to the Kempa and Dubé pattern and contradicts Mackay's pattern.

(iii) It will be observed from the figures given at the foot of Table 6.2 that only 12.5% of all the orderings had either the R mode and P mode in the least preferred position. This is indicative of an unbalanced instrument. The similar figures for the other bodies of data in this table which are part of this study are not ideal though they are more balanced. Balance is important since 100% of the orderings with (3, 6, 6, 3) would be accommodated if the R mode and P mode never appeared in the last position of a body of cognitive preference data; if all modes appeared in all positions equally frequently and if 100% of the orderings were accommodated, much confidence would accrue to the configuration which could then be inferred but in the present sample where the frequency is highly distributed, little confidence accrues to the pattern of behaviour that is reported.

(iv) Mackay supported his evidence from unfolding analysis with a Q-sort factor analysis procedure. During this he selected 64 respondents from the 1358 in Group I (the method used for selection was not given) and of these 64 only 4 had either R mode or P mode scores in the least preferred position while 60 had either A or Q modes in the last position. In the analysis 3 of these 4 are rejected because they cannot be accommodated. It seems possible that this lack of balance may have been over-

→ p.124

Table 6.1 Preference Rank Orderings of Individuals

	Mackay (Sample I) N = 1358	CPI-I N = 830 (ipsative)	CPI-II N = 581 (ipsative)	CPI-III N = 439 (ipsative)	CPI-III N = 45 (normative with 24 items)
AQPR	9	19	10	11	29
APQR	9	10	8	5	32
QAPR	5	18	12	8	20
QPAR	13	21	19	11	22
PAQR	12	10	4	6	36
PQAR	20	25	15	13	30
	68	103	68	54	169
RQPA	36	25	19	12	33
RPQA	304	79	59	18	70
QRPA	7	9	8	13	16
QPRA	22	21	19	23	30
PRQA	95	133	85	30	29
PQRA	38	55	42	27	44
	502	322	232	123	222
RAPQ	215	62	37	48	57
RPAQ	353	76	47	31	40
ARPQ	30	24	20	25	95
APRQ	10	19	18	12	40
PRAQ	65	52	41	14	60
PARQ	12	17	13	9	38
	685	250	176	139	330
RAQP	49	33	30	32	100
RQAP	26	7	6	17	47
ARQP	10	49	27	29	95
AQRP	8	46	30	12	47
QRAP	7	6	2	7	31
QARP	3	14	10	14	39
	103	155	105	111	359
	1358	830	581	427	1080

Table 6.2 Unfolding Analysis: %age of orderings accommodated by three different structures and five bodies of data.

	Mackay (Sample 1)	CPI-I	CPI-II	CPI-III N = 439 Ipsative	CPI-III N = 45 Normative 24 items
1. (3, 6, 6, 3) Structure (from which R/P, A/Q structure is inferred)					
no. of orderings not accommodated	45	124	86	63	188
%age of orderings which are accommodated	96.30	85.06	85.20	85.20	82.59
2. (3, 6, 3, 6) Structure (from which R/Q, A/P structure is inferred)					
no. of orderings not accommodated	108	99	73	59	236
%age of orderings which are accommodated	92.05	88.07	87.44	86.18	78.15
3. (2, 6, 4, 6) Structure (from which R/Q, A/P structure is also inferred)					
no. of orderings not accommodated	95	100	68	56	205
%age of orderings which are accommodated	93.00	87.95	88.29	86.88	81.02
%age of orderings with					
R least	5.0	12.8	11.7	12.6	15.6
A least	36.9	39.8	39.9	28.8	20.5
Q least	50.4	30.1	30.2	32.5	30.5
P least	7.5	18.6	18.1	25.9	33.2

Table 6.3 Means, Standard Deviations, Correlation Coefficients and Rotated Factor Loadings of Cognitive Preferences which were recreated from Rank Orderings, Mackay, Table 5.4.1, p.147-149, (1971) (N = 1358)

	Means	Standard Deviations	Correlations ($\times 10^2$)		
R	3.46	0.89			
A	1.93	0.88	-18		
Q	1.73	0.81	-47	-36	
P	2.86	0.81	-38	-50	-08
			R	A	Q

Factor Loadings ($\times 10^2$)

	Factor 1	II
R	<u>95</u>	-07
A	-13	<u>93</u>
Q	<u>-72</u>	-18
P	-28	<u>-76</u>

Note: this data was recreated from frequencies of orderings by awarding 4 points to the mode which was most preferred, 3 to the next, 2, and 1 to the least preferred score in each ordering.

looked and thus if a more balanced sample had been used Mackay would have found that the Q-sort procedure did not confirm the structure which was revealed by the unfolding analysis.

Part B : R- and Q-type Factor Analysis

The report on this piece of work on ipsative scores is presented here because it serves as evidence of the structure which underlies cognitive preference modes. Q-type factor analysis of four total mode scores with 60 respondents is analagous to R-type factor analysis of the scores of four individuals each with 60 measurements: both types of analysis with the same data are presented for comparison. Scrutiny of the loadings in Q-type analysis allows respondents with similar patterns of cognitive preferences to be identified and assigned to groups. It was postulated that the mean scores of the individuals in the groups would reveal evidence of structure.

The data which was used was obtained from respondents who completed CPI-III normally but who were rejected from the principle sample because they failed to complete one or more of the other tests in the battery. A sample of only 60 respondents was used here because factor analysis of a 60 x 60 matrix utilised the full capacity of the memory of the computer. Use was made of the Q-type factor analysis with two factors to assign each of the 60 respondents to one of five groups according to loading (+ and - on each of the two factors and a fifth, with no obvious loading, formed the reject group). 17 loadings were greater than 0.919 ($p < 0.01$ with $N = 4$); and a further 21 had loadings greater than 0.811 ($p < 0.05$). 6 of the remaining 22 had insufficient loadings on one or other of the two factors and were allocated to the reject group. The mean mode scores and standard deviations are shown in Table 6.4.

Table 6.4 R-type and Q-type Factor Analysis of the same data. (N = 60/V/1976).

		Mean	Stan. Dev.	Correlation Coefficients ($\times 10^2$)		
R-type	R	14.13	8.47			
	A	6.93	10.18	27		
	Q	11.97	8.83	-73	-52	
	P	15.35	9.85	-51	-79	32
				R	A	Q

($r > .33$, $p < 0.01$)

Factor Analysis (loadings $\times 10^2$)

Factor	I	II
R	03	92
A	93	13
Q	-16	-89
P	-91	04
Eigen value	2.567	0.955
Cum. %age of variance	64.2	88.0

Q-type : Mean Mode Scores and Standard Deviations with mean 'O' Level Grade

	N	loadings	$\bar{R}$	$\bar{A}$	$\bar{Q}$	$\bar{P}$	$\bar{O}$ Level Grade
Group I	9	+ive on Factor I	13.22(7.97)	21.7 (4.15)	7.88(5.01)	0.00(4.60)	2.77
II	22	-ive on Factor I	12.14(6.43)	-2.18(6.83)	16.14(7.26)	21.77(7.32)	1.91
III	5	+ive on Factor II	1.20(6.83)	11.40(9.69)	21.40(7.40)	14.00(6.96)	2.00
IV	18	-ive on Factor II	20.00(5.91)	10.11(5.61)	5.11(5.41)	12.61(6.09)	2.39
V	(6	rejects)					

Total : 60

The data in Table 6.4 show strong evidence of the same bipolar structure (A /P, R/Q) in the loadings on the two factors of the R-type analysis and in the mean mode scores of the Groups in the Q-type analysis though the evidence is slightly less strong in the latter. Respondents in Group I with high A mode scores and low P mode scores loaded positively on Factor I and those in Group II with these scores reversed loaded negatively on this factor but there is some support for A/Q, R/P structure since the R score of Group I and the Q score of Group II are also high. Groups III and IV, however, lend unequivocal support to the A/P, R/Q structure.

Part C : Cluster Analysis**

The hypothesis that natural groupings occur in the normative data was further examined by cluster analysis using Ward's method (1963). This is a hierarchical technique in which the individual respondents are classified into groups or clusters and the process is repeated at different levels to form a tree. At each step in the analysis, union of every possible pair of clusters is considered and the loss of information which results is measured by the total sum of squared deviations of every point from the mean of the cluster to which it belongs. The two clusters where fusion results in the minimum increase in 'error scores of squares' are then combined and so, ultimately, spherical clusters with minimum variance are generated. (See Everitt (1974, p.15-18)).

The normative scores of 45 respondents on four of the six pages of CPI-III (pages 1, 2, 5 and 6, i.e. 16 cognitive preference items for each respondent) were analysed. At the 10 cluster stage there were three clusters which embraced 30 individuals and 7 more for the remaining 15 individuals. Fusion of these small fragments was not complete until the 3 cluster stage and it was the behaviour of the individuals in these three clusters which was examined by reassembling the 16 items for each of the three clusters separately and converting the mean option scores to Z-scores. The mean mode scores for all three clusters and the Z-scores for Cluster 2 are given at Table 6.5.

** Computed by programme CLUSTAN IB, UMRCC 7600 version.

Table 6.5 : Cluster Analysis of Normative Data with 3 Clusters (N = 45/IV/1976)

Cluster	No. of respondents in Top Group	No. of respondents in Bottom Group	F-Ratio (sum)	Mean Mode Scores (see note) (and standard deviations) R	A	Q	P
1	10	0	3.65	101.00(30.99)	100.55(27.61)	137.11(24.45)	146.00(13.98)
2	9	7	2.68	156.31(25.16)	148.56(21.76)	134.88(14.93)	134.38(9.79)
3	2	17	3.33	148.58(28.50)	135.00(25.606	106.68(20.43)	100.47(21.71)

Note : mean score on all 16 items (x 10^{-1})

Option Z-scores for Cluster 2 (N = 16; 9 from the top group and 7 from the bottom group

		Option Z-scores (x 10^{2})				
Page	Item	R	A	Q	P	
1	1	29	33	37	-10	*
	2	13	38	0	24	Ø
	3	60	52	30	-02	
	4	44	20	39	41	
2	1	63	22	51	79	*
	2	00	20	-11	01	Ø
	3	08	20	16	62	
	4	09	12	-06	-31	*
5	1	14	20	14	-25	*
	2	23	36	34	21	*
	3	25	34	-03	36	
	4	47	65	24	22	*
6	1	32	48	35	46	
	2	45	10	39	51	
	3	58	32	21	51	*
	4	08	32	44	28	*

The option Z-scores for Cluster 1 (with high overall mean mode scores for P and Q) showed strong preference stimuli for P in 11 of the 16 items and

and for Q in 8 of the items and there was strong response for both P and Q in some items. It was not therefore possible to assess whether the 10 individuals in Cluster 1 were activated by strong preference for P (and weak for A) and by strong preference for Q (and weak for R) or whether they were activated by strong preference for P (and weak for R) and strong for Q (and weak for A), i.e. the structure motivating their behaviour was concealed. The same was true for Cluster 3 with high mean mode scores for R and A but the strong stimuli were reversed. It was noticeable that these two clusters were more diffuse than Cluster 2. (The measurement which was used to establish this was the mean of the sum of the four F-ratios, one for each mode, across the 16 items, for each of the three clusters separately. Small F-ratios indicate low variation within a cluster.)

The option Z-scores for Cluster 2 (Table 6.5) show that 9 of the items (marked *) do afford evidence of A/P, R/Q structure while only 2 (marked Ø) supported A/Q, R/P. This middle cluster contains the respondents who experienced strong preference for either P or Q modes but not for both and they thereby revealed the structure of cognitive preference behaviour which was anticipated from other work. It is reasonable to argue by extension that the same structure pervades all respondent behaviour but individuals in Cluster 1 found that in general both P and Q were strong stimuli and those in Cluster 3 found that in general both the P and Q stimulus statements were weak.

Summary of Section 6.1

The evidence which has been gathered together shows that there is no reason to place confidence in the findings which Mackay reported, while the Q-type factor analysis and the cluster analysis do support the findings of Kempa and Dubé; but these should be treated with the same caution that Kempa and Dubé used in interpreting their statistics. The Research Problem (B.1) has, at least, furnished no contradictory evidence.

Section 6.2 : The Evaluation of Pupils : (Research Problems (B.2)-(B.7))

Introduction to Section 6.2

This section is devoted to the analysis of cognitive preference data in the light of other measurements. These were selected because they were thought to have important classroom implications. The justification for this section of the research is that if it can be shown that cognitive preference behaviour has a characteristic bearing on, say, academic achievement or bias, it will be possible to argue in reverse in later work and claim that, because the cognitive preference style of an individual is found to be of a certain type, the academic achievement or the bias may be expected to conform to the pattern that that type has been shown to have. Obviously academic achievement is of considerable importance and its relationship with cognitive preference is explored in Part A using O level grades with Vth Form students and special tests with IIIrd Form students. (Research Problem (B.2)). Research Problem (B.3) is concerned with the possibility that curricular differences may exist in different schools and that such differences in cognitive preference behaviour as exist between schools may be ascribed to their respective curricula. For this reason, a test of global ability in scientific fields was used for the students of all the schools in the major sample and the cognitive preference behaviour at the time of the global ability test and one year later, was examined.

A further and most important dimension of this study was the relationship between cognitive preference and class level (Research Problem (B.4)). There were two approaches to this field of study: the first involved three different samples, one at each of three class levels. All were examined on the same cognitive preference instrument. In the second approach, a single sample was examined at three moments in their schooling. The sample was examined with a different cognitive preference instrument and with a test interval of one year. The data generated by the second of these two approaches has been viewed as a problem of stability and it has been assigned to Chapter 7.

Another Research Problem (B.5) sought to assess the relationship, if any, between cognitive preference and both bias towards science and interest in the study of science, i.e. towards Science Orientation. The rationale for this approach stemmed from the belief that both are essentially affective responses and that some allied behaviour might be expected. Finally it was deemed important to see whether discernable differences in behaviour between schools and between classes in schools could be detected. Positive evidence of a relationship would have had a marked bearing on the work reported earlier in this section and related to curricular influences. In fact the findings proved to be negative and they are reported only for the sake of completion.

This section is concluded with an investigation of the relevance of sex, if any, to cognitive preference behaviour. (Research Problem (B.6))

Part A : Academic Achievement (Research Problem (B.2))

Kempa and Dubé (1973) administered a cognitive preference instrument to a sample of students of 'A' level in Chemistry about six months after those students had taken their 'O' level examinations. They found strong responses on both the 'curiosity' and 'awareness' scales. King (1972) using the same instrument and a very similar sample of students obtained similar results. In both studies students with high rating on the two scales scored well in their Chemistry examinations. Tamir worked in the biological sciences with a substantially larger sample of students than either Kempa and Dubé's or King's and it contained a less highly selective population. When Tamir related the scores on a 30 item multiple choice to cognitive preference scores by analysis of variance he found significance for the R/Q scale,(the 'curiosity' scale) but not for the 'awareness' scale (Tamir, 1976, Tables 2 and 3).

Table 6.6 : Mean Cognitive Preference Mode Scores, Standard Deviations and Analysis of Variance by O Level Grade (N = 439/Vth/1976)

O Grade	N =	R	A	Q	P
A	70	12.06 (8.67)	4.06 (10.96)	13.56 (9.39)	18.26 (10.17)
B	70	15.26 (9.32)	10.13 (10.29)	10.70 (9.95)	12.04 (10.51)
C	70	17.79 (9.38)	13.20 (9.09)	7.59 (8.47)	8.76 (9.15)
Fail	70	15.91 (8.21)	13.67 (8.03)	8.59 (7.23)	10.19 (7.27)
	280				

Analysis of Variance

Source of Variance	Sum of Squares		Mean Squares		F	p
	between grades df = 3	within grades df = 275	between grades	within grades		
R	1551	17807.8	517.0	64.76	7.98	0.01
A	4333	26126	1444.3	95.00	15.20	0.01
Q	1465	21798	488.3	79.27	6.16	0.01
P	3680	24530	1226.7	89.20	13.75	0.01

The data given in Table 6.6 concerns students younger than those in the sample which was used by Tamir. This sample comprised 280 Vth Formstudents and the mean cognitive preference scores are analysed by 'O' Level grades in Chemistry with the failing grades (i.e. Grades D, E and F) grouped together. The figures show that there is little difference in behaviour between those who were awarded a Grade 'C' and those who failed. Analysis of variance, however, shows that there is more significant variance on the A and P mode behaviours than the R and Q modes. This finding is contrary to that of Tamir. It would not be reasonable to argue from these data that for younger students academic achievement is more strongly related to the A/P scale and that for older ones (i.e. those at college) it is the R/Q scale which is relevant because the circumstances of the two studies are so widely different. The finding is interesting, however. A further analysis of similar data, in Part E, shows, also, that the two sexes behave differently and the relationship between achievement and cognitive preference behaviour is studied again in Chapter 10 (see, particularly, Table 10.1).

Analysis of data from still younger students show that they, too, behave differently. In this instance the cognitive preference data was obtained with IIIrd Form students and the exercise which was used (CPI-III) was the same one that was used with the 'O' Level students. The mean preference scores (with standard deviations and analysis of variance) for five groups is shown in Table 6.7. The groups were obtained by partition after rank ordering the students on their score in a test of recall of chemical information. The test was designed specifically for the students of this sample and it contained only information which could be classified as 'knowledge'. The analysis of variance shows that these young students showed significant variance on the R-Mode and on the Q-Mode scores but not on the other two. An identical analysis was performed on five groups previously partitioned by scores on a test of chemical 'understanding': the variance was found to be non significant for all four modes.

Thus, to summarise, Tamir found significant variance on the R and Q modes with college students, while in this study the younger students confirm Tamir's finding and the 'O' Level students contradict it. The new information which emerges from this is meagre; there may be change in behaviour with increasing maturation but further investigation of the data is necessary. This will be attempted in Part C and in the meantime it is reasonable to say that the evidence in response of Research Problem (B.2) is still incomplete.

Table 6.7 : Mean Cognitive Preference Scores of IIIrd Form students partitioned in five groups by scores in a test of Recall of Chemical Information, with standard deviation and analyses of variance. (N = 161/IIIrd/1976)

Means and Standard Deviations

Group	Recall Score	N	R	A	Q	P
1	42.88(5.66)	33	14.91(8.15)	15.06(8.52)	10.39(9.72)	7.39(7.87)
2	33.78(1.92)	32	17.66(9.39)	17.22(8.82)	6.16(9.90)	6.78(9.33)
3	27.03(1.47)	32	20.09(8.14)	18.25(8.41)	5.41(8.96)	4.25(8.51)
4	21.75(1.85)	32	21.19(7.71)	20.16(6.92)	3.34(7.52)	3.03(8.89)
5	14.22(4.55)	32	19.06(7.41)	19.91(6.19)	4.19(5.49)	5.03(6.93)

Variance

	Sum of Squares		Mean Squares		F	p
	Between Groups (df=4)	Within Groups (df=156)	Between Groups	Within Groups		
R	770.6	10792.4	192.65	69.18	2.78	0.05
A	569.7	9900.0	142.43	63.46	2.24	NS
Q	979.1	11619.9	244.77	74.49	3.29	0.02
P	417.7	11214.3	104.43	71.89	1.45	NS

Part B : Aptitude for Science (Research Problem (B.3))

At an early stage in the research, the need for a test to sample the homogeneity of the sample which was drawn from many schools and classes, was foreseen. The first test package of the longitudinal study (i.e for those tested at IIIrd, IVth and Vth Form Level) therefore included the Bristol Aptitude Test, Study Skills, Level 5. It was administered to 9 schools and to 33 classes at IIIrd Form Level. Analysis of variance of cognitive preference scores between groups constructed by partitioning with B.A.T. scores showed no significant differences. Once this homogeneity had been established it was considered acceptable to use the scores from the sample at this level and at subsequent levels as a single body of data. In a subsequent study, two-way analysis of variance using fourteen randomised classes of fifth form students, drawn from 7 schools with two classes of thirteen students in each class, showed no significant variance for any of the four cognitive preference modes. (No doubt streaming was responsible for the highly significant variance of 'O' level grades in these same classes). In another experiment using 7 classes at IIIrd form level, each with 17 students and all in one school, the variance for the P mode only was significant ($p < 0.05$) but an analysis of covariance showed that when the achievement in chemistry score was used as covariate the variance in the P mode score then ceased to be significant.

Careful scrutiny of all the sample populations for significant variance in mode preference behaviour between schools and between classes failed to reveal differences and so it must be concluded that school and class environments do not modify cognitive preference data.

<u>Part C : Class Levels</u> (Research Problem (B.4)

This problem is concerned with the relationship between class levels and cognitive preference. The data gathered for the resolution of this problem were cognitive preference data obtained with the same instrument (CPI-III) from IIIrd, IVth and Vth Form students. The fourth form population proved to be the smallest of the three and a random selection was made from the students of the other two populations to give three populations of equal size. The means and standard deviations and the analysis of variance is reported in <u>Table 6.8</u>. The most interesting fact which is apparent from these data is the significant variance of the A and P-mode scores, the mean values of the A-mode diminishing with increasing maturity while P-mode mean values increase. It must be accepted therefore

Table 6.8 : Means, Standard Deviations, Variance of Cognitive Preference Mode Scores by Class Levels (N = 369/IIIrd, IVth, Vth/1976)

Means and Standard Deviations

Preference Levels Mode	IIIrd Form (N = 123)	IVth Form (N = 123)	Vth Form (N = 123)
R	18.95 (8.47)	17.41 (10.63)	16.69 (8.68)
A	18.14 (7.96)	15.38 (9.21)	12.26 (9.57)
Q	6.28 (8.60)	7.74 (10.51)	8.73 (9.25)
P	4.60 (8.76)	7.37 (9.51)	10.69 (8.79)

Variance

	Sum of Squares		Mean Square			
	Between Levels	Within Levels	Between Levels	Within Levels	F	p
R	328.2	31992.8	164.1	87.41	1.88	NS
A	2127.6	29491.0	1063.8	80.58	13.20	0.01
Q	373.1	33205.9	186.6	90.73	2.06	NS
P	2287.7	30042.2	1143.9	82.08	13.94	0.01

that these data confirm convincingly the classroom experience of many teachers that as maturation increases, so pupils become increasingly interested in the principles which underlie their chemistry and less interested in the uses to which that information may be put. In other words there is significant movement along the 'application' axis of the double bipolarity behaviour which is such a strong feature of other cognitive preference work.

Part D : Orientation Bias (Research Problem B.5)

This problem concerned the relationship between two affective parameters and cognitive preference behaviour. The measures which were used were the rudimentary orientation questionnaire described in Section 4.5 (Part E) and the bias towards chemistry. It was felt that these two measures together would give some insight into the relationship between science orientation and cognitive preference. Both these measures are insensitive but the intention of the work was merely exploratory. It may be recalled from the information in Section 4.5 that bias towards chemistry was assessed by the place assigned to the subject in the list of subjects selected for study at A Level. Strong bias was thought to be shown by a student who preferred it to all other subjects; antipathy was indicated if chemistry was listed among the subjects least liked for study at A Level. This analysis was conducted by selecting 30 students from each of the five categories: chemistry most preferred, chemistry second most preferred, chemistry third most preferred, chemistry omitted from list of subjects liked and disliked and finally chemistry listed among the disliked subjects. The sample size was determined by the fact that one category had only 30 students and the numbers in the remaining categories were pared down by random selection. (See Tables 6.9 and 6.10).

The evidence from these two tables shows that there is significant 'between group' variance in 'application' scores (A and P, Table 6.9) when the groups are determined by Science Orientation and there is, also, a strong indication that the 'between group' variance in 'curiosity' scores is significant when the groups are fashioned by bias towards chemistry (Table 6.10). Those who plan to pursue their chemistry are motivated by curiosity, while those who express a general interest in scientific things are interested in the underlying principles and eschew the applications. This pattern of behaviour conforms to the pattern that might reasonably have been anticipated but it is gratifying to find that logical expectation is confirmed by data.

→ P.139

Table 6.9 : Means, Standard Deviations and Analysis of Variance of Cognitive Preference Mode scores partitioned by Science Orientation (N = 439/Vth/1976)

	Science Interest	N=	R	A	Q	P
1	11.53(1.46)	55	12.76(7.22)	6.11(11.33)	12.93(8.53)	16.20(10.71)
2	9.40(0.49)	55	16.33(10.06)	9.51(10.70)	9.84(9.50)	12.09(10.54)
3	7.58(0.53)	55	13.15(9.52)	10.93(7.82)	11.56(8.35)	12.69(9.69)
4	6.38(0.49)	55	15.24(9.73)	9.24(10.01)	10.75(10.11)	13.13(9.59)
5	4.84(0.65)	55	15.93(8.98)	10.13(8.80)	10.07(8.38)	12.04(8.58)
6	2.51(0.63)	55	17.62(8.92)	12.91(9.13)	7.00(8.62)	10.38(7.19)
7	0.02(0.82)	55	15.60(7.22)	14.16(8.06)	9.31(7.74)	9.07(6.65)
8	-5.00(3.88)	54	16.46(8.94)	13.85(9.70)	8.46(8.03)	9.19(9.68)

Analysis of Variance

Source of Variance	Sum of Squares Between Groups (df=7)	Sum of Squares Within Groups (df=430)	Mean Squares Between Groups	Mean Squares Within Groups	F	p
R	751.8	35117	107.4	81.67	1.32	(NS)
A	2831.3	40018	404.4	93.07	4.34	0.01
Q	1286.9	33034	183.84	76.82	2.39	0.05
P	2100.7	36628	300.09	85.18	3.52	0.01

Table 6.10 : Means, Standard Deviations and Analysis of Variance of Cognitive Preference Mode Scores by the five Chemistry Bias Groups (N = 439/Vth/1976)

Group	Bias	N=	R	A	Q	P
1.	Chemistry most preferred subject	30	12.37(7.32)	9.60(10.26)	11.37(10.15)	14.53(11.07)
2.	Chemistry 2nd most preferred	30	12.53(8.49)	8.03(11.41)	13.40(9.23)	14.33(9.50)
3.	Chemistry 3rd most preferred	30	10.20(7.94)	11.87(9.16)	11.40(8.48)	14.47(6.89)
4.	Chemistry neither most nor least preferred	30	15.53(9.87)	10.67(9.12)	7.97(7.39)	13.40(8.24)
5.	Chemistry names as one of three least preferred A levels	30	18.63(9.04)	15.37(8.79)	7.23(8.48)	7.23(9.01)
		150				

Analysis of Variance

	Sum of Squares		Mean Squares		F	p
	Between Groups (df=4)	Within Groups (df=145)	Between Groups	Within Groups		
R	1714.3	9493.5	428.53	65.47	6.55	0.01
A	907.92	14397.3	226.9	99.29	2.29	NS
Q	793.82	11591.68	198.5	79.94	2.48	0.05
P	1171.6	12276.2	292.9	84.64	3.46	0.05

Part E : Sex (Research Problem (B.6))

The analysis of the difference in cognitive preference behaviour of males and females which was reported by Tamir was noted in Section 2.3 (Tamir, 1976). Analysis of variance of cognitive preference scores grouped by four levels at 'O' Level examination is an 'awareness' measure rather than a 'curiosity' measure, see Table 6.11. (F values for 'A' and 'P' are 9.04 and 7.57 respectively and 'R and 'Q' values are 4.41 and 4.01 respectively). This result is much as expected from the high element of learning as opposed to comprehension in the traditional and, to a lesser extent, in the modern 'O' level syllabuses. (A χ^2 test of difference in 'O' Level success by the two sexes showed that the slightly higher pass level secured by the girls in the sample was not significant at $p < 0.10$. The value of χ^2 was 4.02).

Analysis of variance by the two sexes separately show that the boys discriminate far more strongly on the A/P axis while the response by the girls is much more even. χ^2 values between the sexes in each of the cognitive preference modes separately show marked differences. The F values are shown in Table 6.11 and χ^2 values are given and summarised in Table 6.12. Tamir's (1976) findings are also given in Table 6.12 for comparison. The similarities between the two bodies of data is remarkable. The generalisation which is possible from this data is that even at the end of the Vth Form year girls are more intellectually ambitious and less susceptible to considerations of utility than boys.

Section 6.3 : The Curiosity and Application Scales

The work by Kempa and Dubé (1973) showed that high academic achievers scored well on the Q mode score on the scale designated as 'curiosity' and in favour of the P mode score on the scale designated as 'Application'. (The former score was obtained by subtracting the Q mode score from the R mode score and the latter by subtracting the P mode score from the A mode

→ p. 143

Table 6.11 : Mean Cognitive Preference Scores of Boys and Girls by 'O' Level Grades, with standard deviation and variance (N = 439/Vth/1976)

	O Grade	N=	Mean and Standard Deviation: R	A	Q	P
BOYS	1	39	13.36 (7.29)	4.95 (11.00	12.00 (8.10)	17.41 (10.85)
	2	75	15.12 (9.71)	11.07 (8.80)	9.73 (8.79)	12.36 (9.74)
	3	88	15.84 (8.72)	14.93 (8.32)	8.58 (7.87)	8.36 (8.25)
	Fail	67	14.10 (7.17)	13.72 (8.07)	9.91 (6.60)	10.46 (7.31)
		269				
GIRLS	1	31	10.42 (9.91)	2.94 (10.81)	15.52 (10.48)	19.32 (9.13)
	2	54	14.74 (10.11)	9.54 (9.70)	12.13 (10.96)	11.67 (9.20)
	3	54	19.33 (7.93)	10.33 (9.46)	7.35 (8.10)	11.20 (9.10)
	Fail	31	19.35 (8.88)	11.55 (9.04)	7.34 (8.37)	10.03 (7.85)
		170				

Analysis of Variance

Source of Variance	Sum of Square: between groups (df=7)	Sum of Square: within groups (df=431)	Mean Squares: between groups	Mean Squares: within groups	F	p
Total Sample						
R	2412	33673	344.5	78.1	4.41	.01
A	5442	37059	777.5	86.0	9.04	.01
Q	2108	32333	301.1	75.0	4.01	.01
P	4281	34826	611.8	80.8	7.57	.01
Boys Only	(df=3)	(df=265)				
R	216	19269	72.0	73.0	0.99	NS
A	2955	20984	985.0	79.5	12.39	.01
Q	319	16734	106.4	63.4	1.68	NS
P	2346	21288	781.9	80.6	9.70	.01

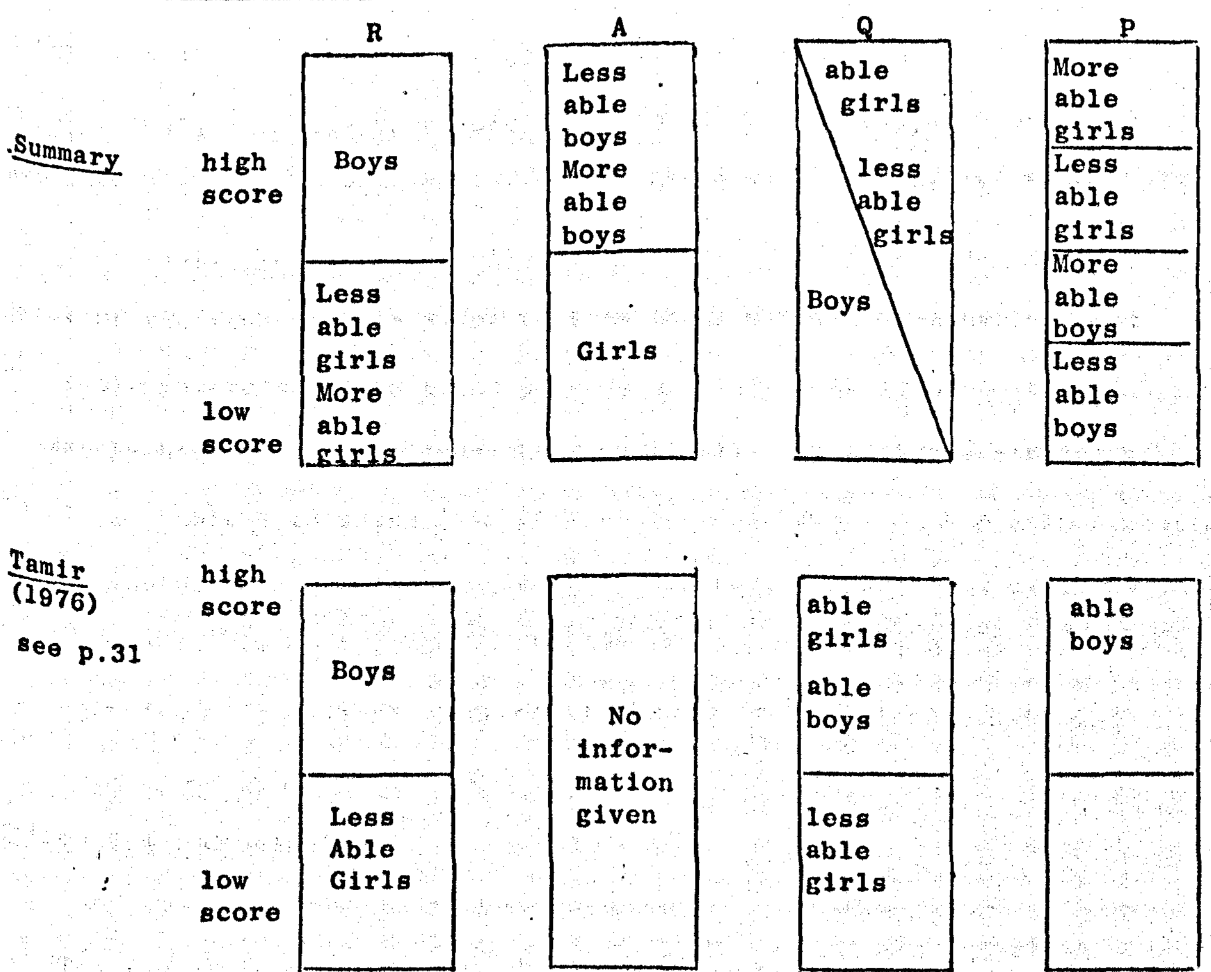
Table 6.11 continued p. 141

Girls Only	(df=3)	(df=166)				
R	1992	144004	664.0	86.8	7.65	.01
A	1453	16075	484.2	96.8	5.00	.01
Q	1744	15600	581.3	94.0	6.18	.01
P	1758	13538	586.0	81.6	7.18	.01

	R	A	Q	P
Sum of Squares				
(s) Between sexes (df=1)	203.6	1034.3	43.9	177.1
(o) Between O Grade Levels (df=3)	1145.5	1987.5	1535.2	3928.7
Interaction, (s) x (o) (df=3)	2207.9	4408.0	2063.1	4103.9
Within cells (df=431)	33672	37059	32333	34826
Mean Squares				
Between Sexes	203.6	1034.3	43.9	177.1
Between O Grade Levels	381.8	662.5	511.8	1309.6
Interaction	735.9	1469.3	687.7	1368.0
Within Cells	78.1	85.9	75.0	80.8
F (interaction)	9.42	17.09	9.17	16.93
p	.01	.01	.01	.01

Table 6.12 : Interaction between Cognitive Preference, Academic Achievement and Sex examined by χ^2 Values. (N = 439/Vth/1976, df = 1)

	Boys More able v. Less able	Boys v. Girls	Girls More able v. Less Able
R	equally high (.83 ; NS)	high low (5.38 ; 0.02)	low R higher R (21.15 ; 0.01)
A	high A highest A (11.2 ; 0.01)	high low (15,13 ; 0.01)	equally low (2.02 ; NS)
Q	equally low (3.26 ; 0.10)	equal (0.10 ; NS)	highest Q high Q (14.15 ; 0.01)
P	fair low (14.56) ; 0.01)	low high (4.51 ; 0.05)	higher ; high (7.70 ; 0.01)



score). This work was repeated with 'O' Level students and with IIIrd form students.

Kempa and Dubé's sample contained only students who had opted for A Level courses in Chemistry; the present sample was larger and embraced all O Level students. It is apparent from <u>Table 6.13</u> that the earlier finding is confirmed since the between group variance in O Level grade, which is the criterion used in both studies for achievement in chemistry, is found to be significant on both scales. Thus those with better O Level grades do tend to show preferences for Q on the 'curiosity' scale and for P on the 'application' scale.

The behaviour of the younger students is different. The variance between groups in scores of recall of chemical information and in understanding chemistry is significant on the 'curiosity' scale but there is no significance on the 'application' scale. Thus,achievement in chemistry at this level is a property associated with 'curiosity' and not with 'application', i.e. the more able students in this sample of young boys have a preference for the 'Q' mode but their preference for the P mode has yet to develop. The data suggests that it will have developed by the time that the students are ready to take their O Level examinations and that their interest in the principles which underly their chemistry will ultimately become dominant over their curiosity. An analysis similar to the one in <u>Table 6.13</u> using IVth Form students (instead of Vth Form students) and the end-of-year examination in Chemistry as the assessment of achievement, shows that the swing towards the P mode was already well developed before the beginning of the final year leading to the 'O' level examination.

<u>Section 6.4 : Conclusions</u>

The two salient points considered in this chapter are whether (a) the evidence substantiates the bipolar structure of information perception behaviour and , if the evidence for bipolarity is unassailable, (b) there is

→ p. 146

Table 6.13 : Means, Standard Deviations and Analysis of Variance of O Level grades of ten Groups partitioned by 'Curiosity' and 'Application' scores (N = 439/Vth/1976)

		Mean O Level Grade partitioned by	
	N	Curiosity Score	Application Score
1.	44	2.89 (0,96)	3.02 (0.99)
2.	44	2.80 (1.29)	2.95 (1.43)
3.	44	2.98 (1.48)	3.16 (1.36)
4.	44	3.07 (1.40)	3.05 (1.43)
5.	44	2.48 (1.29)	3.48 (1.59)
6.	44	3.23 (1.41)	2.75 (1.40)
7.	44	3.05 (1.57)	2.89 (1.23)
8.	44	3.16 (1.43)	2.43 (1.21)
9.	44	2.59 (1.44)	2.64 (1.21)
10.	43	1.98 (1.02)	1.84 (1.29)

	Sum of Squares		Mean Squares		F	p
	Between Groups (df=9)	Within Groups (df=429)	Between Groups	Within Groups		
Curiosity (R/Q)	56.66	791.34	6.30	1.840	3.41	0.01
Application (A/P)	136.29	711.71	15.14	1.66	9.10	0.01

Table 6.14 : Means, Standard Deviations and Analyses of Variance of Recall-of Chemistry and Understanding-Chemistry partitioned into five groups by 'Curiosity' and 'Application' scores (N = 161/III/1976)

	N	CE-R		CE-U	
		Curiosity Score	Application Score	Curiosity Score	Application Score
1.	33	27.55 (9.21)	25.64 (8.45)	10.12 (4.49)	10.15 (4.18)
2	32	24.19 (8.44)	27.25 (10.05)	8.72 (4.01)	9.41 (4.26)
3	32	27.66 (11.10)	27.03 (10.59)	8.91 (4.02)	9.03 (3.21)
4	32	26.00 (9.90)	27.50 (11.17)	7.94 (3.37)	8.22 (5.17)
5	32	33.72 (11.68)	30.97 (10.24)	11.34 (5.98)	9.56 (4.32)

	Sum of Squares		Mean Squares		F	p
	Between Groups (df=4)	Within Groups (df=156)	Between Groups	Within Groups		
Recall of Chemistry						
Curiosity Score (R/Q)	1645.0	17760	411.25	113.85	3.61	0.01
Application Score (A/P)	504.3	165296	126.08	105.96	1.19	NS
Understanding Chemistry						
Curiosity Score (R/Q)	227.8	2919	56.95	18.71	3.04	0.01
Application Score (A/P)	66.7	2940	16.68	18.84	.88	NS

valuable information about student behaviour to be obtained from our knowledge of structure. In Section 6.1 the evidence is shown to confirm Kempa and Dubé's R/Q and A/P structure; the evidence of the alternative structure which Mackay presented is re-examined, found to be incomplete; and yet strongly supportive of Kempa and Dubé's work. Other statistical treatments are also found to be supportive and the evidence in favour of the 'curiosity' and 'application' (or 'utility') scales is found to be strong.

There is significant distribution of O level grades on the 'application' scale but not on the 'curiosity' scale. With third form students, the distribution of recall of chemistry and understanding of chemistry is significant on the 'curiosity' scale but not on the 'application' scale. The changes from IIIrd, to IVth and then to Vth Formers are found to be accompanied by significant changes on the 'application' scale but not on the 'curiosity' scale, while 'interest in science' among Vth Formers behaves differently to bias 'towards chemistry'. Those with an interest in science are significantly more interested in the P mode while those who place chemistry among the three subjects which they would choose to study at A Level are significantly more interested in the Q mode. Computation of an (R-Q) score and a (A-P) score (i.e. actual scores for 'curiosity' and 'application') leads to significance of distribution of O Level grades on both scales and of 'recall' and 'understanding' scores on the 'curiosity' scale only.

There is, therefore, clear evidence of change in behaviour here but if the change is also shown to be unstable it is of no consequence in psychometric studies. The successful evaluation of pupils requires that change in behaviour with age is progressive and therefore predictable. Chapter 7 is, therefore, devoted to the study of stability, while Chapter 10 is used for the evaluation of pupils with both information and perception and information transformation in mind.

CHAPTER 7 : THE STABILITY OF COGNITIVE PREFERENCE BEHAVIOUR

Section 7.0 : Introduction

An extensive examination of stability of cognitive preference behaviour was undertaken because it was considered to be essential to the discussion of the relationship between cognitive preference and cognitive style. The interest here is to demonstrate that cognitive preference behaviour is stable within one administration, that it is stable over a period of years and that the individual's behaviour is stable within the trends in the behaviour of the whole population. A cognitive preference test is valid only if the individual is cognisant of all the factual material which is embodied in it and hence the lapse of time between the beginning of the three year course to 'O' level and the examination itself was selected as the time span for the examination since this allowed time for pretesting and also for three full years of potential change. The fact that this is a time of intense intellectual development was also considered to be important.

It will be recalled that parts of Chapters 5 and 6 were concerned with the quality of the cognitive preference instrument in terms of the evaluation of both instrument and students. This chapter is concerned also with the behaviour of the individual. Measurement of stability of cognitive preference behaviour presupposes a perfect instrument and examines the respondent in its light.

The measurement of reliability and internal consistency has a long and well established pedigree in educational psychology. Both are related to stability of cognitive preference but each has aspects which are unique and demands quantification of relationships during one administration, of changes between one year and the next as well as the treatment of four measurements as one measurement in the expectation that increased accuracy will thereby accrue from the analysis. The examination of the data will be undertaken in four sections.

Section 7.1 will concern the behaviour of each individual on each of the four modes treated separately within the 24 or so items of a single administration. The hypothesis is that each individual shows such a degree of consistency of behaviour from one item to the next that a high level of confidence may be placed in his behaviour. Section 7.2 will also be concerned with the four modes separately. The interest in this part concerns both individual behaviour between one year and the next by examination of mode totals and behaviour and the same items when represented a year after the first presentation. Section 7.3 is devoted to the procedures which achieve an amalgamation of the four modes and to the light which they throw on behaviour. In Section 7.4 the four mode scores are treated as a single 'relationship pattern' and the change of this relationship within the time spanned by the tests is examined and discussed.

Section 7.1 : Mode Reliability

The evidence which is given in this part is presented in order to support the argument that individuals do behave consistently in responding to cognitive preference instruments. The coefficients which are given in Table 5.8 using normative data and ipsative data are the routine measures of test consistency. The interval of 35 days of school holiday which elapsed between the administration of the test and of the retest was considered sufficiently long to obliterate recall and sufficiently short to guarantee no change in cognitive structure.

The test-retest coefficients in Table 7.1 are not as high as the coefficients which are often given by the routine classroom tests, but they do show that the respondents behaviour reflects a positive attempt to express preference in a manner that is genuine and meaningful to the respondent. Similarly, the several coefficients of reliability support the contention that the cognitive preference instruments have an acceptably high degree of internal consistency for each of the four modes separately. It is, therefore, reassuring to find that both the respondent and the measuring instrument are behaving in the manner that was intended by the designers of the method of measurement of cognitive preference.

There is no frame of reference of confidence limit against which these measures may be assessed but experience suggests that they are acceptable in that they afford sufficiently convincing evidence of consistency of behaviour to justify examination of the three remaining measures of cognitive preference behaviour. It is of interest that the relative magnitude of the coefficients in Table 7.1 does not support Brown's prediction which she based on Bloom's hierarchy of cognitive ability (Brown 1975, p.55-57).

Section 7.2 : Mode Stability

The next hypothesis which was examined was that the mode totals of each individual are stable with respect to the remainder of the population. The four preference mode totals for each individual were related to the field independently. The data which was used was gathered by administering three different cognitive preference instruments (CPI-I, CPI-II and CPI-III) at annual intervals to students of 'O' Level Chemistry. 325 respondents completed all three tests.

If the hypothesis is valid (a) preference mode totals should correlate well, (b) each total should remain in each band of the population throughout the test period and (c) it should be possible to subdivide the sample into clusters which contain equal or equivalent scores on all three instruments. The division of the sample into bands was done in two ways. In the first an arbitrary decision was made to divide it into ten bands each with 32 respondents and analyse behaviour on two administrations. In the second, five bands (each with 65 respondents), were examined on the three

Table 7.1 : Coefficients of Mode Reliability using Ipsative Data

Measurement	Instrument	No. of Respondents	Coefficients R	A	Q	P
(a) Test-retest	CPI-III	45	.366	.651	.583	.399
(b) Split-half (after modifcation with Spearman Brown formula)	CPI-III	45	.789	.808	.763	.850
	CPI-I	599	.533	.723	.596	.625
(c) Kuder Richardson (Formula 21)	CPI-III	45	.731	.562	.650	.749
(d) Cronback α	CPI-III	45	.750	.635	.541	.771
	CPI-I	599	.633	.776	.548	.669

Note : see also the coefficients in Table 6.6.

Table 7.2 : Product Moment Correlation Coefficients of Mode Scores (N=325)
(Coefficients x 10^2)

C.P. Instruments	Classes	Years of Administrations	CP Mode R	A	Q	P
I : II	3rd : 4th	1974 : 1975	58	45	41	39
II : III	4th : 5th	1975 : 1976	23	39	26	42
I : III	3rd : 5th	1974 : 1976	11	20	15	28

Note : $p < 0.01$; $r > 0.143$.

test administrations together. The cluster analysis which was selected (for (c) above) contained the largest number of clusters in which there were at least 20 respondents.

A further hypothesis with bearing on stability was examined at the same time. It was postulated that more stable mode totals would be generated by re-use of items in successive instruments than would be obtained by fresh items. Convincing evidence of the validity of this hypothesis would support Brown's contention that response to a cognitive preference instrument is governed more by item content than by compliance with the construct of the cognitive preference test.

Part (A) : Evidence for mode stability from correlation coefficients

Preference mode totals for each of the three administrations were prepared and then subjected to product moment correlation analysis. The coefficients which were obtained are given in Table 7.2. They are positive and all except one are significant at $p < 0.01$. It will be recalled that CPI-III differed slightly in style from CPI-I and CPI-II in order to comply with criticisms by Brown of the earlier style; this difference might be responsible for rather lower coefficients which were obtained with this instrument. The coefficients lend some support to the case for validity of the mode stability hypothesis.

Part (B) : Evidence for mode stability from banding

The rationale for banding analysis starts with the notion that ideal behaviour of both instrument and respondents would generate identical rank orderings of cognitive preference scores in each subsequent administration to a stable population. In practice the behaviour of both is less than ideal and it is expedient to use rank orderings of the first administration to assign the four cognitive preference scores to a band, one for each mode. After the second administration, a count is made of the number of respondents who recur in the same band; a high count indicates a high degree of mode stability. Table 7.3 shows these counts. The first portion of the table concerns the change which occurred between the administration of CPI-I in

Table 7.3 : Band counts with 10 Bands (N = 325)

		R	A	Q	P
The no. of respondents in the same band of CPI-I and CPI-II with	0 neighbouring bands	56	52	53	52
	1	130	136	117	125
	2	193	189	172	187
	3	236	232	220	226
CPI-II and CPI-III	0	64	43	31	47
	1	138	124	108	129
	2	180	190	161	187
	3	228	223	206	222
CPI-I and CPI-III	0	55	33	33	49
	1	108	97	94	115
	2	171	163	139	175
	3	210	204	185	220
Random data	0	30			
	1	87			
	2	144			
	3	182			

July 1974 and CPI-II in July, 1975. The first figure, 56, shows that only 56 of the 325 respondents in the sample (17%) recurred in the same R mode band. As 10 bands were used, random behaviour alone accounts for 10% of this 17%. (The figures in the lowest portion of Table 7.3 were obtained with random numbers in lieu of cognitive preference scores. They show close agreement with chance counts.)

The analysis was extended by widening the count to include those who remain in the same band with those in the neighbouring bands, thus both the one above and the one below were counted. Every count except the top and bottom ones had three bands. These two had only two bands. The analysis was yet further extended to include two neighbouring bands, a maximum of five bands and 3 neighbouring bands, i.e. a maximum of 7 bands in all. The figures will be found in Table 7.3. One year separated the two tests in the second portion of the table (CPI-II and CPI-III) and therefore two years separated the two tests in the third portion (CPI-I and CPI-III). It should be noted that three different instruments were used. Inspection of the data shows that the band count is approximately 1.7 times greater than the random count when the interval is 1 year but the ratio falls to 1.4 (and much less for A and Q modes) when there is a two year interval. The ratio also decreases as the number of bands being counted is increased. On average, therefore, it is reasonable to claim that the counts are 1.5 or 1.6 times greater than the chance value. A χ^2 test shows that the significance level of these ratios is between 0.10 and 0.50 but they are, of course, derived from rank ordered data. There is, therefore, some measure of support for the notion of mode stability and, no doubt, there would be more if the same instrument had been used for all administrations.

A similar analysis in which five bands were used with the same data, shows that while strong preferences and strong rejections of modes are consistent, there is almost random behaviour in the middle bands. (Random data gives counts of 13 for each band in Col. (1) and Col. (2) with a total of 5 x 13, i.e. 65, and in Col. (3) random data gives counts of 3 and a total of 15.) Much of the low count behaviour in Col. (2) can be ascribed to the difference in the style between CPI-II and CPI-III.

On balance it is claimed that there is some evidence for mode stability particularly at the extremes of cognitive preference. This method of analysis is original to cognitive preference studies.

Part (C) : Evidence for mode stability from cluster analysis

The 4 preference mode scores from each of the three administrations were subjected to cluster analysis. A number of different analyses were obtained and the one with 8 clusters was used because each cluster had at least 20 respondents. n Table 7.5 the mean score on all 12 modes for these clusters are given with the mean score after conversion to a Z score and with a letter to indicate the rank order of the variation of

Table 7.4 : Band Counts with 5 Bands

(N = 325)

Mode	Band No.	Number in the same Band in (1) Year 1 and Year 2	(2) Year 2 and Year 3	(3) Year 1, and Year 3
R	1	24	11	4
	2	19	13	2
	3	13	13	4
	4	16	17	7
	5	25	17	9
	Total	97	71	26
A	1	19	14	10
	2	16	13	5
	3	14	15	2
	4	11	7	3
	5	28	18	9
	Total	88	67	29
Q	1	29	20	9
	2	12	11	2
	3	19	13	6
	4	16	11	4
	5	21	13	7
	Total	97	68	28
P	1	30	25	17
	2	11	8	1
	3	13	15	2
	4	16	15	3
	5	20	19	10
	Total	90	82	33

that mean score in the cluster. The letters range from a to l : a signifies least variation and therefore that measurement is a good diagnostic of the cluster, the letter l signifies that there is wider variation on that measurement than on any of the other measurements and therefore less confidence can be placed in it. The procedure is far more sensitive than banding in that the size of each group or cluster is determined by the data.

Inspection of the Z scores in Table 7.5 shows that in each of the four clusters which are numbered 3, 4, 6 and 7 the Z scores for the three R values, for the three A scores, for the three Q scores and for the three P scores are in the same direction. This means that in these clusters which together embrace 60% of the population the modes are reasonably stable. However, this cursory inspection has failed to make effective use of all the data which is in the table. This is due to the fact that the intention of this analysis was to treat mode scores independently of one another. A re-examination of this data with a different purpose will be presented in a later part (Section 7.4).

Part (D) : Link Data

A test was set up in order to determine whether the items which reappeared unchanged in successive test administrations behaved differently from the remaining items of the instruments. Seven items (called, for convenience, 'link items') which were thought to be suitable for inclusion in all three cognitive preference instruments were prepared. They were included in CPI-I and CPI-II but the reshaping of CPI-III to meet criticisms made by Brown would have required alterations in both phrasing and content of these items and so the link items were omitted from this instrument. (The link items appear as numbers 2, 3, 7, 9, 8.5 and 22 in CPI-I and 5, 8, 9, 19, 23, 27 and 28 in CPI-II respectively). The scores on these items in each of the two administrations were correlated and these and other coefficients are given in Table 7.6. There is only one link item

Table 7.5 : Cluster Analysis : Mean Mode Scores, Z Scores and Variance of 8 Clusters (N= 322)

Cluster Number	N =	CPI-I R	CPI-I A	CPI-I Q	CPI-I P	CPI-II R	CPI-II A	CPI-II Q	CPI-II P	CPI-III R	CPI-III A	CPI-III Q	CPI-III P	Mean O Level Grade (stan.dev)
1	20	45.7	23.6	26.4	48.4	48.0	32.3	33.4	55.3	8.65	-7.4	18.4	29.2	1.89
		85b	-80i	-92b	100g	54h	-90e	-51c	79j	-86k	-176a	921	175f	(1.41)
2	51	40.8	20.7	32.5	50.1	43.5	34.0	38.9	51.1	16.2	12.1	9.6	9.8	2.78
		22k	108b	-09h	118a	14g	-72j	09i	381	05e	19c	-06d	-25f	(1.32)
3	60	39.1	37.6	32.8	34.4	47.3	48.7	34.0	38.1	24.0	17.1	3.9	3.4	3.02
		00h	57i	-05c	-55e	48g	791	-40a	-89j	89k	69d	-71b	-126f	(1.43)
4	27	33.4	26.1	39.2	45.2	31.4	35.9	47.1	53.6	4.3	-1.5	23.8	22.9	2.19
		-72a	-56j	81c	641	-92f	-53g	92b	62e	-122b	-117i	153h	110d	(1.00)
5	36	38.7	32.6	35.9	36.7	42.4	38.7	32.6	55.2	19.8	3.6	5.4	19.8	2.50
		-04j	081	36c	-30i	-04f	-24d	-54h	78e	44k	-66b	-54g	78a	(1.23)
6	55	31.9	37.4	40.1	34.5	28.9	48.0	48.3	42.5	10.1	12.3	12.0	8.5	3.04
		-92k	55h	931	-54f	-115c	72g	103e	-46a	-60i	21d	70j	-38b	(0.71)
7.	53	45.6	34.4	27.0	36.9	50.1	42.7	32.2	43.0	14.5	14.3	7.0	11.7	3.36
		84f	25g	-84d	-28h	72j	18a	-58k	-41e	-121	41c	-36i	-05b	(1.56)
8.	20	37.4	36.0	29.2	33.1	41.8	32.2	37.9	56.3	10.5	14.9	14.1	8.5	2.95
		-21k	411	-54j	-70i	00a	-91f	-01b	89c	-56h	47e	44g	-38d	(1.54)
Population	mean	39.1	31.8	33.2	39.4	41.9	41.0	38.0	47.2	15.7	10.2	10.2	12.2	
"	S.D.	7.8	10.2	7.4	9.0	11.3	9.7	9.0	10.3	9.3	9.9	8.9	9.7	

Note (1) : Z scores x 10^2

(2) : Diagnostic Quality is indicated by a letter: the letter 'a' indicates least variance within the cluster, the letter 'l' indicates most variance.

which exceeds the minimum required by Hicks (1970) for significance of correlation coefficients of ipsative data. The evidence in this table either reflects the inadequacy of the items or refutes the contention that response is governed by item content rather than by item construct and it affords no support for stability of behaviour on the four modes.

Table 7.6 : Product Moment Correlation Coefficients obtained with Items linking CPI-I and CPI-II (N = 340)

	Correlation Coefficients (x 10^2)					
	CPI-I : Link Items	CPI-II Non-Link Items	CPI-I Link Items	: Remaining Items of CPI-I	CPI-II Link Items	: Remaining Items of CPI-II
R	23	36	44		58	
A	32	49	55		41	
Q	22	34	30		35	
P	46	39	46		53	

It might be expected that factor analysis of 'link' and 'non-link' scores in the two tests with 4 factors would load the link item scores on the same factors. In practice only the A mode link item scores were found to have significant loadings ($p < 0.01$) on the same factor. In other respects the analysis resembled the factor analyses reported by Kempa and Dubé.

Summary of Section 7.2

Evidence for support of the concept of mode stability was sought from data which were gathered from the administration to the same population of three cognitive preference tests at annual intervals. The data was analysed in four different ways. Product moment correlation coefficients between like mode scores of the three tests were found to be highly significant. Some evidence was found in banding analysis but this required that the data must be rank ordered and it proved difficult to establish confidence limits. Cluster analysis showed that the behaviour of 60% of the population is stable over the three tests. The link data were too incomplete to furnish corroborative evidence. Together these analyses present four interesting ways of dealing with longitudinal data and they give tentative support to the notion of mode stability.

Section 7.3 : Mode Consistency

Mode consistency coefficients quantify the behaviour of each respondent separately on the four modes together, i.e. they are a measure of the individual not of the instrument. An individual with a high mode consistency coefficient will be consistent in giving four votes to the R mode, for instance, in giving three votes to the A mode, two to the Q and one to the P. A large coefficient will occur only when all the stimuli in the options of the instrument are sufficiently strong and the implicit mode sufficiently pure and unequivocal for the respondent to respond consistently to the stimuli even though they are randomly distributed in the items. Further, his preferences must be sufficiently strong and well grounded in his cognitive structure to motivate his selection procedure reliably. It is reasonable to hypothesize that these conditions will be fulfilled only by experienced and mature chemists when working a perfect instrument. Two methods of assessing mode consistency will be examined.

Kendal's coefficient of concordance (W) is one suitable statistic. It is calculated from the total (ipsative) mode scores by the equation:

$$W = 12s/(k^2(n^3 - n)$$

where s equals the sum of squares of observed deviations from the mean score, n equals the number of modes and k equals the number of items ($n = 4$ and $k = 24$). It varies from 0 to +1 and the statistic ($k(n - 1) W$) is used to test the value of W since it is distributed as χ^2 with ($n - 1$) degrees of freedom. (A suitable reference is Galfo and Miller (1970)). Values for W are shown in Table 7.7. It was found that 27 of the 45 respondents had significant coefficients with $p < 0.01$, 4 with $P < 0.05$ and 4 more with $p < 0.1$. The remaining 10 respondents (i.e. 22% of the sample) either appeared to have no preferences or they scattered their votes randomly.

Kendal's coefficient utilises the mode total score but fails to take account of the distribution of votes which leads to the totals. Another coefficient (V) was devised which is of more direct relevance to cognitive preference tests. It embraced the number of times each vote was given to each mode and it was postulated that this would reflect mode consistency much more accurately. An example will best illustrate the calculation of this coefficient.

Student No. 19 voted as follows:

	4 votes	3 votes	2 votes	1 vote	Total Score
R mode	3 times	2 times	8 times	11 times	45
A	0	4	8	12	40
Q	12	8	3	1	79
P	9	10	5	0	76
					240

The coefficient is calculated by taking each line of the table separately. The largest number on the line is multiplied by 3, the second largest by 2, the third by 1 and the smallest by 0 and the figures which are obtained are added to obtain the line sum. For an instrument with 24 items this must exceed 36 and it cannot exceed 72; 36 is deducted from the line sum and the resultant divided by 36 to get the line coefficient. The average of the four line coefficients is the desired coefficient (V).

Thus: (11x3)+(8x2)+(3x1)+(2x0) = 52, 52-36 = 16, 16 ÷ 36 = 0.444

(12x3)+(8x2)+(4x1)+(2x0) = 56, 56-36 = 20, 20 ÷ 36 = 0.555

(12x3)+(8x2)+(3x1)+(1x0) = 55, 55-36 = 19, 19 ÷ 36 = 0.528

(10x3)+(9x2)+(5x1)+(0x0) = 53, 53-56 = 17, 17 ÷ 36 = 0.472

1.999

Mean = 1.999 ÷ 4 = 0.499 = V

The coefficient varies from 0 (for 6, 6, 6, 6, i.e. random voting) to +1 (for 24, 0, 0, 0 voting) and the statistic (4kV/(n-1)) is distributed as χ^2 with (n-1) degrees of freedom. The mean value of V for the 45 respondents was 0.353 (0.102 of this was contributed by the R mode and the remainder by the other three modes equally, i.e. 0.084 each). If any one of the following voting patterns was repeated across all modes and all respondents it would also give a value of $V = 0.3$:

(8, 8, 8, 0) (12, 4, 4, 4) (10, 6, 6, 2)

(10, 7, 5, 2), for instance, gives a value of $V = 0.361$. The value obtained for V lies between these two and so these voting patterns reflect the mean behaviour employed on the instrument. ((10, 7, 5, 2) is significant at $p < 0.01$ and the other three patterns are significant at $p < 0.05$)).

The distinction between the two coefficients lies in the treatment which each will give to a pattern like (4, 9, 6, 5) or (0, 12, 12, 0). W will treat them both as though they were (6, 6, 6, 6) since all patterns sum to 60, while V will acknowledge the consistency in selection of the second most preferred option. W is the more practical statistic if the data is only available as mode totals; V is preferred if the preference data is available on computer cards.

The cognitive preference scores of the 26 respondents with values of V with significance greater than $p = 0.01$ all had two highly preferred modes and two highly disliked modes; 18 had high R mode and high A mode scores and 8 had high Q mode and high P mode scores. 6 of the remaining 19 had one highly preferred or disliked mode.

The mean value of V with another body of students was found to be 0.286. (N = 717/III/1974). If every student in this sample had had a voting pattern of (9, 7, 5, 3) V would have equalled 0.280. The value of 0.286 is significant at $p < 0.05$. The fact that it is less than the value obtained with CPI-III and reported in Table 7.7 may be due to the greater maturation of the sample which was tested with CPI-III.

Confidence was placed in the mode reliability figures given in Table 7.1 but no level of significance was claimed for them. This confidence seems well founded in the light of the significance levels claimed for the mode consistency coefficients in this part.

Summary of Section 7.3

The two statistics show that highly significant behaviour operates when at least half of the population (26 out of 45 respondents) respond to a cognitive preference test and for a third of the remainder (6 out of 19) at least one mode has positive significance. The 30% of the sample who have low coefficients are motivated either by equal preference for the four modes or scored their instrument randomly. Some of the later group may have had difficulty in interpreting their instructions. The previous and subsequent behaviour of this group of students and their eagerness to cooperate with the test, suggests that few, if any, cast their preferences aimlessly. There is, therefore, in every sample a significant proportion of respondents who have four equally forecful preferences.

Table 7.7 : Values of W (Kendal's coefficient) and V.
(N = 45/IV/1976)

No. of the respondent	W	V	No. of the respondent	W	V
1	.044	.210	23	.231	.323
2	.118	.205	24	.202	.335
3	.222	.380	25	.085	.290
4	.356	.453	26	.252	.398
5	.182	.358	27	.279	.420
6	.083	.230	28	.278	.468
7	.030	.290	29	.453	.518
8	.097	.330	30	.238	.420
9	.017	.323	31	.042	.245
10	.247	.425	32	.270	.350
11	.042	.260	33	.045	.285
12	.182	.300	34	.429	.455
13	.094	.205	35	.263	.433
14	.009	.190	36	.356	.455
15	.491	.518	37	.113	.253
16	.386	.475	38	.274	.363
17	.238	.413	39	.042	.268
18	.109	.378	40	.183	.385
19	.431	.495	41	.042	.288
20	.059	.258	42	.205	.340
21	.015	.418	43	.181	.273
22	.325	.385	44	.291	.413
			45	.363	.495

Mean value of W = 0.342 (^{2} = 24.6, df = 3, p $<$ 0.01)

V = 0.353 (^{2} = 11.3, df = 3, p $<$ 0.01)

Section 7.4 : Relationship Stability

Evidence of stable cognitive preference behaviour within each mode in each administration and within each voting pattern in a single administration was given in Section 7.1 and Section 7.3 and claims were made for significant stability of these two measures. No such claim was made for the stability of behaviour on each of the four modes separately when these wereexamined between annual administrations (as they were in Section 7.2). In this section an examination is made of the four modes together and a claim is made for stability between annual administrations. The intention is (a) to demonstrate stable behaviour of individuals relative to the sample population, (b) to explore the direction and extent of change of the population as a whole and (c) to examine the behaviour of groups within that population change. No attempt is made to examine individual behaviour in terms of the cognitive preference model but the behaviour of groups is related to the model.

Part A : Stability of behaviour in successive cognitive preference tests

This was done by a procedure which is best described as discrepancy analysis. The total mode scores for each respondent were considered in rank order and the number of discrepancies between the rank order obtained in one administration and the rank order obtained in a subsequent administration were counted. A discrepancy is defined as a change in rank order and it was calculated thus:-

Example Rank order in Year 1 : A > R > P > Q

2 : Q > R > P > A

Year 1	Year 2	No. of discrepancies
A > R	R > P > A	2
A > P	P > A	1
A > Q	Q > R > P > A	3
R > P	R > P	0
R > Q	Q > R	1
P > Q	Q > P	1
		9.

A complete reversal of the rank order leads to a maximum discrepancy count of 10.

Four digits can be rank ordered in 4! ways, i.e. 24. A 24 x 24 matrix, with 576 cells, each with the appropriate discrepancy count was prepared. The frequency of each of the 11 possible discrepancy counts was made and their cumulative frequency is presented as a proportion of 576 in Col.(c) in Table 7.8. The behaviour of 325 respondents on three cognitive preference tests at annual intervals was examined with this matrix. An analysis was made of the frequencies of the discrepancies which occurred between the first administration and the second, between the second and the third and between the first and the third. This data is presented in Cols. (e) to (m) in Table 7.8. The analysis was repeated with random numbers instead of cognitive preference scores and the cumulative frequencies of the discrepancies between one set of random figures and another set are shown in Col.(d). The figures in Col.(c) and Col.(d) agree closely together but differ from Cols.(g), (j) and (m). The two horizontal lines in the table indicate the division between the body of respondents for whom f_o exceeds f_e. Below the lines random behaviour is increasingly evident. The numbers of students above the lines are 204 (62.8%), 179 (55.1%) and 119 (37.7%) for the three analyses respectively: 119.7, 119.7 and 68.8 are expected from random scores.

The values of t in Table 7.8 are large and highly significant but it could reasonably be argued that this test is insufficiently sensitive to the loss of intensity which has arisen through rank ordering of the scores. It is not possible to place any other quantitative measure on such heavily skewed data but value judgement must confer a large degree of support for the hypothesis of relationship stability where 62.8% of the sample have 3 or less than 3 discrepancies between one year and the next. (A discrepancy of 3 is, for instance, the difference between a score of Q > P > R > A and another with Q > A > P > R in which the A mode has moved from 4th in the order to being the second most preferred score).

Table 7.8 : Frequencies of Discrepancies (N = 325)

No. of discreps.	Expected Matrix f_e	cum. f_e	cum. f_e (proportion)	Random cum. f_e. (proportion)	Year 1 ^ Year 2 f_o	cum. f_o	cum. f_o (proportion)	Year 2 ^ Year 3 f_o	cum. f_o	cum. f_o (proportion)	Year 1 ^ Year 3 f_o	cum. f_o	cum. f_o (proportion)
Col.	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)
0	13.5	13.5	.04	.04	36	36	.11	32	32	.09	28	28	.08
1	41.8	55.3	.17	.13	85	121	.37	70	102	.31	66	94	.29
2	13.5	68.8	.21	.17	22	143	.44	23	125	.38	25	119	.37
3	50.9	119.7	.37	.35	61	204	.63	54	179	.55	38	157	.48
4	35.0	154.7	.48	.44	26	230	.71	29	208	.64	28	185	.56
5	21.3	176.0	.54	.51	26	256	.79	20	228	.70	19	204	.62
6	29.3	205.3	.63	.60	18	274	.84	16	244	.75	24	228	.70
7	56.4	261.7	.81	.79	32	306	.94	36	280	.86	45	273	.84
8	13.5	275.2	.85	.85	4	310	.95	11	291	.89	9	282	.87
9	38.4	313.6	.96	.97	11	321	.99	22	313	.96	33	315	.97
10	11.3	324.9	1.00	1.00	4	325	1.00	12	325	1.00	10	325	1.00
Mean discrepancies				5.16	3.23			3.84			4.31		
Standard deviation				2.78	2.58			2.96			3.05		
Values of t (random ^ observed)					9.17			5.86			3.71		
Significance of t					$p < 0.001$			$p < 0.001$			$p < 0.001$		

Part B : Change in cognitive preference behaviour by the sample populations

It was reported in Part A above that while a substantial measure of stability of individual behaviour in the sample does occur it is not possible to calculate a confidence limit. This portion of Chapter 7 concerns cognitive preference behaviour of the whole population. The interest now lies in the extent to which the whole population changes on maturation, on the direction of that change and on the behaviour of subgroups within that sample population. Data from two sources was available. The first was the data which has already been used extensively in this chapter and it was generated by 325 respondents with three cognitive preference instruments which were used in successive years. Factor analysis of inter-mode correlation coefficients and cluster analysis of total mode scores was used. The same cognitive preference test (CPI-III) was also administered to three different bodies of respondents; one was at each of the three stages which were used with three different cognitive preference tests. These bodies had 161, 123 and 439 respondents in the 3rd, 4th and 5th year classes respectively and the two larger samples were reduced to the size of the smallest for the analysis which is reported in Table 7.9. The selection was random. (Analysis with all 723 respondents gave essentially similar results).

The mean mode scores in Table 7.9 are contradictory. In Part (I) of this table, the A mode mean scores are shown to increase over the three years in question and the P mode mean scores decrease. In both cases the changes are significant. The mean scores for the R and Q modes are confused. In Part (II) of the table the A mode mean scores decrease and the P mode increase while the changes to the R and Q modes are, again, non significant. The data in Part (I) concerns one population and three different cognitive preference instruments but in Part (II), the data from three similar bodies of students who were subjected to the same instrument are reported. It appears that the conflict stems from the disparity between the three tests particularly in view of the alteration that was made to CPI-III both to alleviate the faults reported by Brown and also to the adjustment that had to be made to allow for different numbers of items in the three tests, rather than to mismatching of the three samples of students. The students in the third of these three samples were from nine schools but identical results to the ones reported here were obtained when samples from the IIIrd, IVth and Vth Forms of the same school were used.

There is, therefore, some doubt about the changes which occur in cognitive preference behaviour if the data in Part (I) are accepted. The quality of the data in Part (II) are such that it dispels any doubt about the changes; the consistency of the mean scores and, hence, the large values for the ratio of variances is convincing evidence. The same instrument would, if administered to the same students at annual intervals, give results that closely resemble the results in Part (II).

The significant factor loadings for the respondents in 3rd year classes in the four factors analysis in <u>Table 7.10</u> are on Factors I and II and they are characteristic of the orthogonal bipolar axes reported by Kempa and Dubé (1973). The same students at the end of their 5th year classes generated mode totals which are largely loaded in similar manner on Factors III and IV. The loadings for scores from students at the end of their 4th year classes show a degree of transition; significant loadings on all four factors occur and they reflect the behaviour of both the preceding and following years. Together these loadings show that the population is in a state of flux.

Table 7.9 : Mean Cognitive Preference Scores (N = 325 and 369)

(I) Different Cognitive Preference Instruments administered to the same respondents

Level	CPI	R	A	Q	P
3rd Form	I	39.05(7.8	31.82(10.2)	33.22(7.4)	39.38(9.0)
4th Form	II	35.93(8.3)	35.14(8.9)	32.55(7.8)	40.43(8.9)
5th Form	III	39.68(9.3)	34.17(9.9)	34.17(8.9)	36.18(9.7)

Variance	Sum of Squares		Mean Square			
	Between Levels (df=2)	Within Levels (df=972)	Between Levels	Within Levels	F	p
R	2628.0	1279162	13140	1316.0	1.00	NS
A	1893.6	159319	947.0	163.9	5.78	0.01
Q	431.3	123502	215.6	127.1	1.70	NS
P	3181.8	175040	1590.9	180.1	8.83	0.01

Note: the number of items in CPI-II differed from the number in CPI-I, 28 and 24, respectively and the scoring procedure for CPI-III differed from the procedure used to score the other instruments. Allowance has been made for these differences in the data in this part of the table.

(II) The same Cognitive Preference Instrument (CPI-III) administered to administered to three different levels

Level	CPI	R	A	Q	P
3rd Form (N = 123)	III	18.95(8.47)	18.14(7.96)	6.28(8.60)	4.60(8.76)
4th Form (N = 123)	III	17.41(10.63)	15.38(9.21)	7.74(10.51)	7.37(9.51)
5th Form (N = 123)	III	16.69(8.68)	12.26(9.57)	8.73(9.25)	10.69(8.79)
F ratio		1.88(NS)	13.20 ($p < 0.01$)	2.06(NS)	13.94 ($p < 0.01$)

(The variance of this data is reported in Table 6.8).

Table 7.10: Two Factor Analyses of Inter Mode Correlation Coefficients with Three Cognitive Preference Tests. (N = 325)

Class	CPI	Preference Mode	Factor Loadings ($x 10^2$) I	II	III	IV	I	II
3rd	I	R	-30	76	20	-01	-15	71
		A	89	-06	02	05	65	-16
		Q	07	-81	-14	12	07	-68
		P	-82	05	-09	-12	-67	11
4th	II	R	13	78	-22	31	01	80
		A	63	-09	52	-02	76	-17
		Q	-14	-74	27	-34	00	-79
		P	-64	-12	-51	-02	-75	-03
5th	III	R	06	06	18	90	48	44
		A	14	02	84	20	69	09
		Q	-07	-17	-28	-82	-51	-49
		P	-16	05	-78	-35	-72	-09
Eigenvalues			3.56	2.73	1.50	1.13	3.56	2.73
Variance (cum.%age)			29.6	52.4	65.0	74.4	29.6	52.4

Note : Loadings > 0.168, $p < 0.01$ loadings> 0.152, $p < 0.01$

Section 7.5 : Conclusions

The data on mode reliability confirms that respondents to cognitive preference instruments do behave with a substantial measure of reliability. This claim is supported by the traditional measures of reliability as well as by the test : retest coefficients which though routine to most tests of intellectual ability are new in this field. The coefficients support the claim of the test designer that the respondents are aware of the structure of the instrument, albeit unconsciously, and they do express their preferences consistently. The suggestion that the administration of the test initially fashions or conditions preferences and then, in later items, reinforces the preferences is refuted but evidence was found to suggest that a few of the early items in an instrument do have less stabili.y than the body of the items in the remainder of the instrument.

The conclusion which must be drawn from all the evidence for mode stability, i.e. on stable behaviour in each mode independently over a span of years, is meagre. There are at least three possible explanations for this. It seems probable that the isolation of the individual mode scores from one another for the purpose of the analysis in Section 7.2 is ill-advised, or it is possible that the behaviour is unstable or, even, that the instruments which were used to measure it are insensitive, inaccurate or both. An analysis of the four mode scores together as a single relationship was therefore examined.

The data on mode consistency is based on two coefficients which are used here for the first time in cognitive preference studies. The coefficients which have been reported show that each individual continues to behave consistently when the four modes are examined together. It was shown in an earlier section that consistent behaviour is shown for each mode separately. The high level of consistent behaviour in each 'tetrad' of votes to each item is reported for the first time.

The attempt to demonstrate the stability of each individual's cognitive preference in the web of the whole population over a period of sufficient time for that web to show alteration was attempted by counting discrepancies and examining mean scores. The discrepancies showed that there was greater stability between the first and second year of the study (i.e. between class levels of IIIrd and IVth forms) than between the second and third years (i.e. IVth and Vth years). This is due, in part, to the fact that the students in the sample in the Vth form had chosen their fields for future study and this act of choice tended to polarise their cognitive preferences. This study is unusual in that it embraced students from the IIIrd forms with uncertain academic aspirations and followed them through to the Vth forms; previous studies have tended to examine respondents who had already embarked upon their preferred specialisations. The exploration of behaviour by counting discrepancies between one year and the next is novel and it is, apparently, a sound and reliable method of determining individual behaviour against the background of behaviour which might be predicted by mere chance or when random scores are used. The differences between mean observed scores and mean random scores are found to be highly significant and the actual behaviour revealed is markedly different from the behaviour anticipated by chance.

The mean scores in Table 7.9 furnish conflicting evidence but it is probable that the scores generated with the three samples of students at different class levels and using the same instrument is more reliable than the data obtained with the same students using three different instruments, one at the end of each of the last three years to 'O' level. The former body of data shows that the population as a whole does change significantly on the 'application' scale while there is little change on the 'questioning' scale. The significant change towards P at the expense of A suggests that this is an inevitable consequence of training and experience in chemistry in the classroom; the stability of the response to R and Q suggests that these preferences are manifestations of genetic make-up, i.e. that 'curiosity' is a property that cannot be as readily developed by teaching. This difference in behaviour on the two scales is a finding of some consequence in that it suggests that the teaching of chemistry with the avowed aim of increasing 'curiosity' is doomed. The second body of data does not support or refute these observations but the evidence is confused.

The cluster analysis demonstrates that when the shifts of behaviour which have just been described are annulled, 60% of the population moves in a progressive manner and only 6% fluctuate wildly. It must now be argued that this data lends support to the contention that cognitive preference behaviour should, in future, be established as a cognitive style. (In practice the act of annulling the shifts makes little difference and it serves merely to clarify cluster progression). There is no limit of confidence which can be placed on this observation of progression for justification of stylistic attributes but all the evidence amassed in this chapter, relating as it does to stability, to consistency and to the response being deemed a single measure, does lend profound support in favour of the treatment of cognitive preferences as a cognitive style.

CHAPTER 8 : THE JUDGEMENT EXERCISES

Section 8.0 : Introduction

The background for this work was sketched in Chapter 3 and the individual problems requiring elucidation were listed in Section 4.3. It has been explained that only two of the three broad problems of measuring judgement have been investigated. The first is the conceptualisation of what constitutes judgement and it does not lend itself to research. Once the notion of judgement has been established the notion can be researched. The second problem, then, is to produce an instrument of acceptable validity and the third is to expose a suitable population of students both to the instrument and to other, and more conventional, tests of abilities to discern the relationships between the new tests and the old (and proven) ones.

Section 8.1 is devoted to an investigation of validity using a very different instrument which is claimed to work in an allied field. The problems associated with the hypotheses which are integral to the notion of scientific judgement are explored in Section 8.2 and in Section 8.3 the measurements of judgement are investigated with six facets of school life.

Section 8.1 : Validity of the Judgement Exercise (JE) (Research Problem (D.1))

The JE claims to measure intellectual skills within the confines of a scientific discipline: it was hypothesised that specific previous knowledge should play no part in problem solving. The instrument will have *face validity* only if the scientific method is invoked and possible hypotheses examined in the light of general previous knowledge and experience. The face validity of the instrument was not questioned by the panel of qualified adults who examined the instrument. It is possible that for some respondents, particularly for some of the younger ones, difficulty was experienced with semantics. Spot checks revealed no problem with the words used for evaluation of hypotheses (adequate, relevant, justified) but these terms may have caused difficulty in operation. The face validity would also be diminished if respondents had experience of identical or closely allied situations. The quality of the experiments which they devised was of such low calibre that it seems unlikely that face validity was significantly impaired by previous knowledge or experience. (The experiments are reported in Section 9.5).

The skills which the JE claims to measure are defined clearly and it was thought by the designers of the instrument that the situations sampled the judgemental situations which a young chemist might encounter quite adequately but no frame of reference exists for complete confidence on this aspect of *content validity*. The test was examined by a panel of 21 adult chemists and much useful comment was produced but there was disagreement over the correct keying of items. The adults with wide experience had difficulty in recalling which experiences would be available to chemists with pre-O Level experience and conflict ensued between the knowledge of the adults and the projection into genuine judgemental situations of the adolescents. The inability to agree unquestionably right answers is a significant weakness which should not be overlooked. It was, however, minimised by a very thorough and careful pretest and by a number of pilot studies.

The JE was developed because there is no recognised instrument in this field. For this reason *concurrent* validity cannot be established. An attempt was made to see whether the measurements which are made by the Study Skills section of the Bristol Achievement Test have relevance to those made by the Judgement Test since both tests strive to assess intellectual skills in an essentially scientific form. Analysis of the data furnished by students on the BAT test and on the improved JE test (JE-III) is given in Table 8.1. Inspection of the figures shows that correlation coefficients between subtests of the two are weak and that factor

Table 8.1 : Means, Standard Deviations, Correlation Coefficients and Factor Analysis of scores on Bristol Achievement Subtests and JE-III subtests. (N = 364/IIIrd and Vth/1974 and 1976)

	Mean	S.D.	BAT-1	BAT-2	BAT-3	BAT-4	BAT-5	A	B	C	D	E	F
BAT-1 Props. of Materials	8.49	2.17											
BAT-2 Structures	11.39	1.96	35										
BAT-3 Sequences	7.25	1.63	42	31									
BAT-4 Explanations	6.65	1.59	32	24	35								
BAT-5 Interpretations	7.10	1.84	28	20	38	30							
JE-A Sensible) Theory	3.87	1.81	05	05	07	16	19						
JE-B Justified) Theory	4.70	2.08	10	19	18	16	27	23					
JE-C Adequate) Theory	2.77	1.84	10	04	05	05	02	02	06				
JE-D Relevant) Theory	4.05	1.86	01	08	20	12	12	09	16	16			
JE-E Beyond & Rel.) Lab	5.92	2.02	08	14	23	31	29	30	11	01	21		
JE-F Beyond & Irrel.) Lab	5.51	2.00	24	12	23	34	16	19	17	10	27	30	
JE-G Already Available) Lab	5.35	2.40	04	18	11	12	23	25	20	21	21	34	46

(r > .158, p < 0.01)

Factor Loadings (x 10^2)

	I	II	III	IV	V	VI
BAT-1 Props. of Materials	12	-76	22	07	-09	-16
BAT-2 Structures	40	-41	25	47	-12	-25
BAT-3 Sequences	01	-72	-06	15	29	-09
BAT-4 Explanations	19	-67	-29	-11	-03	16
BAT-5 Interpretations	10	-56	-43	32	03	18
JE-A Sensible) Theory	18	09	-73	28	-06	-10
JE-B Justified) Theory	04	09	-18	82	13	09
JE-C Adequate) Theory	16	04	11	05	06	92
JE-D Relevant) Theory	18	-05	-03	10	93	07
JE-E Beyond & Rel) Lab	34	-26	-63	-13	18	-08
JE-F Beyond & Irrel) Lab	74	-25	-15	-08	20	05
JE-G Already Available) Lab	76	06	-25	21	06	20

(Loadings > .193, p < 0.01)

	I	II	III	IV	V	VI
Eigenvalues	3.107	1.609	1.110	0.986	0.862	0.842
Cumul. Percent.	25.9	39.3	48.5	56.8	63.9	71.0

loadings largely distinguish JE subtests (Factors I and II - VI) from BAT subtests (Factor II). It was anticipated that from the evident similarity between the descriptions of the two tests (see Chapter 4) that some common measurement would be obtained. Thus both tests explore relevance as a scientific parameter and both are demanding some measure of judgement.

Another method of approaching the data was to use the BAT scores as a standard to which the JE scores can be related. (The thorough standardisation of the BAT with a wide range of populations is reported in the BAT manual.) The scores obtained by the population which worked both the BAT test and the earlier from of the JE test were examined for analysis of variance and it was found that the 'between group' variance of JE scores when subdivided into ten groups by BAT score was significantly greater than the 'within group' variance. The chief source of variance is found to be in the groups with poor scores on both tests while those who performed best on the BAT did not also perform with distinction in the JE-II. The explanation for this difference in behaviour may be related to the fact that the BAT level 5 test is designed for students with younger chronological age than those in the sample population and there was therefore little distinction between the two test for those of moderate ability. Subsequent analysis of the JE-II data revealed weaknesses in facility which may have tended to suppress the true distinction between the best and the moderates. It is postulated on the evidence given that while the total score in the JE does show the variance anticipated, the individual subscores do not reveal any meaningful behaviour that is common to the subtest scores of the BAT. It must be admitted that this negative finding is a disappointment since it would have been useful to have a nationally recognised instrument with standardised scores to serve as a frame of reference for the new test which is being reported here. (A shorter version of the Bristol Test using only items with good statistics was prepared and used in later investigations. See Section 8.3).

Table 8.2 : Variance of JE-II Score in Groups by BAT Standard Score (N = 421/IIIrd and IVth/1974 and 1975).

	BAT Score		JE-II Score	
	Mean	S.D.	Mean	S.D.
Group 1 N = 43	129.19	2.34	31.07	4.98
2 42	124.52	1.01	32.19	5.12
3 42	121.79	0.64	30.52	5.10
4 42	119.74	0.58	28.60	6.05
5 42	118.17	0.53	29.52	5.20
6 42	116.45	0.62	30.26	5.98
7 42	114.00	0.72	29.67	4.08
8 42	111.02	0.86	27.07	4.93
9 42	107.12	1.47	26.62	4.34
10 42	99.69	3.56	24.36	4.26

Analysis of Variance

Source of Variance	Sum of Squares	df	Mean Square	F
Between Groups	1456	9	161.79	6.21
Within Groups	10731	411	26.11	
		420		

Significance : $p < 0.01$

Section 8.2 : The Quality of Respondents' Hypotheses (Research Problems (D.2) - (D.4))

It was theorised from the low scores achieved by respondents to JE-II that the quality of their hypotheses might be at fault. Further, it was postulated that covert exposition of these hypotheses could be obtained by inviting the respondents to devise experiments of their own which they thought would earn an 'E' grading (an experiment which "will lead to information beyond that already given, and this information is relevant to the problem stated.") The interest in this exercise was only to explore the parameters which the respondent considered to be worth investigation. This was the purpose of the third part of JE-III. It was scored independently of the other two parts of the JE. The respondents were offered five problem situations which were similar in principle to those which they had already experienced in Part II and they were instructed to devise no more than ten experiments that they considered to be worth the 'E' grading. In Part II of the JE the experiments were outlined for assessment by the respondent; in Part III the experiments were outlined by the respondent and their quality was assessed by the investigator. No parallel assessment was made by another adult so the findings which are reported must be tentative but the experience and discussion initiated by the construction of two editions of Part II ensures that the findings in Part III are well worth consideration.

The respondents were permitted to distribute their ten proposed experiments between the five situations in any pattern. Table 8.3 shows the responses of both third and fifth form respondents, classified by the selected parameter, to Problem J. This problem is reported here because it received most attention from the respondents. (It concerns the growth of a holophytic organism in aqueous manganese(II) sulphate solution. The growth arises through airborne infection. Other, and apparently similar, common laboratory reagents do not appear to provide a hospitable medium for the organism.)

The responses of 289 Vth Form students and 161 IIIrd Form students gave mean scores of 3.67 (SD = 2.12) and 2.37 (SD = 2.19) respectively. It was predictable that in both sets of data this test would have correlated strongly with the JE-Part II subtest scores but the test also correlated well with Test A of Part I (the 'helpful' and 'justified' statements). Correlation with the test of achievement in chemistry by the younger students was strong ($r = 0.362$) and it was also strong with the IQ scores (r = 0.393). It could be argued that strong correlations reflect willingness to co-operate with instructions as much as any other quality but it appeared that even those who showed by the number of suggested experiments that their dedication to the task was strong produced a high proportion of unsatisfactory responses. It is of interest that the percentage of 'E' responses in Table 8.3 (42.7% and 45.4% respectively) accurately reflects the overall success in the other two parts of JE-III.

The new information which is generated by this work is the low quality of the understanding of likely causes of simple scientific phenomena; it is remarkable that six intelligent and willing students in each of the samples which are quoted in Table 8.3 believe that the green deposit originates from interaction between the solution within the bottle and the glass from which the bottle was made. The modest scores achieved by many on the first two parts of the JE must be the result of many elementary misunderstandings of this type; those who were more successful were also able to suggest experiments that were helpful (i.e. 'E' type) in Part III. They were the brighter students in testsof chemical ability and they had the highest general ability as measured by IQ. (Between-group variances of chemical achievement and IQ when grouped by scores on the third part of the JE were significantly greater than within group variances.)

In many scientific situations alteration of one parameter may alter another through a chain of events. Thus in the candle experiment it is reasonable to expect that alteration to the chemical nature of the candle wax will alter the volume of water which ultimately enters the bell jar since the temperature of the flame is governed by the composition of the wax, the temperature determines the proportion of air expelled by heating, and the later contraction of this air on cooling causes the water to enter to adjust the differences in pressure when the flame is extinguished. If a respondent is asked whether experimentation with the chemical

Table 8.3 : Classified Responses to Problem Situation J in JE-III
(N = 161,160/IIIrd & Vth/1976)

Parameter Examined	Classification	% age IIIrd Form (217 responses) (N = 161	% age Vth Form (259 responses) N = 160
Analysis of the green material	E	20.3	15.0
Location - light	E	9.5	6.9
- dry	F	5.0	5.2
- disperse	G	1.1	1.2
- bottle sequence	G	1.1	0.6
Air - no oxygen) - no carbon (IV) oxide)	E	5.4	8.5
- remove stopper) - secure stopper)	G	1.5	0.7
Water - concentration) - purity)	E	2.3	8.1
Contamination by users	F		8.6
Surface area of solution	E	0.6	
Chemical nature of X	E	2.8	0.6
Other green solids	F	4.6	2.0
Microscopic nature of solid	E	1.8	
Bottle - chemical nature of glass	F	3.7	3.7
Shelf - height) - position) - paint on surface)	F	2.8	2.0
Temperature	E		6.3
Anion) Cation)	F		6.9
Tautological responses		13.8	6.9
Unclassifiable		23.7	16.8
		100.0	100.0
Responses Classified as 'E' type		42.7	45.4
'F'		16.1	28.4
'G'		3.7	8.8

Classification E the proposed experiment will lead to information that is 'beyond' the information that is already available and also relevant
F 'beyond' but irrelevant
G 'already available' information

composition of the wax, he must first decide whether there is a chain of scientific events linking the wax to the volume of water and if alteration of the former will modify the latter.

A new angle on students engaged in judgement tests was examined by making use of the experiments that had been suggested by former students of equal experience and ability. The responses to the third part of the JE were built into a new test, JE-IV, to see whether a relationship exists between the quality of judgement and the number of events in the 'cause and effect' chain. The outcome which was anticipated was that the greater the number of steps in the chain the lower would be the success rate for that item. It was then hypothesised that an introduction to the notion of the chain would enhance success since judgement in this exercise demands both an understanding of the science and an ability to assess relevance and importance.

JE-IV contained 70 items of this type. The test has been described in Section 4.5 Part (B) and it appears with the response key at Appendix G. The sample population was assigned at randomto one of four groups and each group was given a different response sheet. Copies of these also appear at Appendix G. The information on the response sheets proffered different strategies for problem solving. Group A was given only an example with answers, Group B was given the example and the notion of chained events was explained, Group C were asked to categorise each item as in Part II of JE-III (with the addition of a fourth category) and Group D received information on both the chaining and the categorisation. (A copy of the instructions is also given in Appendix G.) The sample had had no previous experience of any judgement test. No information was given on the number of correct, i.e. 'E' type, responses and no record was taken of whether or not the respondent made use of the information on the response sheet. Record was made of success on items keyed as E_1, E_2, and E_3. The subscript identified the number of steps which unified the problem with the situation. The mean scores and standard deviations of the four groups are shown in Table 8.4 and the data is analysed for variance in Table 8.5.

Table 8.4 : Mean Scores and Standard Deviations on three 'E' type categories of JE-IV (N = 141/III/1977)

	Groups				
	A No Strategy	B Chains	C Categ-ories	D Chains & Catego-ries	Total
No. of students	35	36	35	35	141
End of Year Chemistry Exam (%)	53.34(16.01)	54.89(14.28)	52.17(14.47)	51.86(15.11)	52.83(15.05)
JE-IV- E_1	5.91(1.86)	5.83(2.22)	5.03(2.24)	4.80(2.07)	5.40(2.16)
JE-IV- E_2*	6.46(2.18)	6.11(2.07)	4.74(2.27)	4.91(2.73)	5.56(2.44)
JE-IV- E_3	3.91(1.52)	4.39(1.78)	2.57(1.08)	3.37(1.62)	3.57(1.67)
Total	33.57(7.90)	33.00(8.56)	25.46(7.78)	26.34(8.04)	29.62(8.89)
Intelli-gence Quotient	120.83(10.06)	124.36(12.00)	122.77(9.09)	123.94(10.84)	122.99(10.65)

(* E_2 has 14 items, E_1 and E_3 have 10 each; max. score 70)

Table 8.5 : Analysis of Variance of the data in Table 8.4

Source of Vari-ance	Sum of Squares		df				F	p
	Between Groups	Within Groups	Between Groups	Within Groups	Between Groups	Within Groups		
Chem.	209	31728	3	137	69.8	231.6	0.30	(NS)
E_1	34	622	3	137	11.22	4.54	2.46	(NS)
E_2	77	761	3	137	25.66	5.55	4.62	0.01
E_3	64	327	3	137	21.33	2.39	8.92	0.01
JE-IV Total	1940	10494	3	137	646.67	76.60	4.72	0.01
I.Q.	265	15736	3	137	88.32	114.86	0.77	(NS)

The analysis of variance in Table 8.5 shows that the end of year scores in the Chemistry examination and I.Q. scores were not significant. This was expected since the students were randomly assigned to groups. The variance on the single-step chains (E_1) was also not significant whilst on the two-step and three-step chains (E_2 and E_3) the variance is significant and it is because of these two types of item that the total scores also show significant variance. The interesting conclusion which must be drawn from this information is that the introduction to the idea of the chain of events which may link the effect with the cause is sufficient to make noticeable difference to the group scores. The request to categorise the items into four types also helped students to select the E type responses more accurately (Group C) but Group D students who had the advantage of the idea of the chain and the information on categories of response fared less well than those in Group B who had had the introduction to the chain and no information on categories.

The essence of this work in this Section has been to show that the general quality of the understanding of cause and effect on fairly elementary chemical problem situations is surprisingly low but better scores are obtained if instruction is given on linked causes. The implication of this is that students would benefit considerably from classroom tests incorporating physical and chemical problems of the type developed here and their understanding of science in general and of chemistry in particular would be much improved by subdivision of major problems into component pieces. It is one of the aims of all teaching in science to expose students to challenging situations and it would appear that the Judgement Exercices do achieve on paper much that is difficult and laborious to achieve in the laboratory.

Section 8.3 : Measurement of Judgement (Research Problems (E.1)-(E.3))

The two previous sections were concerned to show that both the validity of the instruments and the intellectual challenges that they make are satisfactory and meaningful in a classroom situation. This section relates how other important data related to scores on the Judgement Exercises. In particular, six possible influences on the measurement of judgement will be considered. The sample populations furnished information on the improvement in total score over a period of time, on the relationship between judgement and achievement in chemistry as measured by the O Level grade, on the differences in success on the test by the sexes, and on the influence of scientific awareness, bias, intelligence, and intellectual maturation on the JE scores. When this data has been presented and analysed it should then be possible to examine the wider usefulness of the instrument in classroom situations.

(a) JE scores and class level (Research Problem E.2)

This data was obtained by administering JE-III to three samples of students, one at the end of the IIIrd Form year, the second comprising IVth Form students only and the third Vth Form students only.
So far as it was possible to gauge, the three samples were equal in every respect and the data reported here were obtained by random selection of individuals from a pool of over seven hundred students to give a sample with equal numbers in each of the class levels. The data in Table 8.6 show that in all scores there is an increase in ability with class level and that this increase is significant for all subtests except Subtests A, D and F. The parameters examined by these two subtests are 'justified and helpful' in Part I and 'irrelevance' in both Part I and Part II. (Tests D and F respectively). It can be inferred that understanding of what is relevant to the whole body of science does not develop significantly during the time span in question; it is unlikely to have developed in the preceding years

Table 8.6 : Means, Standard Deviations and Analysis of Variance of the JE-III subtest scores by three Class Levels
(N = 369/IIIrd, IVth, Vth/1976)

Subtest	Class Level		
	III	IV	V
A (sensible)	4.11 (1.60)	4.08 (1.53)	4.52 (1.78)
B (justified)	3.63 (1.88)	3.93 (2.03)	4.58 (2.04)
C (adequate)	2.55 (1.29)	2.55 (1.61)	3.08 (1.78)
D (relevant)	4.20 (1.76)	4.28 (1.83)	4.31 (1.73)
E (beyond and relevant)	5.94 (1.26)	6.05 (1.64)	6.51 (1.99)
F (beyond and irrelevant)	5.18 (1.79)	5.65 (2.03)	5.72 (1.92)
G (already available)	4.90 (2.22)	4.93 (2.41)	6.02 (2.44)
JE-III Total	30.51 (6.02)	31.63 (6.85)	34.71 (8.27)

(Maximum Scores : All Subtests : 10., Total : 70)

Analysis of Variance

	Sum of Squares		df		Mean Squares		F	p
	Between Classes	Within Classes	Between Classes	Within Classes	Between Classes	Within Classes		
A	15.49	993	2	366	7.74	2.71	2.85	(NS)
B	57.85	1102	2	366	28.92	3.99	7.25	01
C	23.00	1102	2	366	11.50	3.01	3.82	.05
D	1.00		2	366				(NS)
E	22.96	1199	2	366	11.48	3.28	3.50	.05
F	21.50	1355	2	366	10.75	3.70	2.90	(NS)
G	98.85	2049	2	366	49.42	5.60	8.82	.01
JE-III Total	1162	18658	2	366	580.85	50.98	11.39	.01

and it must be assumed to be an aspect that is developed by the 'A' Level course rather than the 'O' Level course. It is, however, reasonable to postulate that in so far as the students understand the implications of the exercise, they do all equally understand what is justified by the information given and what is not. The two subtests which show the most significant change with time are subtests B and G. Subtest B is concerned with an assessment in Part I of whether or not important information is needed but not supplied while Subtest G is asking for an assessment in Part II of whether the information which an experiment will furnish is already available or not. The subtests are similar in spirit and both are related closely to the student's background and experience in chemistry. It is therefore entirely reasonable that these two measures should show significant development with time and it is interesting that the time in question is now shown to be duration of the O Level course.

(b) JE Scores and O Level Grades in Chemistry (Research Problem E.1))

The JE was seen to be an instrument that measured traits which are distinct from the qualities which are required of an O Level candidate and high level correlation between O Level grades and JE scores was not expected. In the event the analysis of the Vth Form data showed it to be quite high (r = 0.341, N = 439) The data is further analysed in Table 8.8 by analysis of variance. Analysis of variance of subtest scores by grades showed that all the subtests with the exception of subtest C made a significant contribution/establishing a relationship between JE Score and O Level grade. A very similar result was obtained with IIIrd Form students when the test of achievement that was used was the end-of-year examination. In both sets of data the ability to assess 'adequacy' (subtest C) made no contribution and one must conclude that this skill develops later if at all. It is evident that mastery of a far wider body of data is required than is available to an O Level candidate to assess whether information is mere tautology and therefore useless,or whether it embodies some useful material.

Table 8.7 : Means, Standard Deviations and Analysis of Variance of JE-III Scores by O Level Grades (N = 280/Vth/1976)

	O Grade			
	I N = 70	II N = 70	III N = 70	IV (fail) N = 70
A	4.54 (1.70)	4.04 (1.83)	3.94 (1.72)	3.60 (1.82)
B	5.80 (1.72)	4.82 (2.09)	4.47 (2.08)	3.65 (1.78)
C	2.69 (1.90)	2.70 (1.92)	2.70 (1.62)	2.73 (1.51)
D	4.79 (1.89)	4.24 (1.83)	4.21 (1.69)	3.46 (1.51)
E	6.71 (1.91)	6.30 (1.68)	6.00 (1.97)	5.21 (2.22)
F	6.53 (1.87)	5.83 (2.14)	5.59 (1.77)	4.63 (2.05)
G	6.30 (2.20)	5.99 (2.46)	5.47 (2.09)	4.37 (2.43
JE Total	37.39 (7.38)	33.84 (7.52)	33.04 (11.14)	27.36 (7.21)

Analysis of Variance

	Sum of Squares		df		Mean Square			
Source of Variance	Between Grades	Within Grades	Between Grades	Within Grades	Between Grades	Within Grades	F	p
A	32	870	3	276	10.73	3.15	3.40	0.05
B	162	1025	3	276	54.00	3.71	14.54	0.01
C	-3		3	276				NS
D	458	438	3	276	152.73	1.59	96.24	0.01
E	78	1059	3	276	26.08	3.84	6.80	0.01
F	124	1070	3	276	41.43	3.88	10.69	0.01
G	144	1465	3	276	48.17	5.31	9.07	0.01
JE Total	3623	19878	3	276	1207.7	72.02	16.707	0.01

The temptation to 'play safe' in responding to subtest C items is paramount.

(c) Sex

It has been shown in Parts (a) and (b) above that JE scores are modified by class level and by achievement. Inspection of the scores obtained by boys and girls showed that their mean JE scores differed quite substantially (the mean scores were 31.17 and 35.40 respectively and the standard deviations were 7.75 and 8.01) but two-way analysis of variance of JE scores by sex and by academic achievement as measured by the four O Level categories used above failed to show that the girls' performance was significantly better than the boys'. (There were no/other significant differences between the sexes in any test in the test battery).

(d) Maturation (Research Problem (E.3))

There is , within one class level, a sample of students with chronological ages all within one year but with widely differing maturation levels. A simple test of maturation with obvious relevance to chemistry was used with a sample of 161 IIIrd Form students. (The test, the Piaget Combinatorial Test, is reported in Chapter 4). This sample was divided into five groups by the scores on the test and the JE scores were examined for variance.

Table 8.8 : Means, Standard Deviation and Analysis of Variance of JE Scores grouped into five groups by scores or Piaget Combinatorial Test (N = 161/IIIrd/1976)

	N=	PCT		JE	
		Mean	Stan. Dev.	Mean	Stan. Dev.
Group 1	33	7.00	0.00	28.00	5.18
2	32	6.22	0.41	30.88	5.43
3	32	5.22	0.41	28.91	5.89
4	32	4.06	0.66	30.72	6.01
5	32	2.47	0.66	32.53	6.06

Analysis of Variance

Source of Variance	Sum of Squares	df	Mean Square	F	p
Between Groups	380	4	95.09	2.82	0.05
Within Groups	5267	156	33.76		

The data given in Table 8.8 shows that the variance in JE scores in the five groups is significant ($p < 0.05$) but it also shows that the maturation test scores are skewed towards the less mature students and some caution should therefore be exercised. However, this result is interesting in that it suggests that the significant improvement in JE scores with class level (reported in Part (a) above) is attributable in part to increasing maturation.

It is reasonable to suggest that success in the JE is related to awareness and experience of scientific things and to global intelligence. It was not possible in this work to follow the change of scientific awareness and experience which inevitably comes with increasing class level but the behaviour within one sample of students in the 'one shot' study was used to explore the relationship between JE scores and scientific awareness and between JE scores and a measure of I.Q.

(e) Scientific Awareness and Bias

Table 8.9 (a) : Means, Standard Deviations and Analysis of Variance of JE Scores grouped into five groups by scores on the BRI test (N = 161/IIIrd/1976)

		BRI		JE	
		Mean	Stan. Deviation	Mean	Standard Deviation
Group 1	33	20.94	1.58	33.73	5.93
2	32	18.28	0.51	29.69	5.56
3	32	16.25	0.43	29.53	5.81
4	32	14.22	0.78	28.22	5.34
5	32	11.22	1.60	29.69	5.50

Analysis of Variance

Source of Variance	Sum of Squares	df	Mean Squares	F	p
Between Groups	537	4	134.25	4.10	0.01
Within Groups	5110	156	32.75		

It is apparent from the data given in Table 8.9 (a) that the variance in scores between the five groups formed by using the BRI test with IIIrd form students is significantly greater than the within group variance ($F = 4.10$, $p < 0.01$). It may be recalled from Chapter 4 that the BRI test was a test with 25 items with acceptable statistics from the Bristol Achievement Test, Study Skills, Level 5. The full test purports to measure scientific experience in five subtests and standard scores are available: the modified test did not distinguish the five subtests and it is unique to this study. It is claimed that the shortened form of the test is at least as effective in measuring scientific experience with this sample as the full test and, subject to that caveat, it is now evident that JE scores are in part dependent on scientific experience.

Table 8.9 (b) : Means, Standard Deviations and Analysis of Variance of JE scores of five groups by A Level Chemistry Bias (N = 439/Vth/1976)

Group	N=	A Level Chemistry Bias Chemistry ...	JE Mean	JE Standard Deviation
1	30	Most preferred A Level	34.67	6.84
2	30	2nd most preferred	38.03	7.06
3	30	3rd most preferred	34.60	8.33
4	30	No comment	31.43	7.44
5	30	Chem. among the three least preferred subjects	32.90	9.48
	150			

Analysis of Variance

Source of Variance	Sum of Squares	df	Mean Square	F	p
Between Groups	767.6	4	191.90	2.98	0.05
Within Groups	9339.4	145	64.40		

The mean JE scores in Table 8.9 (b) show that Vth Form students with a strong bias towards chemistry achieve significantly better scores than those with antipathy to chemistry. It is noteworthy that those who placed Chemistry second in their preferences for subjects for A Level study scored more strongly on the JE than those who placed Chemistry first; the explanation lies in the greater affinity for Physics among those that are in Group 2. The 'best' scientists (i.e. best in terms of the JE) tended to put Physics first, then Chemistry and then, probably, Mathematics. For this reason the finding in Table 8.9 (b) that bias does influence the performance, is to be expected and the significance of the F value ($p < 0.05$) is certainly weaker than it would be shown to be by a more sensitive test of attitude to chemistry.

(f) I.Q.

Table 8.10 : Means, Standard Deviations and Analysis of variance of JE Scores grouped into five groups by scores on the Otis α and β I.Q. Test (N = 161/IIIrd/1976)

	N	I.Q. Mean	I.Q. Standard Dev.	JE Mean	JE Standard Dev.
Group 1	33	138.00	4.23	31.94	6.46
2	32	127.47	1.97	31.53	5.88
3	32	120.59	1.89	31.53	5.13
4	32	113.69	2.38	29.81	5.11
5	32	103.72	5.35	26.22	4.94

Analysis of Variance

Source of Variance	Sum of Squares	df	Mean Squares	F	p
Between Groups	724	4	181.00	5.71	0.01
Within Groups	4943	156	31.69		

The instrument which was used to measure I.Q. was the Otis α and β instrument which comprises items which demand literacy/and numerical and geometrical reasoning. It would seem to have little obvious bearing on the JE but it has been shown in Table 8.10 that the relationship is quite a strong one. The inference is that the tests both measure intelligence albeit of different styles,but evidently those who have the one type are richly endowed with the other.

Summary of Section 8.3

The data which has been reported in this section shows that there are significant relationships between scores on the JE and class level, achievement in chemistry, maturation, scientific experience and I.Q. The attempt to show that the sexes scored significantly different scores failed though it was found that the girls in the sample did slightly better than the boys.

Section 8.4 : The Judgement Exercise related to Academic Achievement and other variables

It has been shown in Section 8.3 that the JE scores are related to success on the tests of academic achievement, scientific awareness, maturity and I.Q. The next stage in the analysis was to extend the investigation to embrace the possibility of joint influence of these facts of behaviour on the JE score. The justification for this particular piece of work stretches back to the notion touched on in Chapter 1 that the JE is probing into the intellectual behaviour termed *information transformation*. It was postulated at the onset of this work that a variety of intellectual qualities would have direct bearing on the transformation process and evidence for this was sought by including a variety of tests in the test battery. In all the work which follows in this section (and in Chapter 9) the success of each student in the test of academic achievement is considered with the JE score because both achievement and judgement necessitate transformation of chemical information. In the work which follows these two scores are considered together and in relation to a number of other parameters in 2 x 2 x 2 contingency tables. It was hoped that this approach would reveal more about judgement behaviour than the several tests revealed separately.(This approach is further developed in Chapter 9 where the third parameter is the *information perception* data.)

The data in Table 8.11 was obtained by rank ordering the sample of IIIrd form students on each of the three tests in question; the students were then assigned to the top or bottom half of the sample on each measurement and the contingencies were counted. It was found that in all three analyses the biggest cells were the two which contained either the top students or all three measurements on the bottom students in all three. In each of the analyses these two cells together accommodated about 47% of the sample: it is not unexpected that those who score well on two tests should also do well on the third.

The significance levels of the χ^2 value for each of the three tables in Table 8.11 were found to be 0.05, 0.10 and 0.01 for the awareness test, for the maturity test and for I.Q. respectively.

The data in these three tables show that there is marked division of the sample into three groups. The top quarter are successful on all three tests in each part of Table 8.11, they are the best students intellectually and they are located in all top groups. Similarly the bottom quarter are found in all three bottom groups and the remainder, which accounts for about 50% of the sample, are undifferentiated by this treatment of the data. The inference that is drawn is that success in the end-of-year examination in chemistry is independently related to both information transformation and to each of the three variables in Table 8.11 separately; those with success in both information transformation and in any one of these three succeed in CE, poor scores in both are associated with low CE marks, while success in one delineates the partially successful candidates in the chemistry examination. It is, therefore, suggested that these three intellectual attributes are in fact subsumed into the information transformation process; general classroom experience would confirm the validity of this observation.

Section 8.5 : Conclusions (The references in brackets are to Research Problems)

The validity of the Judgement Exercise was established by 23 adult assessors. They found the task of assigning the statements to the 'correct' category (i.e. to the category intended by the test constructors) difficult and extensive refinement was undertaken in the light of their comments. It was, however, generally agreed that the tasks which were imposed on the respondents conformed to the test design. The attempt to establish validity (D.1) by the use of an allied instrument (Bristol Achievement Test, Study Skills, Level 5) was not successful except in that there was some significant compatability between the two tests for less able students.

The investigation into the hypotheses which the respondents generated for themselves was successful (D.2). It was found that success in Parts I and II of JE-III correlated strongly with the score on Part III, i.e. those students who were able to reveal large numbers of correct and relevant hypotheses of their own in Part III also did well when the hypotheses did not require explicit statement in Parts I and II. In general, the sample population produced an unexpectedly large proportion (about 55%) of hypotheses which were worthless in terms of leading to correct answers in the JE-III but were often useful in indicating the nature of the misunderstandings which lead to incorrect answers in this test.

It was found that problems which required chaining of two, and of three, links in a sequence caused greater difficulty than single step problems (D.4). The facilities for these were found to be 0.54, 0.40 and 0.36 for single, double and triple step problems respectively. It was also found that even simple

Table 8.11 : 2 x 2 x 2 Contingency Tables of Academic Achievement Scores, Judgement Scores and one other variable (modified Bristol test, Piaget Combinatorial Test and Intelligence Quotient).
(N = 161/IIIrd/1976

		JE-Top		JE-Bottom	
(a) Modified Bristol Test					
		CE-top	CE-bottom	CE-top	CE-bottom
BRI-top	CE	52.14(11.66)	29.33(5.66)	46.34(6.82)	25.14(6.63)
	JE	(36)35.75(4.36)	(12)33.83(3.61)	(7)28.21(3.42)	(22)27.14(2.65)
	BRI	19.11(2.24)	19.08(1.31)	19.10(1.85)	19.32(1.73)
BRI-bottom	CE	47.76(7.28)	31.00(3.61)	44.12(5.74)	24.32(7.28)
	JE	(17)34.76(3.61)	(13)33.92(2.24)	(17)25.88(3.16)	(29)25.68(2.45)
	BRI	13.71(1.41)	12.77(2.09)	14.35(1.41)	13.39(2.24)
		$\chi^2 = 4.35$, df = 1, p = 0.05			
(b) Piaget Combinatorial Test					
PCT-top	CE	49.39(8.29)	28.36(5.07)	44.69(5.39)	24.00(7.28)
	JE	(18)34.94(4.35)	(11)34.72(3.32)	(13)25.62(3.46)	(35)26.25(2.45)
	PCT	6.00(0.59)	6.36(0.67)	6.92(0.28)	6.43(0.65)
PCT-bottom	CE	51.55(11.83)	31.46(4.12)	45.58(6.32)	25.75(6.02)
	JE	(34)35.71(4.12)	(13)33.23(2.65)	(12)25.00(2.83)	(16)26,69(2.83)
	PCT	3.26(1.08)	3.76(0.93)	3.92(1.16)	4.06(1.00)
		$\chi^2 = 2.73$, df = 1, p = 0.10			
(c) Intelligence Quotient					
IQ-top	CE	51.22(11.66)	29.38(5.06)	46.13(5.66)	27.15(7.06)
	JE	(37)35.70(4.24)	(13)33.54(2.45)	(15)25.53(2.65)	(13)26.46(2.24)
	IQ	131.68(7.42)	129.69(5.10)	132.33(7.55)	129.73(6.40)
IQ-bottom	CE	50.07(7.81)	31.08(4.36)	43.60(5.83)	23.94(6.48)
	JE	(15)35.07(4.00)	(12)34.25(3.32)	(10)25.00(3.87)	(37)26.51(2.65)
	IQ	115.67(4.12)	113.00(7.07)	111.50(5.39)	108.84(7.14)
		$\chi^2 = 10.95$, df = 1, p = 0.01			

Mean Values and standard deviation for IIIrd Form sample

CE	:	37.24 (14.17)
JE	:	30.19 (5.94)
BRI	:	16.21 (3.52)
PCT	:	5.01 (1.68)
IQ	:	120.60 (12.41)

Note CE : End of Year Examination in Chemistry training in the nature of the problems presented by the exercise produced significant improvement in the two- and in the three-step problems but made no difference in one-step problems (D.3).

Exploration of the relevance of class level (E.2) on the subtest scores of the JE revealed no significance for the A, D and F subtests and only significance at $p = 0.05$ for Subtest E. These four subtests together explore the scope aspect of the JE, i.e. they examine the science in the problem in relation to 'global' science. The other three subtests which can be bracketed under quality and are concerned with the immediate problem and the extent to which use is made of the data (Subtests B, C and G) had significant difference at $p = 0.01$, $p = 0.05$ and $p = 0.01$ levels respectively. The interesting and, perhaps unexpected, finding is that class level has little or no bearing on scope but a direct bearing on quality. It might be expected that progression through the three class levels would improve understanding of chemical situations and expose conflicts between the chemistry which is local to a problem and the whole body of scientific information. It is more predictable that discernment of quality, i.e. data use within a problem, should become more acute with training, and that Vth form students might be expected to perform better.

The relationship between achievement in chemistry (within one class level) as measured by O grade (E.1) and JE score is found to be significant for all subtests at $p = 0.01$ except for Subtest A where the significance is $p = 0.05$ and for Subtest C which is not significant. (Indeed, the O Level failures were better at assessing 'adequacy' than were any passing grade).

When the relationship between intellectual maturity and the success in the JE is investigated (E.3), the evidence is weak. Awareness of Science and I.Q. are equally unconvincing. If, however, the data is studied in terms of 2 x 2 x 2 contingency tables which embrace CE scores, JE scores and any one of these three measures, it is found that there is evidence of a relationship between these intellectual qualities and information transformation. No difference in behaviour was found for the two sexes.

CHAPTER 9 : INFORMATION PERCEPTION AND INFORMATION TRANSFORMATION

Section 9.0 : Introduction

The Introduction to Chapter 1 of this work is devoted to a three-stage sequence of classroom learning. The theoretical structure was advanced in order to set cognitive preference styles and judgement styles in a psychological framework. This framework of learning explored aspects of information reception and transformation. In Section 1.1 cognitive preference behaviour was interpreted in terms of information reception and perception styles and in the next section the process of judgement was connected with information transformation. Chapter 2 was used to further develop the cognitive preference aspect and Chapter 3 describes the concept of judgement in so far as it relates to this study.

No immediate connection can be established on theoretical grounds between information reception and perception on the one hand and information transformation on the other, except that the former must precede the latter. However, information perception modes do not lead, as their consequence, to particular information transformation modes. Indeed, it is quite possible that the two are quite independent, not only from a theoretical point of view (which may, of course, be the consequence of an inadequately developed theory), but also empirically. The present investigation allowed this issue to be examined, albeit in a limited and tentative manner.

It has been found possible on the basis of an analysis of cognitive preference scores to identify population sub-groups having distinctly different cognitive preference traits. The methods which are used are described below. If it is assumed that some relationship exists between cognitive preference styles (as a measure of information perception and reception styles) and judgement behaviour (as a measure of information transformation behaviour), the judgement scores of different cognitive preference sub-groups should show significant differences.

There are, however, four different ways of approaching the cognitive preference data. The first to examine the scores in the judgement exercise of groups which have been identified by the score of respondents to each of the four cognitive preference modes separately. In this analysis, which will be found in Section 9.1, the CP scores are presumed to be independent of one another. Secondly, the evidence which is reported in Chapter 6 shows that there are, in fact, dichotomous relationships between the R and Q modes and the A and P modes. In Section 9.2 judgement scores are examined in the light of these two dichotomies separately and together. Another approach is to consider the cognitive preference data as delineating one single trait for each individual. In principle, this means that all four scores are considered as one 'trait score'. In Section 9.3 judgement scores are related to the 'trait score'. The fourth approach is to subject the cognitive preference data to cluster analysis and then to examine the behaviour of each cluster on the JE. The merit of this last system is that no predetermined outcome is placed on either the structure of the cognitive preference data or the size of each cell. It is reported in Section 9.4 and the findings of all four approaches are summarised in Section 9.5.

Section 9.1 : CP Scores, as four separate modes, related to JE Scores

Two samples of data are reported in this section because they give different results. The first sample is analysed in Table 9.1. It was gathered from IIIrd Form students and it shows that when JE scores are partitioned by cognitive preference mode scores the between-group variance is significantly greater than the within-group variance for groups partitioned on their Q score but there is no significance in the variance of the other three mode scores. The JE scores are seen to increase with increasing Q score for the first three groups but not for the last two groups

Table 9.1 : JE Scores grouped into 5 groups by each CP Mode score separately, with analysis of variance (N = 161/IIIrd/1976)

Group Mode Score (mean and S.D.) and Group JE Score (mean and S.D.)

	N=	R	JE	A	JE	Q	JE	P	JE
1	33	28.58(2.36)	29.55(5.39)	28.70(4.03)	30.45(6.31)	19.76(4.47)	33.52(6.78)	17.61(4.82)	30.64(5.96)
2	32	23.69(1.40)	28.69(4.96)	22.25(1.32)	30.44(4.39)	9.09(1.61)	30.13(5.45)	8.97(1.51)	30.75(6.38)
3	32	19.97(1.07)	29.50(5.65)	18.78(1.02)	29.03(5.75)	5.09(1.28)	28.34(5.48)	5.13(0.96)	29.50(5.33)
4	32	14.31(1.67)	31.81(6.01)	14.59(1.85)	28.75(5.62)	0.72(1.61)	28.91(4.49)	0.75(1.50)	30.34(6.56)
5	32	6.59(3.54)	31.09(6.62)	6.91(4.25)	31.94(6.45)	-4.78(2.64)	29.66(5.45)	-6.19(3.13)	29.72(5.80)
	161								

Group Variance of JE Scores

.	Sum of Squares Between	Sum of Squares Within	Degree of Freedom Between	Degree of Freedom Within	Mean Square Between	Mean Square Within	F	p
R	210.9	5326.6	4	156	52.72	34.14	1.54	NS
A	210.9	5326.6	4	156	52.72	34.14	1.54	NS
Q	535.4	5024.9	4	156	133.8	32.21	4.15	0.01
P	39.74	5637.3	4	156	9.93	36.14	.27	NS

Table 9.2 : JE Scores grouped into 8 groups by each CP mode score separately, with analysis of variance (N = 439/V/1976)

Group Mode Score (mean and S.D.) and Group JE Score (mean and S.D.)

	N=	R	JE	A	JE	Q	JE	P	JE
1	55	30.40(3.30)	34.11(6.85)	25.25(2.91)	31.80(7.38)	25.[illegible](3.39)	36.49(8.34)	28.44(4.87)	37.69(7.08)
2	55	23.71(1.34)	32.51(7.00)	20.33(1.32)	30.18(8.73)	18.45(1.97)	32.64(8.51)	19.67(8.74)	32.18(8.21)
3	55	19.84(1.02)	30.56(7.27)	16.27(1.03)	31.09(6.56)	13.51(1.01)	32.62(15.53)	15.85(1.13)	33.51(7.58)
4	55	16.62(0.98)	32.04(7.82)	13.15(1.10)	31.64(7.20)	10.85(0.96)	32.02(8.52)	12.69(0.91)	33.02(15.82)
5	55	13.89(0.71)	33.76(15.71)	10.16(0.80)	29.00(7.89)	7.84(0.78)	31.40(8.19)	9.98(0.80)	30.89(7.68)
6	55	11.04(1.03)	32.24(8.73)	6.47(1.22)	34.22(15.21)	4.85(1.05)	31.25(7.20)	7.38(0.96)	30.95(7.95)
7	55	6.96(1.35)	32.15(9.08)	2.05(1.55)	36.45(6.94)	1.64(1.00)	31.58(6.86)	3.35(1.05)	29.84(6.53)
8	54	0.80(3.22)	34.26(7.91)	-6.61(3.92)	37.31(7.74)	-2.69(3.21)	33.83(6.83)	-2.48(4.55)	33.43(7.47)
	439								

Group Variance of JE Scores

.	Sum of Squares Between	Sum of Squares Within	Degree of Freedom Between	Degree of Freedom Within	Mean Square Between	Mean Square Within	F	p
R	608.70	37227.3	7	431	86.9	86.37	1.00	NS
A	3402.9	34447.1	7	431	486.1	79.92	6.08	0.01
Q	1162.2	36704.8	7	431	166.0	85.16	1.95	NS
P	2254.6	35458.4	7	431	322.08	82.27	3.91	0.01

and to decrease with increasing R score (this latter relationship has a correlation coefficient = -0.197). The interesting feature which arises from this data is that JE scores are, to some extent at least, distributed along the 'curiosity' (or R/Q) scale. It will be recalled that these students had not completed their O level course in chemistry.

The second sample comprises Vth Form students and the analysis of their data is reported in Table 9.2. The procedure used for the two analyses was the same but with the second group it was found that between-group variance on the JE are significantly greater than within-group variance when A and P mode scores are used to fashion the groups but not as previously for Q. Thus it is seen that when the students had completed their O Level courses in chemistry their behaviour on the JE is distributed along the 'approach' scale; those who score high P preference scores are more likely to do well on the JE than those who score high A preference scores.

When I.Q. scores were examined for variance between groups differences were non significant with all four modes.

Section 9.2 : CP Scores used dichotomously for the purpose of grouping JE scores

The data which is contained in Tables 9.3 and 9.4 for the IIIrd and Vth Form students respectively is predictable from the data in the previous section. When IIIrd Form students are examined it is found that the between-group variance in JE scores is significantly greater than the within-group variance along the 'curiosity' axis but not along the 'approach' axis. The position is reversed for the older boys in the Vth Form samples.

When two-way analysis of variance is performed using both the dichotomous axes with the Vth Form sample it shows that the interaction between the 'curiosity' and 'approach' scores is significant. The evidence is that those with high P and high Q scores (top RH corner) have scored best while those with low P and high Q (bottom RH corner) have scored least well.

Table 9.3 : JE Scores grouped into 5 groups by CP scores used dichotomously, with analysis of variance (N = 161/IIIrd/1976)

Group Mode Score (mean and S.D.) and Group JE Score (mean and S.D.)

	N=	(R-Q)	JE	(A-P)	JE
1	33	32.36(4.05)	30.52(5.68)	33.48(6.33)	30.82(5.92)
2	32	22.59(2.38)	28.31(4.29)	20.50(2.06)	28.84(4.50)
3	32	14.69(1.86)	29.00(5.81)	13.88(1.75)	30.91(6.14)
4	32	5.25(3.77)	30.16(5.76)	5.88(3.25)	28.72(6.20)
5	32	-11.09(10.32)	32.63(6.60)	-8.44(9.94)	31.31(5.87)
	161				

Group Variance of JE Scores

	Sum of Squares Between df=4	Within df=156	Mean Square Between	Within	F	p
(R-Q)	350.7	5186.8	87.68	32.25	2.63	0.05
(A-P)	196.3	5341.2	49.08	34.24	1.43	NS

Table 9.4 : JE Scores grouped into 10 Groups by CP scores used dichotomously, with analysis of variance (N = 439/Vth/1976)

Group Mode Difference Score (mean and S.D.) and Group JE Score (mean and S.D.)

	N=	(R-Q)	JE	(A-P)	JE
1	44	32.70(5.34)	34.02(6.63)	28.14(5.28)	32.73(8.01)
2	44	23.50(1.84)	33.05(6.76)	17.52(1.94)	30.86(7.91)
3	44	17.91(1.89)	31.61(6.99)	11.36(1.48)	31.00(7.28)
4	44	12.16(1.48)	30.41(6.81)	6.48(1.23)	30.50(6.30)
5	44	7.70(1.37)	32.00(8.76)	2.43(1.48)	29.25(7.23)
6	44	3.73(1.16)	32.55(17.41)	-1.93(1.39)	31.43(8.35)
7	44	-0.30(1.10)	32.30(7.97)	-6.43(1.24)	34.18(16.86)
8	44	-5.80(1.67)	30.95(8.41)	-11.25(1.84)	32.77(8.49)
9	44	-13.41(2.77)	34.18(8.68)	-19.75(3.13)	36.16(5.97)
10	43	-23.95(6.13)	36.00(7.89)	-36.07(7.73)	38.23(7.68)
	439				

Group Variance of Scores

	Sum of Squares Between	Within	Degrees of Freedom Between	Within	Mean Square Between	Within	F	p
(R-Q)	1093.8	36743	9	429	121.53	85.65	1.42	NS
(A-P)	3022.5	34814	9	429	335.83	81.15	4.14	0.01

Table 9.5 : Two-way analysis of variance of JE Scores partitioned by (R-Q) score and (A-P) score. (N = 439/Vth/1976)

Mean Score and S.D. of Cells

		(A-P) Score			Group Mean - Population Mean		
		Top	Middle	Bottom			
(R-Q) Score	Top	31.63(7.62)	32.44(6.33)	33.92(6.09)	-0.04	0.77	2.25
	Middle	33.45(7.99) N = 49	29.98(7.23) N = 48	30.54(6.60) N = 48	1.78	-1.69	-1.13
	Bottom	32.43(8.09) N = 49	31.79(7.54) N = 48	28.88(7.43) N = 48	0.76	0.12	-2.79

Analysis of Variance

Source of Variance	Sum of Squares	df	Mean Squares	F	p
(A-P) Scores	157.3	2	78.65	(1.47)	NS
(R-Q) Scores	214.0	2	106.9	(1.99)	NS
Interaction	656.6	4	164.1	3.06	.02
Within Cells	22865.0	426	53.67		
		434			

It is therefore apparent that the interaction is significant only because the 'approach' score has so much bearing on the JE score. The implications of the data in this section and in the previous section are discussed in Section 9.5.

The same two-way analysis of data was performed on JE scores and on I.Q. on the IIIrd Form sample and the interaction was found to be non significant for both. The non significance of the I.Q. score, which is a measure of general ability, suggests that groupings are homogeneous for traits other than the ones in question.

Section 9.3 : CP scores treated as a 'trait score'

The third method of handling CP Scores was to treat all four CP scores as though they portray a single cognitive preference 'trait'. The scores are, of course, ipsative i.e. they inter-related mathematically. The purpose was to break a sample of CP data into groups without making any condition about group size or cognitive preference structure. The method used was to rank order the CP scores and then to assign respondents to their 'trait' group by a top half/bottom half decision on each of the four scores. Thus a student with a higher than average preference for the R and A modes and a lower than average preference for the Q and P modes would be given a 'trait score' of 1122. The ipsative nature of the scores preclude 'trait scores' of 1111 and 2222 but all of the other possible groupings were found in the sample. The incidence of each 'trait score' is reported in Table 9.6, with the JE score both as a mean score (with S.D.) and as a Z-score.

Any four measurements can be related to one another as axes within a tetrahedron. If this structure is used for the four CP scales, and the tetrahedron is assumed to be regular, a 1222 'trait score' will occupy a position at the R apex of the tetrahedron, 2122 at the A apex, etc. 1112 will lie in the middle of the face which has RAQ at its corners. This is because these are the above average scores and only P is below average;

Table 9.6 : Mean JE scores, S.D., and Z-score of groups containing individuals with same CP 'trait' score. (N = 439/Vth/1976)

Group No.	CP 'Trait Score'	N	Mean JE Score	S.D.	JE Score as Z-Score (note 1)	Rank Order
2	1112	16	30.87	7.08	-0.20	10
3	1121	10	33.50	5.63	0.09	4
4	1122	90	31.71	7.30	-0.10	7
5	1211	18	35.35	9.29	0.29	2
6	1212	20	30.25	7.18	-0.26	12=
7	1221	46	33.93	6.19	0.13	3
8	1222	20	31.45	6.93	-0.13	8
9	2111	20	31.00	7.65	-0.18	9
10	2112	38	30.32	7.61	-0.26	12=
11	2121	32	28.94	7.88	-0.40	14
12	2122	13	33.31	7.78	0.07	5
13	2211	84	36.77	13.42	0.44	1
14	2212	22	30.36	9.01	-0.25	11
15	2221	10	33.10	9.02	0.04	6
		439				

Note (1): the Z-score was calculated with the population mean (32.70) and the population standard deviation (9.29).

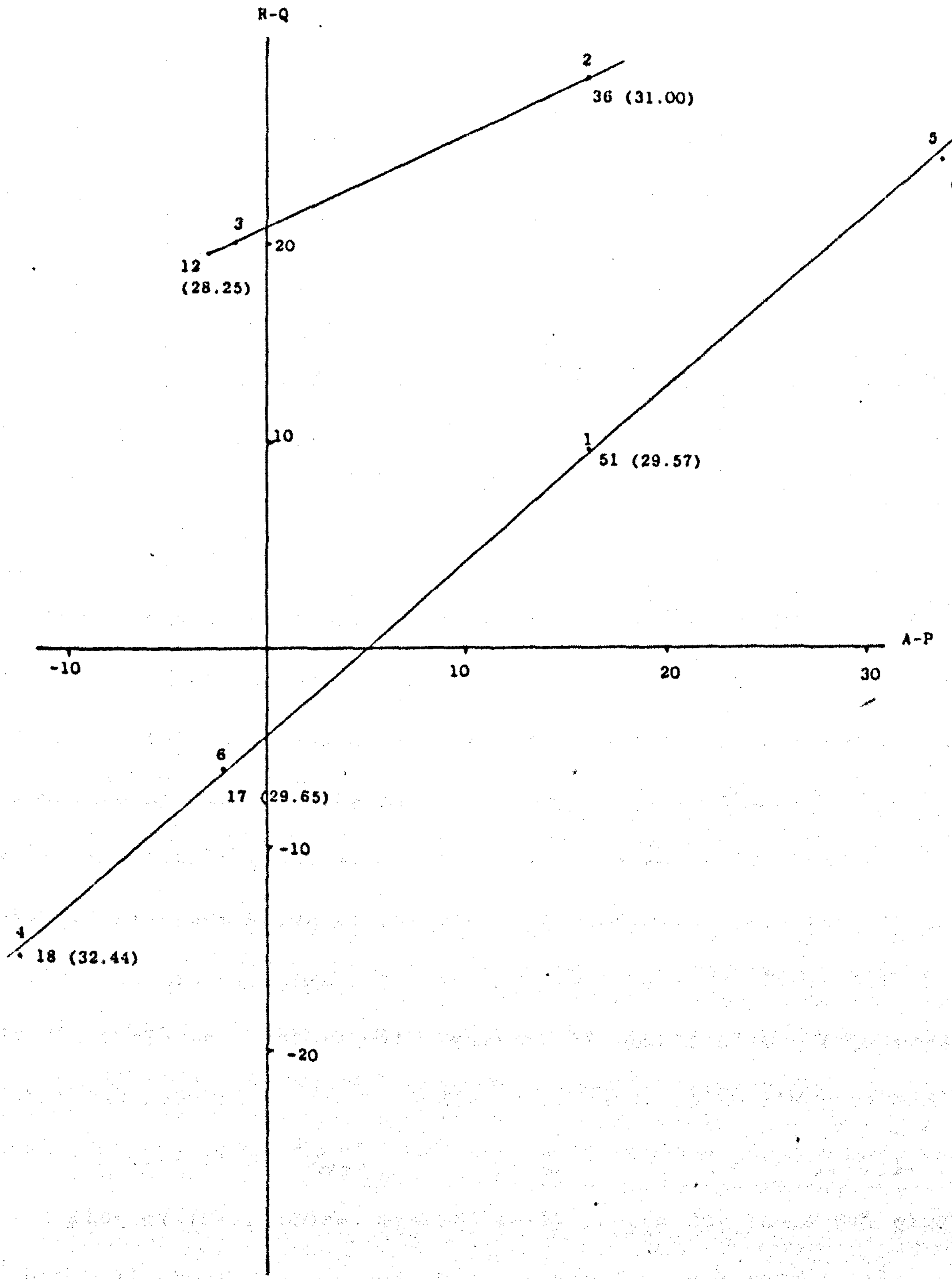
P is therefore below the face. 1212 has above average R and Q scores and below average A and P scores; it lies in the edge of the tetrahedron which joins the apex of R and the apex of Q. (Most statistical procedures strive to relate data to fewer axes without unacceptable loss of intensity; this process retains all the four axes. The approach which relates CP structure to two dichotomous axes which is used widely in this study, selects two edges of the tetrahedron and relates the data to them. The geometrical nature of a regular tetrahedron requires that the dichotomous axes shall be mutually perpendicular.)

The purpose of the tetrahedral axes in this instance is to provide a frame of reference for the 'trait scores'. Each group of respondents with the same 'trait score' has a mean JE score which, for convenience, is also presented as a Z-score. It is reasonable to ask whether the tetrahedron, if it is turned in the order given by the JE group scores after they have been rank ordered, turns in a regular manner and also whether this exercise has any useful purpose. In fact the first five in the rank order largely approximate to a turn round the P axis and the last six to a turn round the A axis. No doubt this could have been predicted from the evidence in the previous sections of this chapter but the exercise does prove to be valuable when a third variable is added to the two embraced so far, i.e. when academic achievement is added to the CP and JE data that is under discussion. This is reported in Chapter 10.

Section 9.4 : Cluster Analysis

Cluster analysis is another method of aggregating individuals into groups without predetermining the size of the groups. In using this method, the only restriction on group size was the necessity that the number of individuals in each group be sufficient to be statistically viable. It is unlike the 'trait score' method which was described in Section 9.3 in that there is extensive reduction of data.

Figure 9.1 : A plot of (R-Q) v. (A-P) Scores for the six clusters in Table 9.7 (N = 161/IIIrd/1976



Note: the figures in brackets are the mean JE scores of the six groups (see also Table 9.7).

Table 9.7 : JE scores of 6 Clusters. Clusters identified by cluster analysis of CP Data. (N = 161/IIIrd/1876)

Cluster	N=	R	A	Q	P	JE
1	51	16.90(4.67)	20.00(4.21)	7.45(3.58)	3.84(4.04)	29.57(6.54)
2	36	25.64(3.60)	20.75(3.26)	-2.83(3.56)	4.08(4.07)	31.00(6.57)
3	12	20.83(2.37)	12.50(3.75)	0.92(1.98)	13.83(3.08)	28.25(7.34)
4	18	2.33(4.67)	7.17(5.80)	17.89(7.27)	20.11(4.95)	32.44(6.25)
5	27	25.48(4.78)	27.55(6.03)	1.30(6.43)	-6.59(3.29)	29.07(5.94)
6	17	12.24(2.41)	7.59(3.74)	18.53(4.51)	9.71(3.04)	29.65(4.06)

The technique used previously in this chapter, of partitioning samples into groups of equal size is arbitrary. The di tribution (in Table 9.7) which was obtained by subjecting the four mode scores to cluster analysis, is found to be uneven; the largest group contains almost a third of the sample, the smallest has less than 10%. It must be admitted that the number of clusters which is selected for examination is also arbitrary: in this case six clusters were examined because an analysis with more than six includes clusters with unacceptably small size. (The seventh cluster was reduced to six by adding 2 students to 25 students and fashioning Cluster No. 5).

A plot of (R-Q) scores against (A-P) scores for these six clusters (Figure 9.1) shows Cluster Nos. 5, 1, 6 and 4 on a straight line (correlation coefficient, r, = 0.9999). These four clusters contain 70% of the sample and the remaining 30% is contained in two clusters (Nos. 2 and 3) which are found to lie at a distance of + 20 units from the line.

The data which was used for forming the clusters in Table 9.7 and Figure 9.1 was the four cognitive preference scores. The mean JE score of the students in the clusters on the line are found to progress from 29.07 for Group 5 to 32.44 for Group 4. (These mean JE scores were calculated by hand after the students had been assigned to clusters.) This is, therefore, both a powerful demonstration of the interaction between information perception and information transformation for the students on the line where high Q and P mode scores are associated with high JE scores, and an apparent inversion for 30% of the sample. The important conclusion is that a positive approach to any of the four modes is associated with gain on the JE; for 70% it is increasingly positive Q and P mode scores which lead to success, for 30% it is increasingly positive R and, particularly, A mode scores which lead to higher JE scores.

Section 9.5 : Conclusions

It has been shown in this chapter that JE scores are related to R and Q scores to (R-Q) scores for younger students (IIIrd Forms) and they are related to A and P scores and to (A-P) scores for older students (Vth Forms). Further, the interaction between (R-Q) and (A-P) is found to be significant for the Vth Formers but it is not significant for the IIIrd Formers.

The most interesting, but still tentative, conclusions which can be drawn arise from the comparison of these two bodies of data. In younger students it is curiosity which is dominant in those who have the greatest capacity for effective information transformation while in the older boys, with two years of additional experience and a structured course to O Level that was both complete and thoroughly revised, it is the interest in the principles of chemistry which distinguish those with the better performance on the JE. The conclusion which this leads to is that younger students are differentiated in their higher levels of intellectual thought along the curiosity scale but as they pursue their studies there is a change. They become less interested in the applications of chemistry and more interested in principles. The suggestion was made by King (1972) on the evidence gathered from 1st year VIth Form students that the 'curiosity' scale behaviour is not susceptible to alteration by different styles of teaching while the behaviour or the 'approach' may be modified by teacher traits. The new information neither supports nor refutes the suggestion. Now it is suggested that what is in fact altered depends to some extent on the stage that the student has reached in his examination programme. A profound preference for principles at the expense of application is retained at the end of the O Level course; the interest in the pursuit of aspects of chemistry that are tangential to the body of information in the syllabus, i.e. curiosity, is equivocal.

It is worth noting further that the data gathered by Kempa and Dubé (1971) and by King (1972) shows that there is abundant interest in the aspects defined as 'questioning' and it is now suggested that this recandescence of what is loosely described as curiosity is a property of classroom atmosphere arising out of the new terrain that was being explored at the time when the tests were undertaken. This introduced the new notion to the study of cognitive preference behaviour that respondents whose training on a syllabus is only half complete will behave differently from those whose training on that syllabus is complete, i.e. that in the 'O' Level - 'A' Level - degree sequence there will be evidence of cyclical behaviour. The alternative hypothesis that the cognitive preference is static throughout is untenable but research on different academic levels is difficult because numerous tests with several different instruments over a long span of time would be required.

The significance of the (R-Q), (A-P) interaction shows that there is a relationship between information perception and information transformation but the discussion hitherto has taken no account of other classroom attributes that might well be implicated. This omission is remedied in Chapter 10.

CHAPTER 10 : COGNITIVE PREFERENCES, JUDGEMENT BEHAVIOUR AND CLASSROOM LEARNING BEHAVIOUR

Section 10.0 : Introduction

In the broadest sense, research into cognitive styles has potentially promising applications. For, if it were possible to identify learners' cognitive styles in an adequate and reliable manner, educational (pedagogical) strategies could be so designed as to match learners' thinking and reasoning behaviours. The relationship between learning behaviour and students' cognitive styles has in recent years become an educational research field of increasing importance (see, for example, Kempa 1979). Research knowledge is still too scant to allow the matching process just referred to to be put into operation. Nevertheless, the educational potential in this area seems promising, especially with attempts to "individualise learning" more than is currently possible.

Although the "predictive" applications of cognitive styles research are still remote, interest in correlational research seeking to examine the relationships between cognitive styles and, e.g., achievement in studies, motivational traits, etc., is high. The main reason for this interest lies in an endeavour to identify correlates and possible determinants of success in the study of academic disciplines, other than conventional IQ measures.

The purpose of Chapter 9 was to establish the existence of a relationship between the two procedures for intellectual manipulation of information which have been identified in this research. Significant interaction between the information perception processes and information transformation has not been established. The burden of Sections 1 - 4 of this chapter is to relate these two procedures directly to classroom behaviour. The intention is to examine the relationship from two directions. The first approach is an examination of the interaction between information perception and information transformation on student performance and in the second, the implications of selected attitudes and qualities as well as established academic achievement on the interaction, are examined. It will be evident, then, that the interaction is used to expose the determinants of academic success. The data is examined broadly by correlation and subsequent factor analysis in Section 10.1, in Section 10.2 the first approach is used and Section 10.3 is devoted to the relationships between academic achievement and information transformation together on cognitive preference for each mode separately, for dichotomous behaviour and for all four modes as a single trait.

The data which will be used is the sample that was obtained from the Vth Form students. For some analyses the data has been reduced to give four academic achievement categories of uniform size. The data from all the students who scored Grade A at 'O' Level has been retained, and as these numbered 70, equal numbers of students were selected at random for the 'O' Level Grades B, C and 'failing' grades. (A 'failing' grade is a Grade D, E or F). Section 10.4 explores differences between the behaviour of younger students, i.e. IIIrd Forms, and the behaviour that has been demonstrated by Vth Form students with the intention of examining the extent to which information perception, information transformation and the interaction can be considered stylistic in nature. In the final section before the summary, the incidence of different behaviours within a student population and the bearing on their future performance is described.

Section 10.1 : Correlation and Factor Analysis

The data which is used for the statistical analyses in Table 10.1 comprised part of the Vth Form Sample: 70 of the students in the full sample were awarded 'A' grades at O Level and so 70 of the students who scored 'B', 'C' and 'failing' grades were randomly selected from the sample to make a total population of 280. (This restricted sample was also used for the analysis in Table 6.6).

Table 10.1 : Correlation and Factor Analysis of CP, JE and Other Data.
(N = 439/Vth/1976)

Means, Standard Deviations and Correlation Coefficients (x 10^2) (N = 280)

	Variables	Mean	S.D.							
1.	O Grade (see Note)	2.50	1.12							
2.	JE	32.84	8.36	41						
3.	Scientific Awareness	4.33	5.31	33	32					
4.	A Level Chemistry Bias	2.64	1.33	-54	-27	-57				
5.	R	15.25	9.15	-17	-04	-05	24			
6.	A	10.39	10.43	-35	-25	-31	33	23		
7.	Q	10.11	9.11	22	15	14	-27	-75	-49	
8.	P	12.31	10.04	31	18	24	-31	-48	-80	30
				1	2	3	4	5	6	7

($r > 0.16$, $p < 0.01$)

Note : the sign has been changed to facilitate comprehension

Rotated Factor Analysis (loadings x 10^2)

	Factor No.:	I	II	III	IV
	Variables				
1	O Grade	47	15	17	59
2	JE	-11	-00	-09	-93
3	Sci. Awar.	-84	05	-15	-13
4	Bias	86	20	13	17
5	R	05	93	18	-02
6	A	18	18	90	17
7	Q	-10	-89	-20	-12
8	P	-14	123	-91	-06
	Eigenvalues	3.309	1.540	0.963	0.799
	Cum. %age	41.4	60.6	72.6	82.6

(loadings > 0.20, $p < 0.01$)

The relationships which are revealed by the data in Table 10.1 are much as expected. The close bearing between the general measure of Scientific Awareness and the more specific one of Chemical Bias is shown by the loadings on Factor I. The (R-Q) and (A-P) loadings on Factors II and III respectively are expected from the dichotomous relationship studied in Chapter 6. The strong correlation between academic achievement and the JE score is also expected from earlier work (Section 8.3).

This data is given chiefly to show that the variables when examined separately do not serve to elucidate further our understanding of the nature of those variables. In subsequent sections they are examined in conjunction with one another.

Section 10.2 : Interaction between Perception of Information, Information Transformation and Academic Achievement by χ^2 Tests

The justification for the analyses which follow stems from the relationships which were demonstrated in Chapter 9. It was shown that with Vth Form data there was some increase in success in the JE with increasing P mode score and decreasing success with increasing A mode score. The relationship between the R and Q mode scores and JE scores were shown to have variance that was not significant.

In this section the significance of the difference in incidence between observed and forecast samples is tested by examination of χ^2. Table 10.2 shows this analysis for the Vth Form sample when divided into top and bottom groups with three variables. The first variable is one of the four cognitive preference scores, the second is the JE score and the third is the O Level grade which is taken as the measure of academic achievement. Table 10.3 is similar except that the cognitive preference score is one of the two 'dichotomous' scores and in Table 10.4 both 'dichotomous' scores are used and the contingency table has the dimensions 2 x 2 x 2 x 2. It was hoped that this analytical technique would reveal whether the relationship between cognitive preference and judgement has bearing on the quality of O Level grade.

It is reasonable to start from the assumption that there is no relationship between information perception and information transformation beyond the interaction which has been demonstrated already (Chapter 9). Table 10.2 shows that when the sample is divided into high achievers and low achievers there is no significance in the relationship between the R mode and JE scores and between the Q mode and JE scores. There is also no relationship of significance between the A modes and JE and between the P mode and JE when the low achievers are investigated but for the high achievers the relationship is strongly significant. If both the A mode score is low and the JE score is high, the high achievers are more numerous than expected; if either the JE score is low or the JE score is high the high achievers are less numerous than expected. The pattern for the high achievers and the P mode score is similar; if both P mode score is high and JE scores are high, the high achievers are more abundant than expected (note that it is low A mode scores or high P mode scores that are involved here), but if either high P mode or high JE score are obtained the number of high achievers is less than expected.

This evidence is substantiated when the analysis is done with the dichotomous scores (Table 10.3) but the effect described above is strengthened to include the low achievers also, thus both the high and the low achievers are found to be significantly more numerous when both the (A-P) score is low (because P is high and A is low) and the JE is higher than might be expected. There is some evidence, though it is weak, of similar behaviour in favour of Q at the expense of R in the (R-Q) portion of Table 10.3. The relationship becomes even more marked when both (R-Q) and (A-P) are used together (Table 10.4). In brief then, low A scores and high P scores together with high JE scores cause a larger number of high achievers than anticipated but if any one of these conditions is not fulfilled

Table 10.2 : 2 x 2 x 2 Contingency Tables derived from scores in the Cognitive Preference and Judgement Exercise and O grade (each measure was partitioned into top and bottom groups) (N = 439/Vth/1976)

	R		A		Q		P	
High Achievers	JE High	JE Low	JE High	JE Low	JE High	JE Low	JE High	JE Low
Cog. Pref. High	54	44	47	45	75	37	88	34
Cog. Pref. Low	76	39	81	40	55	47	42	51
χ^2 =	2.24 (NS)		4.84 (p = 0.05)		3.28 (NS)		14.95 (p = 0.01)	
Low Achievers								
Cog. Pref. High	49	68	43	81	43	58	39	54
Cog. Pref. Low	36	64	42	51	42	73	46	77
χ^2 =	0.55 (NS)		2.03 (NS)		0.59 (NS)		0.27 (NS)	
Mean Scores and S.D.s								
High Achievers								
O Grade	1.76(0.61)	2.07(0.54)	1.79(0.58)	2.07(0.53)	1.55(0.55)	1.92(0.59)	1.51(0.56)	2.03(0.57)
JE	39.48(3.29)	27.73(3.61)	39.66(3.32)	27.11(4.46)	40.47(4.66)	26.08(4.90)	40.15(4.28)	27.03(4.81)
Cog.Pref.	22.96(5.69)	23.05(5.08)	18.09(4.19)	18.27(4.28)	19.29(6.55)	17.84(6.36)	21.35(7.65)	19.00(6.32)
N=	(54)	(44)	(47)	(45)	(75)	(37)	(88)	(34)
O Grade	1.55(0.55)	1.95(0.60)	1.56(0.56)	1.95(0.59)	1.76(0.60)	2.11(0.52)	1.90(0.53)	2.00(0.56)
JE	39.92(4.60)	25.53(5.28)	40.35(4.29)	26.53(4.80)	38.80(2.95)	27.57(4.24)	38.74(3.44)	26.71(4.50)
Cog.Pref.	7.09(5.53)	8.51(5.42)	0.47(6.20)	4.45(6.42)	3.40(3.95)	2.74(4.59)	4.21(5.45)	5.61(4.12)
N=	(76)	(39)	(81)	(40)	(55)	(47)	(42)	(51)
Low Achievers								
O Grade	3.51(0.86)	3.97(1.18)	3.65(0.96)	4.05(1.22)	3.49(0.87)	4.22(1.20)	3.49(0.84)	4.24(1.23)
JE	38.27(3.57)	26.16(4.51)	38.51(3.74)	25.04(4.35)	38.19(3.09)	24.59(4.62)	38.46(3.26)	25.20(4.90)
Cog.Pref.	22.82(6.10)	22.22(5.00)	19.26(5.04)	19.28(5.14)	17.02(5.46)	14.55(3.89)	18.18(4.48)	16.76(3.99)
N=	(49)	(68)	(43)	(81)	(43)	(58)	(39)	(54)
O Grade	3.56(0.93)	4.27(1.24)	3.40(0.79)	4.22(1.21)	3.57(0.90)	4.04(1.22)	3.57(0.92)	4.04(1.20)
JE	38.47(3.30)	24.33(4.76)	38.19(3.14)	25.65(5.24)	38.52(3.79)	25.78(4.76)	38.26(3.62)	25.29(4.61)
Cog.Pref.	8.25(4.71)	9.52(4.39)	2.62(5.90)	7.00(4.16)	2.24(4.05)	3.29(3.97)	4.17(5.55)	4.48(5.04)
N =	(36)	(64)	(42)	(51)	(42)	(73)	(46)	(77)

Note: the figures in brackets denote numbers in each group

Table 10.3 : 2 x 2 x 2 Contingency Tables derived from scores in the Cognitive Preference and Judgement Exercises and O Grade using (R-Q) and (A-P) separately
(N = 439/Vth/1976)

	(R-Q)		(A-P)	
High Achievers	JE High	JE Low	JE High	JE Low
Cog.Pref.High	58	48	44	44
Cog.Pref.Low	71	36	84	41
χ^2 =	2.55 (NS)		5.67 (p = 0.02)	
Low Achievers				
Cog.Pref.High	41	68	44	85
Cog.Pref.Low	44	64	41	47
χ^2 =	0.11 (NS)		3.42 (0.05)	
Mean Scores and S.D.s				
High Achievers				
O Grade	1.72(0.61)	2.08(0.57)	1.84(0.56)	2.00(0.52)
JE	39.28(3.25)	27.65(3.89)	38.77(3.36)	26.95(4.39)
Cog.Pref.	17.97(9.64)	18.88(9.02)	13.48(8.95)	12.75(7.24)
N=	(58)	(48)	(44)	(44)
O Grade	1.56(0.55)	1.92(0.55)	1.54(0.57)	2.02(0.60)
JE	40.18(4.66)	25.64(5.27)	40.29(4.40)	26.71(4.88)
Cog.Pref.	-11.58(11.07)	-9.89(10.82)	-19.85(13.74)	-12.02(11.96)
N=	(71)	(36)	(84)	(41)
Low Achievers				
O Grade	3.56(0.91)	3.94(1.19)	3.57(0.89)	4.12(1.23)
JE	38.44(3.56)	25.84(4.52)	38.23(3.68)	25.26(4.58)
Cog.Pref.	20.93(8.87)	18.66(8.90)	13.89(9.26)	13.01(10.44)
N=	(41)	(68)	(44)	(85)
O Grade	3.50(0.87)	4.30(1.22)	3.49(0.89)	4.11(1.19)
JE	38.27(3.36)	24.67(4.86)	38.43(3.85)	25.30(4.98)
Cog.Pref.	-6.11(9.18)	-3.45(7.18)	-14.00(10.31)	-9.04(7.54)
N=	(44)	(64)	(41)	(47)

Note: the figures in brackets denote numbers in each group.

Table 10.4 : 2 x 2 x 2 x 2 Contingency Table derived from scores in the Cognitive Preference and Judgement Exercises and O grade, using (A-Q) and (A-P) together
(N = 430/Vth/1976)

Achievers	(A-P) High				(A-P) Low			
	High		Low		High		Low	
	JE High	JE Low	JE High	JE Low	JE High	JE Low	JE High	JE Low
(R-Q) High	31	25	30	48	26	23	11	20
(R-Q) Low	13	18	14	37	56	18	30	27
χ^2 =	0.95 (NS)		1.21 (NS)		5.80 (p = 0.02)		1.73 (NS)	
O	1.94(0.56)	1.96(0.54)	3.63(0.93)	4.00(1.19)	1.46(0.58)	2.22(0.60)	3.36(0.92)	3.80(1.20)
JE	38.74(3.30)	27.52(3.70)	38.27(3.75)	25.04(4.43)	39.88(3.20)	27.78(4.24)	38.91(3.30)	27.75(4.27)
$(R-Q)_1$	17.65(9.73)	19.64(9.03)	21.40(8.73)	18.83(9.05)	18.65(9.74)	18.04(9.35)	19.64(9.97)	18.25(8.74)
(A-P)	14.94(9.48)	13.44(6.89)	14.60(10.50)	16.71(10.79)	-15.31(13.36)	-8.83(8.15)	-14.45(9.14)	-10.35(8.02)
N=	(31)	(25)	(3)	(48)	(26)	(23)	(11)	(20)
O	1.62(0.51)	2.06(0.54)	3.43(0.85)	4.27(1.27)	1.55(0.56)	1.78(0.54)	3.53(0.88)	4.33(1.18)
JE	38.55(3.63)	25.94(5.30)	38.14(3.80)	25.54(4.75)	40.57(4.88)	25.33(5.52)	38.33(3.21)	23.48(4.84)
$(R-Q)_2$	-5.54(7.15)	-6.50(8.49)	-1.71(5.80)	-3.92(7.36)	-13.16(11.31)	-13.27(12.31)	-8.17(9.79)	-2.81(7.02)
(A-P)	10.00(6.53)	11.94(8.18)	12.36(6.38)	8.22(7.65)	-21.77(13.75)	-16.11(15.07)	-13.83(10.83)	-8.07(7.31)
N=	(13)	(18)	(14)	(37)	(56)	(18)	(30)	(27)

Note: the figures in brackets denote numbers in each group.

there are fewer high achievers than expected. The effect is far less marked among low achievers and the similar effect for R and Q modes is also much weaker.

One attribute which is common to high P/low A, high JE and high achieving subset is likely to be motivation but this cannot be the only explanation because low achievers with sufficiently high motivation to fall into the high P/low A and high JE subset do not show the relationship. Another possible quality which has a bearing is aptitudes and these will each be examined in the next section (Section 10.4). There is a large proportion of factual material integrated into the syllabuses of all O Level examinations and it is possible that it is mastery of this material which differentiates the low from the high achievers in this sample. It can then be argued that information perception and information transformation have relevance only if the factual detail is both sufficiently abundant and accurately known by the student. The implication is that the high achievers can pursue their appetite for the principles of chemistry and they can operate meaningfully in information transformation exercises; those who can do both these things and also do well at O Level are more numerous than those who do only one of them and yet also do well at O Level. Another possible facet of this work is the relevance of experience in chemistry. It is also interesting to reflect upon why there seems to be no relationship with curisotiy, i.e. the (R-Q) mode, a property that most successful chemists might be expected to display in abundance. It is also possible that experience and curiosity are related. Evidence to substantiate or refute this was sought and is reported in Section 10.5.

It is also of interest that an analysis of variance of O grade of nine cells formed by partitioning (R-Q) scores in top, middle and bottom groups and (A-P) scores divided similarly, give values for F which are found to be insignificant and there is no significant interaction. Yet it has just been shown that when JE scores are included and the relationship is sought for in terms of χ^2 values a limited relationship is found; limited because it is apparent for high and not for low achievers.

Section 10.3 : Interaction between Perception of Information, Information Transformation and Academic Achievement by Analysis of Variance and by 'Trait' Score

The approach to the data in this section differs from the approach in the previous section in that the Vth Form sample is first divided into four academic achievement cells of equal size and then each cell is divided into 'high JE' and 'Low JE' subcells. Mean mode scores were calculated and the significance of the variance between and within the eight cells was determined. The data is shown in Table 10.5 for each of the four modes separately and in Table 10.6 for the (R-Q) and (A-P) scores.

The data in Tables 6.6 and Table 10.5 show that the significant variance in cognitive preference scores which was observed with O Level grade cells of equal size is further strengthened when the analysis includes a division of each cell into high and low JE score. When dichotomous scores are used (Table 10.6) highly significant variance is obtained. The mean scores of the students who scored 'C' grades and 'Fail' grades are much more similar than the mean scores of the other academic achievement categories.

The inferences to be drawn from Tables 10.2 and 10.3 differ markedly from the information which is shown in Tables 10.5 and 10.6. This is remarkable in view of the fact that both analyses were conducted on Vth Form data. In the first analysis a χ^2 procedure was used with all the data; in the second analysis of variance was done with O Level grade cells of equal size. The χ^2 analysis showed significant differences with high achievers on A, P and (A-P) mode scores; the analysis of variance showed significance on all six cognitive preference scores even though low achievers tended to behave similarly. The divergence between the two methods of analysis is in degree only and it is probable that if a more uniform measure of academic achievement than the O Level grade had been used some of the distinction would have vanished. It is, however, true to say that if the sample is restricted to cells of equal size, the dichotomous behaviour is much accentuated.

Table 10.5 : Mean Cognitive Preference Scores, with Standard Deviations and Analysis of Variance by O Level Grade and by JE score (N = 439/Vth/1976)

O Grade	N=	R		A		Q		P	
		High JE	Low JE	High JE	Low JE	High JE	Low JE	High JE	Low JE
A	70	11.29(8.55)	12.82(8.41)	4.09(11.31)	4.02(10.47)	14.37(9.80)	12.74(8.08)	18.17(9.96)	18.34(10.94)
B	70	12.80(10.19)	17.71(9.30)	5.49(11.91)	14.77(10.79)	15.49(9.83)	5.91(9.62)	14.26(12.64)	9.83(9.72)
C	70	14.20(10.31)	21.37(8.97)	6.57(12.63)	20.83(10.49)	14.57(9.71)	0.60(9.71)	12.69(13.21)	4.83(9.63)
Fail	70	15.26(10.68)	16.57(9.27)	8.34(13.60)	19.00(10.07)	12.97(10.69)	4.20(8.98)	11.83(13.30)	8.54(8.84)
	280								

(N = 35 in each cell)

Analysis of Variance

Source of Variance	Sum of Squares Between Groups (df=7)	Sum of Squares Within Groups (df=272)	Mean Squares Between Groups	Mean Squares Within Groups	F	p
R	2589	20828	369.9	76.57	4.83	0.01
A	11386	19073	1626.6	70.12	23.20	0.01
Q	7877	22137	1125.3	81.39	13.83	0.01
P	5293	22916	756.2	84.25	8.97	0.01

Table 10.6 : Mean Dichotomous Cognitive Preference Score, with Standard Deviation and Analysis of Variance by O Level Grade and by JE score (N = 439/Vth/1976)

O Grade	N=	(R-Q)		(A-P)	
		High JE	Low JE	High JE	Low JE
A	70	-3.09(17.34)	0.09(15.46)	-14.09(20.48)	-14.31(20.53)
B	70	-2.69(19.07)	11.80(18.15)	-8.77(23.92)	4.94(19.78)
C	70	-0.37(19.20)	20.77(17.75)	-6.11(25.11)	16.00(19.28)
Fail	70	2.29(20.50)	12.37(16.73)	-3.49(26.20)	10.45(17.53)
	280				

Analysis of Variance

Source of Variance	Sum of Squares Between Groups (df=7)	Sum of Squares Within Groups (df=272)	Mean Squares Between Groups	Mean Squares Within Groups	F	p
(R-Q)	18687	62813	2669.8	230.9	11.56	0.01
(A-P)	31151	74099	4450.1	272.4	16.34	0.01

It will be recalled from Section 9.4 that a significant proportion (70%) of the sample (of IIIrd Form students) behaved in one way while a minority behaved in another when the analytical procedure which was used did not require the sample in cells of predetermined size. It will also be recalled from Section 9.3 that the 'trait score' procedure also permitted disparity of cell size. Table 10.7 is identical with Table 9.6 except in that the mean O grade (with standard deviation and Z-score) has been added. Correlation between the mean JE score and mean O grade of 14 'trait scores' gives a coefficient = -0.8695 but it is clear from the distribution of the 'trait' groups when plotted on mean O grade and mean JE score axes that two populations exist. Each population has 7 'trait' groups; those on Line 1 are nos. 6, 9, 4, 3, 5, 13 and 7 and for these the correlation coefficient is -0.9865; those on Line 2 are 11, 14, 2, 10, 8, 12 and 15 and for them the coefficient is -0.9692. The Line 1 groups contain 66% of the sample and they have, in general, higher R and P mode scores. There is, perhaps, only an artificial distinction between the two populations on the two lines in that any population can be split into two subpopulations with correlation coefficients higher than the original coefficient, but the division of the sample into two in both this analysis and in the cluster analysis is revealing. The conclusion to the cluster analysis was that positive response on any mode tended to be associated with better JE scores but the P and Q mode preferences favoured JE more frequently than R and A preferences. In the present situation where there are three variables (O grade, CP scores and JE scores), for any given O grade, positive response to R and P tends to be associated with better JE scores.

Section 10.4 : The Sample of Younger Students

The purpose of this section is to compare the behaviour of the IIIrd Form sample with the behaviour of the Vth Form sample which has already been described. The interest in this approach is to expose the changes which

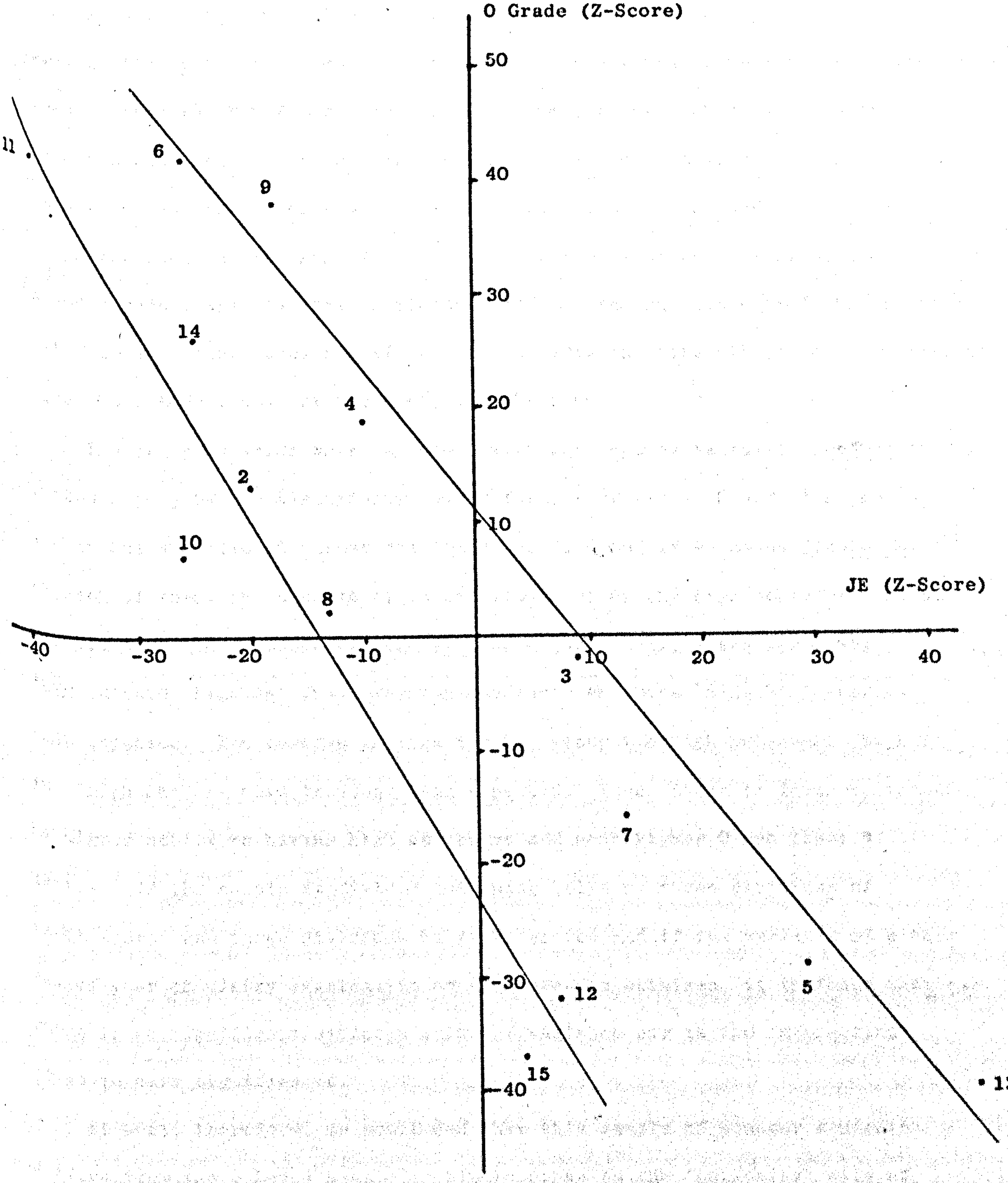
Table 10.7 : Mean JE scores, S.D., and Z-score of groups containing individuals with same CP 'trait' score. (N = 439/Vth/1976)

Group No.	CP 'Trait Score'	N	Mean JE Score	S.D.	JE Score as Z-Score (note 1)	Rank Order	Mean 'O' Grade	S.D.	'O' Grade as Z-score (note 2)
2	1112	16	30.87	7.08	-0.20	10	3.00	1.21	0.13
3	1121	10	33.50	5.63	0.09	4	2.80	1.54	-0.02
4	1122	90	31.71	7.30	-0.10	7	3.08	1.21	0.19
5	1211	18	35.35	9.29	0.29	2	2.41	1.37	-0.29
6	1212	20	30.25	7.18	-0.26	12=	3.40	1.50	0.42
7	1221	46	33.93	6.19	0.13	3	2.60	1.22	-0.16
8	1222	20	31.45	6.93	-0.13	8	2.85	1.28	0.02
9	2111	20	31.00	7.65	-0.18	9	3.35	1.68	0.38
10	2112	38	30.32	7.61	-0.26	12=	2.92	1.48	0.07
11	2121	32	28.94	7.88	-0.40	14	3.42	1.54	0.43
12	2122	13	33.31	7.78	0.07	5	2.38	1.08	-0.32
13	2211	84	36.77	13.42	0.44	1	2.26	1.23	-0.40
14	2212	22	30.36	9.01	-0.25	11	3.18	1.40	0.26
15	2221	10	33.10	9.02	0.04	6	2.30	1.49	-0.37
		439							

Note (1): the Z-score was calculated with the population mean (32.70) and the population standard deviation (9.29).

(2): the Z-score was calculated with the population mean (2.82) and the population standard deviation (1.39).

Figure 10.1 : Groups in Table 10.7 plotted on Mean O Grade and Mean JE Scores



occur during the O Level period.

Correlation and subsequent factor analysis of the tests which are displayed in Table 10.1 give coefficients and loadings which are essentially similar (Table 10.8). These younger students are found to be less scientifically 'aware' than the Vth Formers and this test correlates much less strongly with both academic achievement and the JE. Bias towards chemistry is weaker and this also has weaker correlation with the achievement test and with the JE. The difference in strength of correlation of the 'awareness' test on the A and P mode scores is worth noticing; the increasing 'awareness' that comes with age and class level significantly increases the strength of the correlation with the P mode behaviour and decreases the A mode.

The analyses which were performed with contingency tables in Section 10.2 are found to give non significant values for χ^2 for the A, P and (A-P) modes scores but significant values are found for R, Q and (R-Q) modes (Table 10.9). In each of these three cases it is the frequency of the high achievers which differs from the expected frequency; for R (and (R-Q)) there are fewer 'high R' and 'high JE' than expected and there are more 'high R' 'lower JE' than expected. The reverse is true for Q; among the high achievers 'high Q' and 'high JE' are more frequent than expected. Thus, there is further evidence of higher achievers having high JE scores and both higher Q and lower R scores. If the sample is divided into nine cells by three divisions of (R-Q) scores and three divisions by (A-P) scores and if the variance of either end-of-year chemistry examination or JE scores is examined, it is found that there is no significant difference in the variance nor is the interaction found to have significance.

It must, therefore, be concluded that this sample of younger students differs from the O Level sample in its response to the 'curiosity' axis of cognitive preference (R-Q) but it shows no differentiation on the 'approach' axis. It was suggested in Section 9.5 that this is a fundamental difference of some importance. It is a difference that is also sustained by the evidence in this section but the conviction is less positive.

Table 10.8 : Means, Standard Deviations and Correlation Coefficients of CP, JE and other data of younger students (N = 161/IIIrd/1976)

		Means	S.D.	Correlation Coefficients						
1	End of Year Chemistry Exam	37.24	14.17							
2	JE	30.19	5.94	43						
3	Scientific Awareness	6.52	3.84	12	-01					
4	A Level Chemistry Bias	3.60	1.28	-26	-09	-46				
5	R	18.47	8.44	-15	-20	-06	19			
6	A	18.14	8.10	-20	-07	-05	31	48		
7	Q	5.97	8.85	22	20	10	-26	-78	-64	
8	P	5.33	8.48	09	04	04	-23	-65	-76	37
				1	2	3	4	5	6	7

($r > .20$, $p < 0.01$)

Section 10.5 : The Distribution of Student Type within one Level

The next analysis was undertaken in order to explore the distribution of groups of students of different type within a sample population by cluster analysis. The IIIrd Form sample was used as thirteen items of data were available on each student. No preconceived number of clusters was envisaged and the analysis given in Table 10.10 was the one containingthe largest number of clusters which were all large enough to be viable. It was found that there were five clusters of approximately equal size and one which was much larger.

Cluster 1 contained 22 students. The mean scores on the academic tests are high, the Combinatorial test shows that they are advanced in their formal thinking, they have strong preference for the Q Mode and little preference for the R Mode and they have mild willingness to attend to Chemistry. Students with strong leaning towards Science and others with strong leaning to Arts subject were clustered together.

Table 10.9 : 2 x 2 x 2 Contingency Tables derived from R Mode, Q Mode and (R-Q) Mode Scores, Judgement Exercises and the End-of-year Chemistry Examination for younger students
(N = 161/IIIrd/1976)

	R		Q		(R-Q)	
High Achievers	JE High	JE Low	JE High	JE Low	JE High	JE Low
Cog.Pref. High	19	15	32	8	20	18
Cog.Pref. Low	32	10	19	17	31	7
χ^2 =	3.51 (.06)		5.18 (.03)		7.21 (.01)	
Low Achievers						
Cog.Pref. High	15	28	12	25	12	27
Cog.Pref. Low	10	23	12	27	13	24
χ^2 =	0.76 (NS)		0.00 (NS)		0.54 (NS)	
Mean Scores and S.D.s						
High Achievers						
CE	48.68(10.54)	45.47(6.22)	51.81(10.67)	45.54(7.76)	48.40(10.07)	29.61(5.27)
JE	33.47(3.76)	25.87(3.23)	36.19(4.34)	28.76(3.32)	33.95(3.41)	33.22(3.14)
Cog.Pref.	26.32(3.92)	24.13(3.76)	15.72(6.53)	14.92(7.01)	27.90(7.52)	24.22(7.12)
N=	(19)	(15)	(32)	(8)	(20)	(18)
CE	52.13(10.76)	44.60(5.34)	48.58(10.31)	30.88(4.44)	52.68(10.78)	31.34(5.28)
JE	36.34(4.02)	24.50(3.31)	34.00(3.50)	33.53(3.12)	36.32(4.27)	32.02(3.16)
Cog.Pref.	11.50(5.95)	10.20(4.26)	-1.47(4.66)	-0.66(4.05)	-3.48(11.69)	-1.20(5.31)
N=	(32)	(10)	(19)	(17)	(31)	(2)
Low Achievers						
CE	30.07(5.16)	24.82(7.31)	46.92(4.96)	24.04(6.19)	44.00(6.12)	24.70(7.43)
JE	33.13(3.14)	26.21(2.53)	25.17(3.43)	25.92(2.75)	25.25(3.28)	26.26(2.57)
Cog.Pref.	25.07(3.47)	25.14(3.14)	12.58(7.12)	10.20(5.19)	24.33(6.97)	25.37(6.14)
N=	(15)	(28)	(12)	(25)	(12)	(27)
CE	30.40(4.27)	24.78(6.01)	42.58(5.53)	25.04(7.51)	46.15(5.49)	24.92(5.91)
JE	35.00(2.45)	26.43(2.95)	25.75(3.19)	26.67(2.60)	25.38(3.38)	26.38(2.90)
Cog.Pref.	14.30(6.31)	8.78(15.08)	-0.33(3.23)	-0.63(3.43)	1.54(11.59)	3.88(9.57)
N=	(10)	(23)	(12)	(27)	(13)	(24)

Table 10.10: Cluster Analysis of Third Form Students (N = 161/III/1976). (Mean Score, Z Score (x 10^2) and Diagnostic Quality (Note (2)

Measurement	Mean & S.D.	(1) (N = 22)		(2) (N = 21)		(3) (N = 63)		(4) (N = 20)		(5) (N = 15)		(6) (N = 20)	
Scientific Aptitude	16.21 (3.52)	18.27	59k	16.90	20i	15.69	-15m	18.00	51k	15.47	-21d	13.60	-74m
Number Facility	22.32 (5.62)	27.13	86l	28.24	105d	20.84	-26c	22.45	02l	19.93	-43f	17.15	-92j
Space Facility	34.98 (8.67)	39.18	48c	35.62	07j	36.05	12j	39.45	52i	34.93	00k	21.90	-151f
English Comp.	12.71 (3.15)	15.18	78a	14.00	41h	11.46	-40g	15.50	89f	11.80	-29b	10.45	-72l
Judgement Ex.	30.19 (5.94)	37.27	119d	29.90	-05f	28.60	-27i	33.30	52g	28.53	-28m	25.85	-73k
Scientific Awareness	6.52 (3.84)	7.63	29m	4.19	-61m	7.02	13k	4.60	-50m	7.20	18a	7.55	27a
COGNITIVE PREFERENCE R	18.47 (8.44)	8.50	-118g	21.00	30e	22.12	43f	21.45	35c	7.73	-127j	20.30	22c
A	18.14 (8.10)	9.59	-105f	15.57	-32b	20.96	35a	28.25	125d	6.53	-143c	19.90	22h
Q	5.97 (8.85)	17.82	134b	2.05	-44g	2.49	-39b	1.50	-51h	17.40	129c	3.90	-23d
P	5.33 (8.48)	12.04	79i	8.67	39a	2.49	-33h	-3.50	-104a	16.53	132i	3.75	-19g
End-of-Year Chemistry Test	37.24 (14.17)	55.13	126e	39.81	18k	35.21	-14d	39.90	19b	34.93	-16l	20.35	-119e
Piaget Combinatorial Test (Note(1))	5.01 (1.68)	3.18	108j	5.66	-39c	5.32	-13l	3.65	81j	5.40	-23h	6.40	-83b
Intelligence Quotient	120.60 (12.41)	127.55	56h	128.81	66l	119.49	-09c	129.25	70c	114.7	-48g	103.60	-136i
% of Cluster opted for science 2 years later		41		24		52		35		60		25	
Convenient Label (Note(3))		Most Able		Maths/Arts		Mean		Strong Arts		Weak Scientists		Low Achievers	

Note (1) : Sign of Z score reversed for convenience
(2) : Diagnostic Quality is indicated by a letter : the letter 'a' indicates least variance with the cluster, the letter 'm' indicates most variance.
(3) : The Convenient label was based on behaviour 2 years subsequent to the administration of these tests.

Cluster 3 was the mean group. It contained 63 respondents and it was found to have undistinguished Z scores on all 13 data items. The low achievers were assigned to Cluster 6. These cognitive preferences were similar to those in the mean group but they fared much less well on the academic tests and they had the lowest I.Q.s.

Cluster 2 scored well on the IQ test and on numerical aptitude, they disliked science and had high P score but low A and high R but low Q. This group would appear to contain the mathematicians and some of the more able Arts students. Cluster 4 also disliked Science but did not score well on numerical aptitude. They had strong preference for A but not for P and some preference for R but not for Q. The few who subsequently opted to do chemistry for A Level all selected non-mathematical subjects to go with it (Biology, Geography, etc.). Cluster 5 was the smallest cluster and contained boys of poor quality in terms of academic tests but they had strong interest in science, and unlike Cluster 6 they had strong preferences for the Q and P Modes. The only cluster with either strong P/weak A and strong Q/weak R or strong A/weak P and strong R/weak Q was Cluster 2 but this was shown by subsequent behaviour to have a mixture of students: those with leaning for the Arts preferred P but rejected Q and were more numerous than those who preferred Sciences and who preferred both P and Q. It seems that these latter were rejected from Cluster 1 because of low JE and Combinatorial scores.

Section 10.6 : Conclusions

The Vth Form sample was tackled in two ways. In the first, the whole sample was divided into 2 x 2 x 2 contingency tables by behaviour in O Level, in the JE and in the six cognitive preference measures (R, A, Q, P, (R-Q) and (A-P)). It was found that for the A, P and (A-P) scores the incidence of students who scored well at O Level, scored a high mark in the JE and were above average in the preference for the P mode and/or below average in the preference for the A mode was significantly greater than expected. Closer examination of the mean scores shows that there are more 'high CP' and 'high JE' among 'high achievers' than might be expected and fewer 'high CP' and 'low JE' among 'high achievers'. It would be reasonable to expect a higher frequency of 'low CP' and 'low JE' among low achievers but this scoring was not observed. It seems therefore that the gain in the 'best' students is balanced by fewer high CP/low JE students rather than an unexpected increase in 'worst' students: in short there is a 'gain' for the best but no 'loss' for the worst. The evidence for interaction from the 2 x 2 x 2 x 2 contingency table using (R-Q), (A-P), JE and O grade for the four variables was positive but it was more muted than expected; again, the high scorers on all four measures were more abundant than chance alone would predict.

In the second approach the same sample was reduced in number to ensure the same number in each of the four academic achievement categories and cells fashioned from high/low JE division for variance of cognitive preference scores. Highly significant variance was found for all six cognitive preference mode scores; R and A were found to increase with less good A grades and Q and P were found to decrease. The finding that there was little distinction, if any, between the 'C' grade sample and the 'Fail' grade sample confirms the finding of distinction in the better half of the sample when tested by the χ^2 test. The fact that all six modes gave significant variance whereas only three were significant with χ^2 shows the effect of 'improving' the sample by artificial selection (i.e. random selection within each grade level). The firmer response to A, P and (A-P) in this sample was confirmed by both types of test but the distinction between these three mode scores and the other three (R, Q and (R-Q)) was more marked with χ^2.

The wholly different approach of treating the four cognitive preference scores as a single 'trait score' was examined. Its merit is that it does not impose uniform cell sizes, and examination of the O Level and JE performance of the fourteeen groups(each containing students with the same 'trait' score) showed that the sample could be divided into two portions with much the same results as those which were obtained by an entirely different technique in Chapter 9. The division of the sample by cluster analysis into more sub-samples using data from the whole battery of tests rather than with the cognitive preference data alonedoes not support the division by 'trait' score but the subsamples identified in this way (i.e. by cluster analysis) do have distinctive character and are rudimentary traces of a means of parcelling students into groups for career counselling and for other pedagogical ends.

CHAPTER 11 : CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

Section 11.0 : Introduction

The broad intention of this work was to coalesce current thought on learning processes and habits with work on cognitive styles. The Learning Model which was introduced was postulated as having a cognitive style of perception and a cognitive style of transformation. Evidence of the former in terms of cognitive preferences was found but the evidence of the latter in terms of qualities of judgement was inconclusive. New instruments were designed and developed; their quality and characteristics, as well as their behaviour with dependent, independent and affective variables, was examined and reported.

Section 11.1 : The Perception of Information

A large proportion of this work centred on the cognitive preference test since this instrument was considered to gauge perception of scientific information with accuracy. For the first time attention was paid to the construct validity of the instrument and the confidence placed in this type of testing was shown to be justified. A pair comparison procedure with interviews was used to establish motives for preference. Strong evidence of the existence of each of the four cognitive preference modes which were first described by Heath (1965) was found and there was no evidence of other modes. Some further evidence of construct validity was found by analysis of each of the items of a cognitive preference test separately and this novel technique exposed some weak items and some weaknesses of the instrument as a whole.

The statistical credentials of the instrument were examined. The several scoring procedures which have been used previously were compared and, since normative statistical procedures have recognised advantages much attention was paid to normative scoring and analysis. Comparisons showed that ipsative scoring facilitates measurement and yet gives results which compare closely with normative scores and analysis. The practice of using ipsative scores with normative procedures followed by cautious interpretation was shown to be justified.

The structure of cognitive preference behaviour was pursued and much evidence in support of previous workers was found. Both R-type and Q-type factor analysis and cluster analysis lent statistical support to the two dimensional approach which has been demonstrated in the past. Re-examination of the contrary finding on structure by Mackay showed that while his method was probably acceptable his findings were incomplete. Deeper investigation of his data showed that his structure was certainly no more acceptable than several of the structures which he rejected.

The attempt which was made to demonstrate the existence of a subject-specific component in cognitive preference behaviour failed but the technique which was employed (namely the factor analysis of normative data from items which contained only a single facet of an extensive syllabus) is worthy of further attention. There was convincing evidence of the relationship between academic achievement and cognitive preference behaviour; able O Level students had a strong preference for the P mode at the expense of A, i.e. they were differentiated principally on the 'application' scale, while able IIIrd Formers were differentiated on the 'curiosity' scale. The differentiation of girls at O Level on the 'application' scale was particularly marked. It is interesting, also, that the interaction of the (A-P) and (R-Q) scores academic achievement was shown to be significant. The relationship, among Vth Form students, in which scientific orientation is differentiated on the 'application' scale and the chemical bias is differentiated on the 'curiosity' scale, casts an unexpected light on the nature of these scales. The interest in chemistry as a subject for further study is related to the behaviour which is loosely described as curiosity while scientific orientation elicits a different response and one that is a function of the utility of scientific information. It was not anticipated that chemistry as a subject is separately identified from the whole mass of scientific information in the minds of the students in this way.

The relationship between cognitive preference and cognitive style was developed. Evidence was found which showed that cognitive preference is stable over each mode separately in each administration of a cognitive preference instrument and between the four modes together. A cognitive preference relationship is also stable over a period of years, i.e. of three years in this instance. This evidence was used in support of the claim that cognitive preference is stylistic in that it is an idiosyncratic and stable pattern of the perception of information and hence its acceptance as a cognitive style is justified. The support for the stability of any one mode over the period of years was not strong.

Section 11.2 : The Transformation of Information

The concept of judgement was related to critical thinking and it was seen as a high level of intellectual activity. It was defined in terms of the generation and subsequent classification of hypotheses and four specific attributes of good judgement were postulated and used in the classification process. The test instrument consisted of items in which chemical situations were described, a problem was posed and the statements and experiments for classification were given. Judgement was considered to be important in both theoretical and practical situations and both were incorporated in the test. The former with four levels of classification discriminated less well than the latter with only three levels of classification. An addendum to one edition of the exercise required the respondents to suggest experiments which would serve to elucidate the problem. The interest in the replies to this section centred chiefly on the actual parameter which the student considered sufficiently relevant to deserve further experiment. It may be recalled that the items in the Judgement Exercises endeavoured to explore situations which were wholly novel to the students. The evidence arising from the poor quality of the experiments which were suggested by the students showed that recall of the chemical situations in the Exercises, and of the scientific explanations of them, was not a contributory factor in test scores. In another experiment the facility was found to be related directly to the number of the hypotheses in the chain which united the problem with the suggested experiment. This was seen to be important since each link in the chain provides opportunity for the respondent to falter.

One of the principal weaknesses of the instruments which were used was the allocation of a single mark for the correct answer. It is suggested that, in future work in this field, the respondent should be required both to state his hypothesis and then to classify the statements included in the item. This would enable the quality of the hypothesis and the classification to be scored separately and some method of grading to distinguish gross errors from smaller ones could be adopted. Some lack of comprehension was expected with an unfamiliar task but the high level of comprehension which is required to cope successfully with the instructions to the test, the situations, the problems and the semantics of the classification procedure ensured a more taxing exercise than had been anticipated. The pre-test furnished warning of this problem and every attempt was made to mitigate it. An improved version might embrace demonstrations, films and pictures and as greater curiosity is generated by practical situations it might be wise to confine new situations to the laboratory. (Discussion of the nature and rationale of the test and its objectives prior to an administration would also serve to improve the efficacy of the instrument).

The question whether the JE presents the respondent with an exercise of divergent thought in a convergent disguise is interesting; the differences in JE behaviour of students previously classified at the extremes of the divergent/convergent spectrum would be revealing. Further, the relationships between quality of judgement and the well documented cognitive styles like breadth of categorising, impulsivity and levelling would probably be illuminating. The rôle which is played by the teacher in stimulating and improving judgement could be assessed by actively teaching some problems to one group and comparing subsequent behaviour in untaught problems with a matched group. It would also be interesting to establish whether the quality of scientific judgement of an Arts student reaches a peak at 'O' Level and then declines or whether it simply levels off. (It seems improbable that it might actually increase with no further training in science but increasing maturation might be the salient component.)

Section 11.3 : The Learning Model

The cognitive and affective hierarchies which have been proposed in the past inspired much valuable and illuminating research into learning processes. They have been valuable for assessment of effective teaching and for diagnostic investigations. This research has been concerned with the learning disciplines of a factual science and it was postulated that information on the cognitive styles of perception and transformation of material might similarly serve to improve both teaching and testing. The behavioural model which was proposed was thought to have three stages. The reception and perception stages were related to cognitive preference and the transformation stage to judgement. The intellectual activities were examined for relationships with the stage of intellectual development and understanding chemistry as distinct from the recall of chemical information. The Learning model was invoked to unify these disparate facets of the learning process.

In the past, research on cognitive preference has been confined to 'one-shot' studies. This longitudinal study shows that the population

BIBLIOGRAPHY

Abercrombie M.L.J., 1960, The Anatomy of Judgement, Hutchinson, London

Atwood, R.K., 1966, A Comparative Study of Achievement in CHEM Study, unpublished Ed.D.Thesis, Florida State University.

Atwood, R.K., 1969, Change in Cognitive Preferences of High School Physics Students, School Sci. & Mathematics, 69, 697-9.

Ausubel D.P., 1968, Education Psychology - A Cognitive View, Holt, Rinehart & Winston, Inc New York.

Barnett H.C., 1974, An Investigation of Relationships among Biology Achievement, Perception of Teacher Style, and Cognitive Preferences, J. of Res. in Science Teaching, 11,2, 141-147.

Bennett J.F. and Hays W.L., 1960, Multidimensional Unfolding: Determining the Dimensionality of Ranked Preference Data, Psychometrika, 25, 27-43.

Bieri J., Atkins A.L., Briar S., Lobeck L., Miller H., and Tripodi T., 1966, The Nature of Cognitive Structures, Clinical & Social Judgement, Wiley, New York.

Bignen E.J., 1973, Cluster Analysis, Tilburg University Press, Groningen.

Bloom B.S., 1956, Taxonomy of Educational Objectives, Longmans, New York.

Brimer A., 1969, Bristol Achievement Tests: Interpretive Manual, Nelson Ltd., London.

Brown S.A., 1974 unpublished Ph.D. thesis, University of Stirling.

Brown S.A., 1975, Cognitive Preferences in Science: Their Nature and Analysis, Studies in Sci. Educ., 2, 43-65.

Bruner J.S., Goodnow J.J., and Austin G.A., 1956, A Study of Thinking, Wiley and Sons Inc., New York.

Cattell R.B., 1957, Personality and Motivation: Structure and Measurement, New York, p.492-499.

Coombs, C.H., 1964, A Theory of Data, Chapters 6, 7 and 8, New York

Closs S.J., 1976, Ipsative vs. Normative Interpretation of Interest Test Scores, Bull. Br. Psychol. Soc., 29, 289-299.

Dubé G.E., 1971, A Study of Cognitive Preference Orientations and Scientific Interest Traits in Students subsequent to O Level Chemistry Courses, unpublished M.Sc. thesis, University of East Anglia.

Ennis R.H., 1962, A Concept of Critical Thinking, Harvard Educ. Rev., 32, 1, 81-111.

Everilt B., 1974, Cluster Analysis, Heinemann Educational Books Ltd., London.

Galloway C., 1976, Psychology for Learning and Teaching, McGraw Hill Book Co., New York.

Gagné R.R., 1970, The Conditions of Learning, Holt, Rinehart and Winston Inc., New York.

Gray J.L. and Knief L.M., 1975, The Relationship between Cognitive Style and School Achievement, J. of Exptl, Ed., 43, (4), 67-71, 1975.

Guildford J.P., 1954, Psychometric Methods, MGraw Book Co. Inc., New York.

Hays W.L. and Bennett J.F., 1961, Multidimensional Unfolding: Determining Configuration from Complete Rank Order Reference Data, Pschometrika, 26, 221-238.

Heath R.W., 1964, Curriculum, Cognition, and Education Measurement, Educ. and Measur., 24, 239-253.

Hicks L.E., 1970, Some Properties of Ipsative, Normative, and Forced-Choice Normative Measures, Psychol. Bull., 74, 167-84.

Kagan J., Moss H.A., and Siegel I.E., 1963, Psychological Significance of Styles of Conceptualisation, Soc. for Res. in Child Dev. Monograph, 28, (2), 73-124.

Kelly, G.A., 1955, The Psychology of Personal Constructs, Volume 1, Norton.

Kempa R.F., 1979, Plenary Lecture at the Cognitive Development Seminar, Leeds University (in press).

Kempa R.F. and Dubé G.E., 1973, Cognitive Preference Orientations in Students of Chemistry, Brit. J. of Educ. Psychology, 43, 279-288.

Kempa R.F. and Hodgson G.H., 1976, Levels of Concept Acquisition and Concept Maturation in Students of Chemistry, Br. J. Educ. Psychol., 46, 253-260

Kempa R.F. and McGough J.M., 1977, A Study of Attitudes towards Mathematics in Relation to Selected Student Characteristics.

King J.N., 1971, A Study of Pair Comparison as a Solution to the Psychometric Problems that arise from the Ipsative Nature of Cognitive Preference Tests, unpublished paper, Univ. of E. Anglia.

King J.N., 1972, Styles of Thinking and Curriculum Comparison, unpublished M.Sc. thesis, University of E. Anglia.

Kogan N., 1971, Education Implications of Cognitive Styles, Psychology and Educational Practice, Ed. Lesser G.S., Scott, Forman and Co., Illinois, USA.

Krathwohl D.R., Bloom B.S., Masia B.B., 1964, Taxonomy of Educational Objectives Handbook II : Affective Domain, London.

Leach P.J., 1967, A Critical Study of the Literature Concerning Rigidity, B. Journal of Soc. and Clin. Psych., 6, 11-22.

Mackay L.D., 1971, unpublished Ph.D. thesis, University of Monash.

Mackay L.D., 1972a, Cognitive Preference Tests, Schools Council Project for Evaluation of Science Teaching Methods, (draft).

Mackay L.D., 1972b, Changes in Cognitive Preferences During Two Years of Physics Study in Victorian Schools, Australian Sci. Teachers' Journal, 18, 63-6.

Marks R.L., 1965, Cognitive Preference Test: High School Chemistry, Indiana, Pennsylvania.

Marks R.L., 1967, CBA High School Chemistry and Concept Formation, J. of Chem. Educ., 44, 471-474.

McElwain D.W. and Keats J.A., 1961, Multi-dimensional Unfolding: Some Geometrical Calculations, Psychometrika, 26, 325-332.

Messick S., 1970, The Criterion Problems in the Evaluation of Instruction: Assessing Possible, Not Just Intended Outcomes, The Evaluation of Instruction, Issues and Problems, Withrock M.C. and Wiley D.E. (Eds), Holt, Rinehart, and Winston Inc., New York.

Nedelsky L., 1965, Science Teaching and Testing, Harcourt, Brace and World Inc., New York.

Peel E.A., 1960, The Pupil's Thinking, Oldbourne, London.

Peel E.A., 1971, The Nature of Adolescent Judgment, Staples Press, London.

Rogel A., 1974, Unpublished Master's thesis, Univ. of Jerusalem. (Quoted by Tamir (1975)).

Rowell J.A., and Hoffmann P.J., 1975, Group Tests for Distinguishing Formal from Concrete Thinkers, J. of Res. in Science Teaching, 12, (2), 157-164.

Schedemann G. and La Shier W.S., 1967-8, Cognitive Preferences of Students and Selected Characteristics of their PSSC Teachers, J. of Res. in Sci. Teaching, 5, 40-42

Scott N.C., 1973, Cognitive Style and Inquiry Strategy: a Five-Year Study, J. of Res. in Sci. T., 10, (4), 323-330.

Tamir P., 1975, The Relationship among Cognitive Preference, School Environment, Teachers' Curricular Bias, Curriculum, and Subject Matter, Amer. Educ. Res. J., 12, (3), 235-264.

Tamir P., 1976, The Relationship Between Achievement in Biology and Cognitive Preference Styles in High School Students, Br. J. Educ. Psychol., 46, 57-67.

Tamir P., 1977a, A Note on Cognitive Preferences in Science, Studies in Science Education, 4, 111-121.

Tamir P., 1977b, The Relationship between Cognitive Preferences of Students and their Teachers, Curric. Studies, 9, (1), 67-74.

Tamir P. and Kempa R.F., 1976, College Students' Cognitive Preferences in Science, J. of Educ. Res., 210-218.

Tamir P. and Kempa R.F., 1978, Cognitive Preference Styles Across Three Science Disciplines, Sci. Educ., 62, (2), 143-152.

Tamir P. and Lunetta V.N., 1977, A Comparison of Ipsative and Normative Procedures in the Study of Cognitive Preferences.

Tamir P. and Lunetta V.N., 1978, Cognitive Preferences in Biology of a Group of Talented High School Students. J. of Res. in Sci. Teaching, 15,1 59-64.

Tamir P. and Schiffman A., 1978, Development of a Medical Cognitive Preferences Inventory, (unpublished paper).

Trippi R.T. and Settle R.B., 1976, A Nonparametric Coefficient of Internal Consistency, Multi-variate Behavioural Research, October 1976, 419-424.

Vinton J.C., 1972, The Relationships among life style, task, and structure in university classrooms, unpublished doctoral dissertation, Case Western Research University, 1972.

Warr P.B., 1970, Introduction, Thought and Personality, Penguin Books Ltd., Hammondsworth, Middlesex.

Watkin H.A., Goodenough D.R. and Karp S.A., 1967, Stability of Cognitive Style from Childhood to Young Adulthood, J. of Personality and Soc. Psych., 7, 291-300.

Watson G. and Glaser E.M., 1964, Watson-Glaser Critical Thinking Appraisal, Form YM, Harcourt, Brace and World Inc., New York.

Ward J.H., 1963, Heirarchical Grouping to Optimize an Objective Function, J. Am. Statist. Ass. 58, 236-244.

Williams C., 1975, A Study of Cognitive Preferences, J. of Exp. Educ., 43, (3), 61-77.

Wish M., 1964, An Unfolding Analysis of Cognitive Preferences, Educational Testing Service, Princeton, USA.

Witkin H.A., Oltman P.K., Raskin, E., and Karp S.A., 1971, A Manual for the Embedded Figures Test, Consulting Psychologists Press Inc., Palo Alto, Calif. USA.

APPENDIX

A. Cognitive Preference Test - I

B. Cognitive Preference Test - II

C. Cognitive Preference Test - III

D. Judgement Exercises - II

E. Judgement Exercises - III (with Part III)

F. Bristol Achievement Test,Level 5, Study Skills

G. Judgement Exercise - IV (with response key and instructions to the four groups of respondents)

H. End-of-year Chemistry Examination - IIIrd Forms

I. 'BRI'

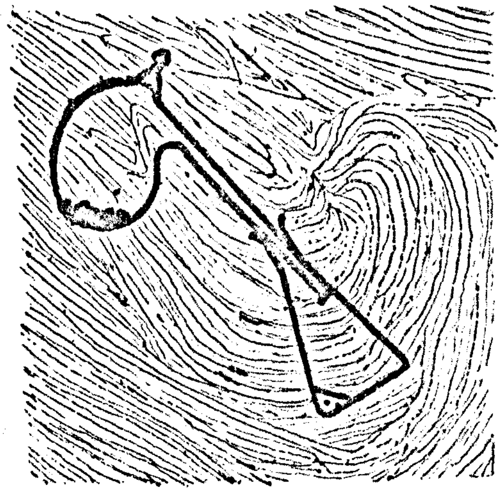
J. Science Orientation Questionnaire

K. Rotated Factor Analysis of Correlation Coefficients of Inter Item Scores

Appendix A

Cognitive Preference Test - I

Thinking about Chemistry 1



Every finger has its own print; similarly each of us has our own way of processing information. This booklet has been written with the intention of exploring fingerprints of the mind. The subject matter which has been used happens to be Chemistry but other factual subjects could have been used instead. THIS IS NOT A TEST; your way of handling data is not necessarily any better or worse than any one else's.

Each section has four statements and every statement in every one of the twentyfour sections is intended to be wholly CORRECT! You are simply asked to say how strong your liking is for each of the four parts by giving it a vote. This may be difficult to get used to because you are not usually asked about your feelings in Chemistry, but don't worry; just go through the booklet thoughtfully.

WHAT TO DO. Start by filling in the top part of the Response Sheet then read several sections to see what they are like. Next go back to the first Section and give

- three votes to the statement which you like most,
- two votes to the statement which you prefer next,
- one vote to the next,
- no votes to the statement which you like least.

Enter your votes in the boxes on the Response Sheet. Don't write anything in the booklet and don't give half votes.

Here is an EXAMPLE:-

Section 40. When coloured crystals are dropped into a solution of water glass, they behave in a curious manner.

(A) Their growth resembles plants growing in a tropical forest.

(B) An understanding of osmosis is necessary for an understanding of the behaviour of the crystals.

(C) The concentration of the water glass is important.

(D) Water glass can be used for preserving eggs.

A student who filled in the Response Sheet like this:

40	3	0	2	1

has shown the strongest liking for statement A, he gave no votes to statement B, two votes to statement C and one vote to statement D.

Work steadily through the booklet and try to finish just as the lesson ends.

June 1974.

J.N.KING.
Tonbridge.

1. A <u>candle flame</u> is both hot and yellow.

(A) Wax burns on the wick to give heat and light.

(B) Flames are the result of chemical reactions between gases which combine to release heat and light energy.

(C) A candle is a convenient source of light when there is a power cut.

(D) The flame is coloured yellow because it contains tiny pieces of hot carbon.

---

2. In winter, <u>salt</u> is spread on icy roads.

(A) The more salty a solution is, the more difficult it is to freeze it.

(B) Salt tends to make ice turn to water.

(C) The snag about salting roads is that extensive damage is done to the metalwork of cars.

(D) The use of salt keeps many roads open which might otherwise be too dangerous.

---

3. The farmers in Britain spend millions of pounds a year on <u>fertilisers</u>.

(A) When a crop is taken from a field, the chemicals which were needed to grow that crop must be returned to the soil.

(B) Fertilisers are generally balanced mixtures of the chemicals which growing plants need.

(C) Neglected soils can be made to yield good crops by careful use of artificial and natural fertilisers.

(D) The use of too much artificial fertiliser or the wrong one can lead to severe damage to the texture of the soil and to the pollution of neighbouring water.

---

4. Many elements and compounds can be obtained in crystalline form. They occur in many shapes and colours.

(A) Crystallisation is a useful method of purification.

(B) It would be interesting to know how the shape and colour of crystals are determined.

(C) A crystal is formed when the solvent is evaporated from a saturated solution.

(D) A crystal is a solid with flat surfaces at characteristic angles to one another.

---

5. Separation of the several dyes in ink is most easily done by chromatography.

(A) Chromatography relies on the fact that each chemical in a mixture can be 'bowled along' at a different speed by the moving fluid (generally a liquid).

(B) Liquids in a mixture of liquids like petrol can be separated by chromatography; they are moved along a tube filled with warm powder by a stream of gas.

(C) The coloured components of chlorophyll can be separated on filter paper with an alcohol solvent.

(D) Chromatography is used to check that only harmless dyes are used in fruit drinks.

---

6. The brown solid which forms on the surface of iron and steel is called rust.

(A) Iron does not rust in wet air which has first been treated to remove carbon (IV) oxide (carbon dioxide) so the explanation for rusting must account for the rôle played by this chemical.

(B) In rusting, iron is reverting to a similar chemical to the one in which it was found in the crust of the earth and energy is released in the process.

(C) Iron will only rust if it is both wet and exposed to the air; paint and grease serve as barriers and delay the process.

(D) In the past many millions of pounds a year have been spent on applying paint to structures which are made of iron, in order to prevent them from rusting.

---

7. Fireworks give a great deal of pleasure and they also have their uses.

(A) The distress signals which are used at sea are large fireworks.

(B) The fuse starts a chemical reaction inside the case of the firework.

(C) The different colours are produced by using the elements which impart colours to a flame.

(D) The firework contains at least two chemicals; one burns and gives out a great deal of heat and gas as it does so, and the other provides oxygen.

---

8. Fire often does a great deal of damage merely because the correct fire extinguisher was not used in time.

(A) Fire extinguishers are brightly coloured and they are placed in convenient and conspicuous places.

(B) Fire extinguishers are used wherever there is a risk of fire. The extinguisher is designed for the type of fire which is likely to occur.

(C) One type of fire extinguisher produces a rush of gas which blows out the flame, another type produces foam and gas, another uses a fine powder, and a fourth uses a blanket.

(D) If a flame is cooled below the temperature at which the gases in it will burn or if it is starved of oxygen, the flame will be extinguished.

---

9. Steel can be made into stainless steel by the addition of elements like chromium and nickel.

(A) The use of stainless steel for knives and razor blades is only fairly recent because the technique for applying sharp edges developed slowly.

(B) Stainless steel is a form of steel which does not rust because other elements besides the carbon which is normally present in steel have been added.

(C) It seems likely that other metals might be protected in the same way.

(D) The elements chromium and nickel both form thin protective layers on their surfaces and on the surface of the steel which protect them.

---

10. Pure water boils at 100°C on some days but on others it boils at temperatures which are slightly above or below this temperature.

(A) Climbers find that it is impossible to prepare hot drinks on mountains because the air pressure is low.

(B) The boiling point of a liquid is the temperature at which the liquid turns into vapour.

(C) The boiling point is reached when there is a stronger pressure of molecules wishing to escape than there is pressure on the surface of the liquid.

(D) The boiling point of an unknown but liquid may serve as a clue to its identity.

---

11. Laboratories are generally equipped with Bunsen burners.

(A) The burner is used in industry to heat fire bricks which then heat ovens etc.

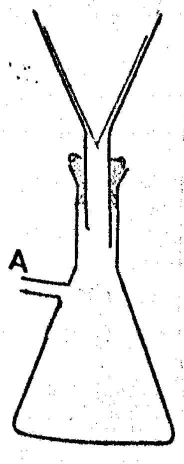
(B) The burner is designed so that the speed of gas coming up the pipe just equals the speed at which it can be burnt in the flame.

(C) The burner is an interesting one because air is drawn in to give a flame which is hot because it has two cones of burning gas.

(D) Gas and air are combined in a Bunsen burner to produce a safe, clean flame.

---

12. The separation of a solid from a liquid and solid is brought about by filtration. The apparatus shown in the diagram is suitable.



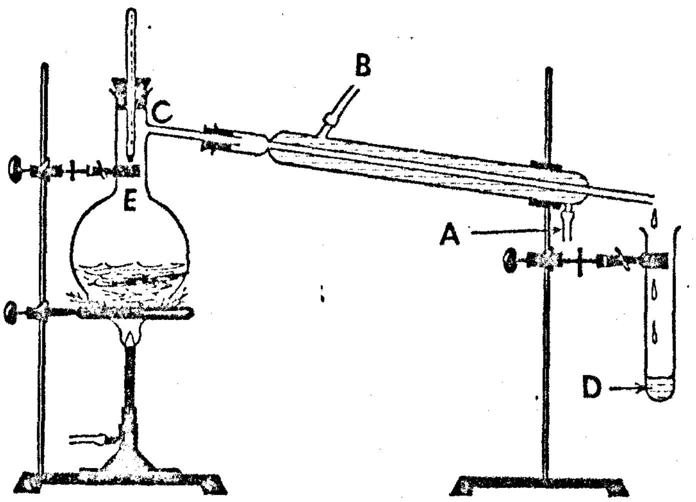
(A) A slow and tedious filtration can be speeded up by attaching a suction pump onto Tube (A).

(B) The fine holes in the filter paper allow the liquid through but hold back the grains of solid.

(C) Beer cannot be clarified by filtration because the particles are too small; precipitation is used instead.

(D) A solid may be separated from a liquid by filtration; the solid (called the residue) is retained by the filter paper, and the liquid (called the filtrate) is collected in the flask.

13. The Flask and Condenser.



(A) The enormous towers which stand beside electricity generating stations serve as condensers for steam.

(B) The vapour which rises to C goes into the side arm, is cooled and collects as a liquid at D.

(C) The condenser allows thermal energy to be transferred from the molecules of the vapour to the liquid which enters at A and leaves at B.

(D) Sometimes vapours which are difficult to condense may be made liquid by using the cold water in the condenser.

---

14. Water which has chlorine dissolved in it will function as a bleach.

(A) A chemical which turns a coloured material white is called a bleach.

(B) Oxygen—hating bacteria are quickly destroyed by dilute solutions of chlorine in water because the chlorine releases oxygen from the water.

(C) Chlorine is a powerful bleach for stained material.

(D) The bleaching action occurs because the chemical which is coloured is turned into another which is colourless.

---

15. Pollution of our environment by all types of waste is a world-wide problem of growing importance.

(A) The careful disposal of every type of waste is a major headache for all of us.

(B) It is noteworthy that no animal besides man causes pollution.

(C) The waste which causes polution includes solids, liquids and gases but it also includes heat.

(D) Pollution occurs because man has the ability to make rapid changes in his environment which Nature can only adjust to slowly.

---

16. Sodium Hydrogen carbonate is added to flour to make it into 'self raising' flour.

(A) Baking powder contains sodium hydrogen carbonate ($NaHCO_3$) which gives off carbon (IV) oxide (carbon dioxide) and steam when it is heated.

(B) It should be possible and it might be more efficient to produce bubbles in semi-liquid foods by injecting them with gas from a cylinder.

(C) A baking powder gives off a gas which makes bubbles in the food while it is still wet. Cooking prevents the holes from collapsing.

(D) Baking powders are used to make food lighter; they prevent scones from turning out like biscuits, for instance.

---

17. A reaction which is slow may benefit from the use of a catalyst.

(A) A catalyst provides an easier pathway for molecules to break the bonds and reform new ones.

(B) Nature's catalysts, called enzymes, are used in some washing powders to hasten the removal of stains.

(C) A catalyst may make a reaction go faster or it may make it go slower but the quantity of catalyst does not change.

(D) It is remarkable that the catalyst for a reaction which is reversible will catalyse the reaction in one direction exactly as efficiently as it will catalyse the reaction in the opposite direction.

18. Burning magnesium ribbon in a crucible is an experiment which is included at an early stage of courses in Chemistry.

(A) So much heat is produced by the reaction with oxygen that a magnesium also combines with the nitrogen in the air.

(B) Magnesium powder burns so brightly that it used to be used for indoor photography.

(C) Solid magnesium and oxygen gas combine to form solid magnesium oxide. The oxide is seen as a smoke and a white residue.

(D) The release of energy in the form of light and heat occurs because a chemical which is more stable is formed.

---

19. Crude oil is a mixture which can be separated by distillation.

(A) Separation of the components in crude oil into 'fractions' is possible because the components have different boiling points.

(B) Oil from the different oil fields round the world differ widely in composition.

(C) None of the products of crude oil are wasted.

(D) The separation of crude oil requires the use of columns of considerable height fitted with trays of 'bubble caps'.

---

20. Mercury (II) oxide (mercuric oxide) is an orange powder which is unstable to heat.

(A) Heating agitates the molecules and breaks the bonds between the mercury and the oxygen.

(B) This is an easy but expensive way of making small quantities of oxygen.

(C) There is no explanation for the dark red solid which is produced when the oxide is just heated.

(D) If the tube is heated oxygen is released and a film of mercury remains.

---

21. Drinks are made fizzy by pumping a gas into the liquid before the mouth of the bottle is sealed by the cap.

(A) Carbon (IV) oxide (carbon dioxide) is suitable because it is colourless, tasteless and not very soluble.

(B) Carbonisation of sweet, coloured liquids is used because it makes them more appetising to drink.

(C) The mass of gas which is dissolved by the liquid in the bottle depends on the pressure used during bottling.

(D) Carbon (IV) oxide (carbon dioxide) forms naturally in sparkling wines after the wines have been bottled and corked.

---

22. The protection of a less reactive metal (like iron) by a coating of another which is more reactive (like zinc) is an interesting way of preventing rusting.

(A) In tin plating the less reactive metal is put on the outside but when the rusting does eventually penetrate down to the iron beneath, the holes which are formed are small but deep.

(B) The protection of iron by zinc is commonplace; corrugated iron sheets for buildings, dustbins and wall nails are examples.

(C) The zinc serves to protect the iron underneath because the zinc is sacrificed before the iron starts to rust.

(D) Iron which is protected by a thin layer of zinc is said to have been 'galvanised'.

---

23. When crystals of copper sulphate are heated they lose the water which is chemically combined within them (i.e. they lose their water of crystallisation).

(A) The white form of copper sulphate may be used as a test for water since it turns blue when in contact with it.

(B) The white form of copper sulphate is a powder so it would seem that the water molecules are involved in the formation of the crystal.

(C) The crystal of copper sulphate is blue before heating, then turns a chalky white after heating.

(D) In crystalline copper sulphate the water is bound into the crystal by chemical bonds and it is therefore an important part of the structure of the crystal.

---

24. Air is a mixture of gases and it is also a most valuable raw material.

(A) Nitrogen is the most abundant and accounts for 4/5ths of the air, while the remaining 1/5th is mostly oxygen.

(B) Separation by physical means of the components in a mixture of gases is possible because each component behaves quite independently of all the others which are mixed with it.

(C) The oil and coal which man has burnt during the last fifty years has made little change to the proportion of oxygen and nitrogen in air.

(D) The removal of nitrogen from the air used in making steel saves money because it reduces the total bill for fuel.

Thinking about Chemistry 2

Some of you will be familiar with this exercise from last time and will know what to do. The purpose is to explore "fingerprints" of the mind. It is NOT A TEST in which some students do well and others do badly: your way of handling chemical information is not necessarily better or worse than anyone else's.

Each section has four statements and every statement in every one of the twentyeight sections is intended to be wholly CORRECT! You are simply asked to say how strong your liking is for each of the four parts by giving it a vote. This may be difficult to get used to because you are not usually asked about your feelings in Chemistry, but don't worry; just go through the booklet thoughtfully.

WHAT TO DO. Start by filling in the top part of the Response Sheet then read several sections to see what they are like. Next go back to the first Section and give

- three votes to the statement which you like most,
- two votes to the statement which you prefer next,
- one vote to the next,
- no votes to the statement which you like least.

Enter your votes in the boxes on the Response Sheet. Don't write anything in the booklet and don't give half votes.

Here is an EXAMPLE:-

> Section 40. When coloured crystals are dropped into a solution of water glass, they behave in a curious manner.
>
> (A) Their growth resembles plants growing in a tropical forest.
>
> (B) An understanding of osmosis is necessary for an understanding of the behaviour of the crystals.
>
> (C) The concentration of the water glass is important.
>
> (D) Water glass can be used for preserving eggs.

A student who filled in the Response Sheet like this:

40	3	0	2	1

has shown the strongest liking for statement A, he gave no votes to statement B, two votes to statement C and one vote to statement D.

Work steadily through the booklet and try to finish just as the lesson ends.

1. Solid carbon dioxide ('dry ice') changes directly into carbon dioxide gas without melting at -78°C.

(A) This is a physical process called sublimation.

(B) The sublimation temperature of carbon dioxide is low because the forces between CO_2 molecules in 'dry ice' are weak.

(C) Solid carbon dioxide is often used for keeping food-stuffs cool.

(D) It would be interesting to know how many other substances change from solid to gas without melting.

---

2. When a salt is dissolved in a solvent a point is reached when no more solute will dissolve in a given amount of solvent.

(A) In a saturated solution the solute molecules in the solid constantly interchange with the solute molecules in solution.

(B) When some of the solute remains undissolved in the solution, the solution is said to be saturated.

(C) The saturation point of salts in solution will vary considerably with temperature.

(D) A variety of useful salts may be prepared from seawater by evaporation of the water.

---

3. Acids are a most important class of chemicals.

(A) Any chemical which gives the H^+ ion or reacts with an alkali is, by simple definition, an acid.

(B) Acids release H^+ ions in solution.

(C) The annual consumption of sulphuric acid by a nation is a good measure of how well its industry is developed.

(D) No chemical will behave as an acid unless water or some similar liquid is present.

When crystals of <u>hydrated salts</u> are heated, they lose their water of crystallisation.

(A) Plaster of Paris hardens and so holds broken limbs in place when the anhydrous calcium sulphate recovers its water of crystallisation.

(B) Since anhydrous copper sulphate is non-crystalline, it would appear that the water molecules are responsible for the crystalline nature of hydrated copper sulphate.

(C) In crystalline copper sulphate the water is chemically bonded to the copper sulphate forming part of the crystal structure.

(D) Hydrated copper sulphate is blue before heating, then turns a chalky white after heating.

In winter, <u>salt is spread</u> on icy roads.

(A) The more salty a solution is, the more difficult it is to freeze it.

(B) The snag about salting roads is that extensive damage is done to the metal work of cars.

(C) Salt prevents ice forming on roads.

(D) The use of salt keeps many roads open which might otherwise be too dangerous to use.

An <u>atom</u> is composed essentially of a small, dense, positively charged nucleus surrounded by levels of diffuse negatively charged electrons.

(A) Man's understanding of the atom is still incomplete since he has yet to understand the forces that hold the protons together in the nucleus.

(B) The chemical activity of an atom is determined by the number and energy of its outermost electrons.

(C) An atom is the smallest particle of an element that can take part in a chemical change.

(D) A knowledge of the structure of the atom can be used to predict its chemical behaviour.

---

7. A pure liquid may be separated from a solution by the process of distillation.

(A) The refining of petroleum involves the distillation of crude oil into its principal fractions.

(B) Distillation is a method of purification frequently used in organic chemistry.

(C) If strong forces exist between the molecules of the different components of a mixture, complete separation is impossible with even the best fractionating column.

(D) By evaporating a solution and recondensing the resulting vapour in a separate vessel we can separate a pure liquid from a solution.

---

8. The farmers in Britain spend millions of pounds every year on fertilisers.

(A) Fertilisers are mixtures of the chemicals which growing plants need.

(B) The careful use of fertilisers ensures good crops from poor or tired fields.

(C) The use of too much fertiliser or the wrong one can lead to severe damage to the field and pollution of neighbouring water.

(D) When a crop is taken from a field, the chemicals which were needed to grow that crop must be returned to the soil of that field.

---

9. Fireworks give a great deal of pleasure and they also have their uses.

(A) The distress signals which are used at sea are large fireworks which make bright light and a great deal of smoke.

(B) The chemical reaction inside the cardboard container of the firework is started by lighting the fuse.

(C) Exciting variety is achieved by using salts of those elements which impart colour to a flame, by using grains of iron to make sparks and by careful packaging to bring about different rates of burning.

(D) All fireworks contain two principle chemicals; the first one burns and gives out a great deal of heat and gas as it does so, the second provides the oxygen for the first to burn in.

---

0. <u>Oxygen</u> is an odourless, tasteless, colourless gas which is slightly heavier than air.

(A) One should consider that O_2 and O_3, both found in air, differ in their physical and chemical properties.

(B) It is commonly prepared in the laboratory by treating hydrogen peroxide with a catalyst like manganese (IV) oxide.

(C) The large amount of nitrogen in the air accounts for the fact that air is lighter than oxygen.

(D) Without oxygen, life as we know it, could not exist.

---

1. Dilute solutions of <u>ammonia</u> in water are sold in bottles bearing the name "Scrubbs Cloudy Ammonia".

(A) Scrubbs Cloudy Ammonia is a valuable chemical for cleaning many pieces of household equipment.

(B) The greater the pressure which is applied to a gas in order to make it dissolve in a certain volume of solvent, the greater the mass of the gas which dissolves.

(C) The label on the bottle does not list the other chemicals which are used and so there is no quick way of knowing why this solution is cloudy.

(D) Ammonia is an alkaline gas, so "Scrubbs Cloudy Ammonia" will turn red litmus blue.

---

2. <u>Chlorine</u> is a highly reactive element and readily forms negatively charged chloride ions.

(A) It is remarkable that chlorine, an electron-receiving element, can react with other electron-receiving elements.

(B) In accepting one electron, a chlorine atom completes its outer electron shell.

(C) One of the important applications of chlorine is in the purification of drinking water.

(D) Chlorine reacts directly with sodium with the evolution of light and heat.

---

13. The rate at which a given quantity of a solid dissolves in a fixed volume of liquid varies with the size of the solid particles.

(A) When making a copper sulphate solution it is better to use many fine copper sulphate crystals, rather than a few large ones.

(B) As a rule, the smaller the particles of solid, the faster they dissolve.

(C) The reason for this is that by decreasing the particle size of the solid one increases the surface area available for interaction.

(D) The rate of dissolution of large solid particles can often be increased by warming and by vigorous stirring.

---

14. Water which has either chlorine or sulphur dioxide dissolved in it will function as a bleach.

(A) A chemical which turns a coloured material white is called a bleach.

(B) Chlorine is a powerful bleach but there is a danger that it may damage the material; sulphur dioxide is less destructive but also less powerful.

(C) Bleaches can also be used to destroy bacteria.

(D) The bleaching action occurs because the chemical nature of the coloured material is changed to a colourless chemical either by the addition or removal of oxygen.

---

5. The Periodic Table is a valuable tool to the chemist since it shows trends in the physical properties of the elements and predicts patterns of chemical behaviour.

(A) The properties of an element can be predicted from a knowledge of the properties of a neighbouring element.

(B) It is disappointing that the Periodic Table does not predict all the different valency states of the elements.

(C) The Periodic Table now contains over 100 elements. These are arranged in chemical families, e.g. the alkali metals, the halogens, the transition metals, etc.

(D) The Table is based on the principle that if elements are arranged in the order of increasing atomic number they show a periodic repetition of properties.

---

16. Pollution is a world-wide problem of great importance.

(A) The waste which man produces and which Nature is slow to deal with, including heat as well as solids, liquids and gases, is the material of pollution.

(B) No animal besides man causes pollution but the oxygen in the air which we breathe could be looked upon as being the result of the pollution of the earth by plants: there was no oxygen in the atmosphere which first surrounded the earth.

(C) Almost every chemical process for which man is responsible results in pollution.

(D) New uses should be found for every kind of waste in order to minimise the damage done by pollution.

---

17. When a silver nitrate solution is added to a solution of sodium chloride a curdy white precipitate is formed.

(A) A precipitate is formed because the solution is overburdened with an insoluble solute.

(B) This reaction is the basis of the test for the presence of chloride ions.

(C) The reaction between silver nitrate and sodium chloride is best described by the equation:

$$Ag^+ + Cl^- = AgCl$$

(D) It would be interesting to find out why the solubility of AgCl differs so much from that of NaCl.

---

18. When atoms either lose or gain electrons, they become ions.

(A) This statement would be more meaningful if it indicated which type of atoms tend to gain electrons and which type tend to lose electrons.

(B) Pure metal may be obtained by electrolysing a solution containing its ions.

(C) Ions are electrically charged particles which in solution conduct electricity.

(D) Since atoms can become ions by electron loss, they must be made up of positively and negatively charged particles.

---

19. Steel can be made into stainless steel by the addition of elements like chromium and nickel.

(A) The elements chromium and nickel both form a thin layer on their surfaces which protects them from further action. They do this when they are mixed in with steel and so the steel is also protected.

(B) At present stainless steels are expensive: an equally effective protection which was also cheap would be a major technological breakthrough.

(C) The use of stainless steel for knives and razor blades is fairly recent because it was not possible to put a sharp edge on implements which were made out of the early stainless steels.

(D) Stainless steel is a form of steel which does not rust.

The reaction between hydrochloric acid and zinc is represented by this equation:-

$$2HCl + Zn = ZnCl_2 + H_2$$

(A) Most metals above hydrogen in the electrochemical series give hydrogen when treated with hydrochloric acid.

(B) The action of hydrochloric acid on zinc is a convenient method of preparation of small samples of hydrogen.

(C) Traces of several other gases may be present in the hydrogen; if their origins were known, it might be possible to eliminate them and so prepare pure hydrogen.

(D) Hydrogen is a light, colourless gas which pops when lit in air.

Some chemical substances can affect the rate of a chemical reaction between other chemical substances.

(A) Such substances are known as catalysts.

(B) An uneconomic process can suddenly be made profitable by the discovery of an effective catalyst. Considerable sums are spent on improving existing catalysts and finding new ones.

(C) Although a catalyst appears unchanged at the end of a reaction, it is always chemically involved in the reaction.

(D) Catalysts increase the rate of chemical reactions by making bond-breaking and bond-making easier.

When smoke particles are suspended in air and observed through a microscope, they can be seen to carry out an irregular zigzag movement.

(A) The fact that this motion can also be observed with solid particles suspended in liquids, shows that gases and liquids have important properties in common.

(B) This motion can be used to illustrate the random movement of the molecules of a gas.

(C) The movement is caused by the bombardment of smoke particles by the rapidly moving air molecules.

(D) This irregular movement of fine particles is called Brownian motion.

23. Fire often does a great deal of damage merely because the correct fire extinguisher was not used in time.

(A) Everywhere where there is a risk of fire, including the home and the car, should have a fire extinguisher and it should be designed for the type of fire which is likely to occur.

(B) Fire extinguishers function in different ways. One type produces a rush of gas which blows out the flame, another produces foam and gas, a third uses a fine powder and a fourth uses a blanket.

(C) There are numerous different types and shapes of fire extinguisher. Most of them are started by turning them upside down and/or by banging the knob against something hard.

(D) Either the flame must be cooled below the temperature at which the gases in it will burn, or it must be starved of oxygen.

---

24. Salt and sugar both dissolve in water; the solution of salt will conduct electricity, the solution of sugar will not.

(A) A solution which will conduct electricity contains ions; one that does not do so contains molecules.

(B) Crystals dissolve because Nature prefers the random motion of the particles in a solution to the orderliness of the same particles in a crystal.

(C) Warming the water in which sugar is being dissolved makes a marked difference to the quantity which will dissolve; it makes little difference to the quantity of salt which will dissolve.

(D) The chemical charges which occur at the electrodes when a current is passed through a solution of salt can lead to the preparation of many useful chemicals like chlorine and sodium hydroxide.

The halogens, group VII in the periodic table, are the most reactive of the non-metallic elements.

(A) The halogens readily combine with other elements.

(B) The activity of the halogens is explained by their tendency to gain electrons.

(C) One should consider that there is a steady decrease in reactivity from fluorine to iodine.

(D) The halogens are never found free in nature but always occur in the form of salts.

---

6. When copper turnings are heated in air, a black powdery compound is formed.

(A) The compound is copper oxide and has the formula CuO.

(B) It is difficult to change all copper into copper oxide, even in excess air or oxygen.

(C) This reaction can be used to determine the percentage of oxygen in air.

(D) Analysis of the black compound shows that copper atoms and oxygen atoms combine in the ratio 1:1.

---

7. Separation of the several dyes in ink is most easily done by chromatography.

(A) Chromatography is used to check that only harmless dyes are used in fruit drinks.

(B) Chromatography relies on the fact that each chemical in a mixture can be 'bowled along' at a different speed by the moving fluid (generally a liquid).

(C) Liquids in a mixture of liquids like petrol can be separated by chromatography; they are moved along a tube filled with warm powder by a stream of gas.

(D) The coloured components of chlorophyll can be separated on filter paper with an alcohol solvent.

---

28. The protection of a less reactive metal (like iron) by a coating of another which is more reactive (like zinc) is an interesting way of preventing rusting.

(A) In tin plating the less reactive metal is put on the outside but when the rusting does eventually penetrate down to the iron beneath, the holes which are formed are small but deep.

(B) Iron which is protected by a thin layer of zinc is said to have been 'galvanised'.

(C) The zinc serves to protect the iron underneath because the zinc is sacrificed before the iron starts to rust.

(D) The protection of iron by zinc is commonplace; corrugated iron sheets for buildings, dustbins and wall nails are examples.

Teaching

Notes

TOPICS FOR DISCUSSION IN CLASS

The information which is given in this booklet has been taken from textbooks and from teachers' notes. There are no intentional mistakes and yet this is in no sense a test. Each page has sixteen topics for discussion in class; all you are asked to do is to say which of these topics you like and which ones you find unsatisfying. To make this job easier the sixteen topics on each page are grouped into four groups each of four topics. You are to think about the four topics in each group at the same time; first decide which one of the four topics you would (or perhaps did) find most satisfying for discussion in form and you are to give it two ticks(✓✓) in the appropriate box of the Answer Sheet. Next decide which of the four topics you like least and put an 'X' in its box. Then decide which you prefer of the two which remain and give it one tick; leave the last of the four boxes empty. Treat the next four topics in the same way and continue until you have dealt with all sixteen topics on the page, four at a time. Turn over and carry on in the same way.

When you have finished a set of four boxes it might look like this:-

(1) ✓✓	(2)	(3) X	(4) ✓

This student found (1) most satisfying, he disliked (3) and he preferred (4) to (2). When he had finished a page it looked like this:-

Page I

(1) ✓✓	(2)	(3) X	(4) ✓
(5) ✓	(6) ✓✓	(7)	(8) X
(9)	(10) X	(11) ✓	(12) ✓✓
(13) ✓✓	(14) ✓	(15) X	(16)

The information which you are giving by doing this test is needed to help us to decide whether we have been teaching you your chemistry in the right way and in the right order. Concentrate carefully on what you are doing and learn as much chemistry from the material given here as you can. Work quite fast and be sure to finish by the end of the time available(about 35 minutes). You may ask for extra time if you need it.

Please start by putting your name and initials at the top of the Answer Sheet and also the name of your chemistry teacher. Please do not make any marks of any kind on the actual booklet itself.

Carbon Dioxide (CO_2) Carbon Monoxide (CO)

Sublimation.

(1) Solids are said to sublime if they turn directly into gas when heated, e.g. solid carbon dioxide heat carbon dioxide gas;

(2) Strong bonds exist between atoms in CO_2 but weak bonds between molecules, heating breaks weak bonds;

(3) Mixture of solid carbon dioxide and a volatile liquid like ether is most effective freezing mixture for obtaining low temperatures in lab.;

(4) SO_2 (sulphur dioxide) gas forms liquid when cooled, i.e. no sublimation, but SO_3 does sublime.

Lime Water.

(5) Carbon dioxide reacts in solution with calcium ions to form insoluble carbonate of calcium, this is precipitated;

(6) A solution of lime water is turned milky by carbon dioxide, may go clear again if lots of gas is used;

(7) A solution of barium ions would serve just as effectively as a solution of calcium ions;

(8) Provides convenient method for distinguishing carbon dioxide from other gases.

Oxides of Carbon.

(9) CO reacts with oxygen to give CO_2 because more stable electron-sharing bonds are formed; energy released;

(10) Both carbon monoxide and carbon dioxide are colourless gases, carbon monoxide takes oxygen from oxides (e.g. lead oxide) and is therefore reducing agent, carbon dioxide is not reducing agent;

(11) Oxygen-free samples of less reactive metals prepared by heating oxide in carbon monoxide, e.g. clean lead from yellow lead oxide;

(12) $CO_2(g) + C(s)$ if heated strongly $\longrightarrow$ 2CO, reverse action impossible.

Preparation of carbon dioxide.

(13) Convenient supply for lab. demonstrations from Kipps' apparatus containing marble and mod. conc. hydrochloric acid;

(14) Weak acids unsuitable even in concentrated solution;

(15) H^+ from the acid displace CO_3^{2-}, H_2CO_3 formed, unstable so $CO_2(g)$ given off;

(16) Every carbonate gives carbon dioxide when treated with either hydrochloric, sulphuric or nitric acids.

Separation of Mixtures, Purity

Distillation.

(1) Separation of mixture of liquids by distillation possible if the chemicals in the mixture have different boiling points, i.e. at any one temperature, the vapours of those chemicals exert different pressures;

(2) Fairly weak bonds between alcohol molecules and water molecules exist and they prevent complete separation of alcohol from water and water from alcohol;

(3) The term distillation means separation by boiling followed by condensation of vapour;

(4) Some alcohol is made by fermentation of sugar followed by distillation of the products of fermentation.

Chromatography.

(5) If the strengths of the interactions between the dye molecules and the water molecules on the fibres of the filter papers could be measured, the order of the spots on a chromatogram could be predicted;

(6) Equilibrium occurs between the molecules of dye in the moving solvent and the molecules of dye in the solvent which is held by the fibres of the paper;

(7) Separations of mixtures by this means is called chromatography;

(8) This is much the easiest way to demonstrate that ink or chlorophyll is mixture.

Crystallisation.

(9) Fairly pure crystals of copper sulphate can be prepared from solution of copper oxide in sulphuric acid by crystallisation;

(10) Crystals form when concentrated solution of crystalline chemical is cooled;

(11) It is not obvious why the formation of some crystals, e.g. sodium thiosulphate, is delayed by supercooling;

(12) The solubility of most solid chemicals decreases on cooling and the crystals form when the solvent becomes saturated.

Melting Point, test for purity.

(13) Accurate melting points of pure chemicals are known, chemicals which melt at temperatures lower than the correct melting point are impure;

(14) If Chemical X is known to be either Chemical A or Chemical B it is possible to determine which one it is by finding the melting point of pure A, of A with a little X, of pure B and of B with a little X;

(15) If a little of C lowers the melting point of D and a little of D lowers the melting point of C, some mixture of C and D must have a melting point which is raised when either C or D is added to it;

(16) A small proportion of solute interferes with the formation of crystal lattices and so makes formation of lattice more difficult.

Chlorine and The Halogens

Preparation of Chlorine. Conc. hydrochloric acid and warm manganese (iv) (di)oxide.

(1) Convenient for small samples of chlorine in lab. for showing props. of chlorine;

(2) Manganese dioxide oxidises hydrochloric acid to chlorine and water;

(3) In acid conditions solid manganese (IV) compounds take electrons. manganese (II) in alkaline conditions receives electrons from chlorine and forms solid manganese (IV) compounds;

(4) Manganese (IV) reacts with cold hydrochloric acid to give an acid containing Mn(IV), when this is warmed it decomposes to give $Cl_2(g)$ and Mn(II) (aq).

The reaction of chlorine with iron.

(5) The compound formed between iron and chlorine (iron (III) chloride) is largely ionic in water but is largely covalent when in the solid state;

(6) If chlorine is passed over hot iron, the iron burns brightly to give a brown smoke of iron (III) chloride;

(7) When iron is oxidised by chlorine a sharing of electrons between iron and chlorine atoms occurs, this serves to complete the shell structure of both atoms;

(8) This is the only possible way to make dry iron (III) chloride for reactions which require it as a catalyst.

Uses of chlorine (e.g. as a bleach).

(9) The bleaching action of chlorine on wet litmus serves as a good test for the gas;

(10) Coloured chemicals are oxidised by chlorine because their structure is changed to different (and colourless) chemicals by altering their electron structure;

(11) Chlorine bleaches by removing electrons, sulphur dioxide (gas) bleaches by adding electrons;

(12) Dilute solutions of chlorine in water or alkali bleach discoloured fabrics and ink stains.

The Halogens. A Periodic Group (F, Cl, Br, I).

(13) Some aspects of the chemistry of fluorine do not conform to the general behaviour in this Group, fluoride ions are particularly small and have a larger charge on each unit of surface area;

(14) All the elements in a Group have the same number of electrons in the outermost shell, they differ only in the number of complete shells;

(15) Grouping like elements together beneath one another on the Periodic Table provides a framework for learning the chemistry of elements;

(16) The elements in a Group have similar physical and chemical properties.

Reactions: fast and slow.

Catalysis and prep. of oxygen.

(1) The catalyst manganese (IV) oxide is needed in both common lab. preps. of oxygen, i.e. from hydrogen peroxide or from potassium chlorate (V);

(2) Catalysts hasten chemical reactions but they can be recovered again in some chemical condition when the reaction has finished;

(3) Catalyst provides reaction with pathway (involving intermediate compounds) which requires less energy, hence more molecules can react so lower temperature is possible;

(4) Chemical intermediates like $KMnO_4$ and Cl_2 are known to form when MnO_2 reacts with potassium chlorate (V), the composition of the intermediates with hydrogen peroxide are not known.

Slow reaction, e.g. rusting.

(5) Unprotected iron rusts in damp air, water and oxygen are both necessary;

(6) Iron in electrolyte (solution of ions) or with uneven exposure to oxygen rusts rapidly;

(7) Rusting is prevented by physical means (i.e. paint or plastic coatings) and by chemical means (i.e. zinc).

(8) An anodic region (electrons released) forms under drop and cathodic region (electrons consuming) forms at edge where oxygen is abundant, rust (red iron (III) oxide) forms between them.

Fast reaction (e.g. fireworks).

(9) Way of showing fast reaction in reasonable safety;

(10) Different colours in flames, atoms (of different elements) with excited electrons, cool and give out different coloured light;

(11) Compact mixtures of oxygen-rich, unstable chemicals in body of firework give large volume of hot gas very quickly when ignited;

(12) Energy released by reaction; energy stored in firework as chemical energy and released in mass of hot gas as kinetic energy.

Nature's catalysts, enzymes in fermentation.

(13) All living cells employ cataylsts, called enzymes, e.g. yeast uses enzymes in conversion of sugar to alcohol and carbon dioxide, process called fermentation.

(14) Rate at which enzyme controlled reaction proceeds depends on the pH, optimum rate for a specific pH, chemical nature of enzyme determines pH.

(15) Yeast oxidises the sugar and chemical energy useful to yeast is released, many different stages therefore many different enzymes;

(16) Fermentation useful in preparing dilute solution of alcohol for demonstrating distillation.

Solutions.

Solutions of solids, solubility.

(1) Solids which dissolve do so because change of orderly crystalline lattice into random scattering of particles in solution releases energy;

(2) Analysis of an unknown chemical often achieved by gathering information on the solubility of salts of that compound;

(3) Solubility is defined as mass of solute required to saturate a specified mass of solvent at a specified temperature;

(4) Sugar and salt are both soluble, sugar as complete hydrated molecules, salt as hydrated ions, water accommodates both types of particle.

Water of crystallisation.

(5) If water vapour pressure inside a crystal is greater than the water vapour pressure outside, water is lost to atmosphere, i.e. efflorescence occurs;

(6) White copper sulphate goes blue when water is added, convenient test for water;

(7) Water trapped inside a crystal by chemical means is called water of crystallisation, such crystals are in no sense damp;

(8) Water in crystals held by electron-sharing bonds in well-defined spaces, and in precise molecular proportion.

Solution of gases.

(9) Gas above solvent comes to equilibrium with gas within solvent so mass of gas which is required to saturate given volume of solvent is proportional to the pressure which is applied to the gas.

(10) In theory the volume of a gas which dissolves is independent of pressure, in practice gases do not conform to this ideal behaviour;

(11) Some important laboratory reagents (e.g. hydrochloric acid, ammonium hydroxide) are solutions of very soluble gases;

(12) All gases are soluble, some much more soluble than others, but all less soluble in warm water than in cold water.

Solution of acids;

(13) Acidity is dependent on the presence of H_3O^+ in solution;

(14) Solutions of acids are frequently needed for the preparation of salts, other chemicals, for catalysts, etc.;

(15) Acid strength is determined by the extent to which acid molecules are split into ions, not by the mass of the acid which will dissolve;

(16) An acid solution reacts with indicators, gives carbon dioxide with carbonates and dissolves metal, oxides and bases.

Ammonia (NH_3).

Physical properties.

(1) Ammonia is a light, colourless soluble gas with a strong smell, easily turned into liquid by slight pressure or cooling;

(2) Changes from liquid to gas and gas to liquid so easily that it is usually selected for the demonstration of how a refrigerator works;

(3) The atoms within a molecule of ammonia are covalently bonded, forces between molecules are weak, hence gaseous nature of ammonia at room temperature;

(4) Hydrogen combines with most elements including nitrogen to give compounds which are gases, hydrogen combines with oxygen to give liquid (H_2O), another type of bond needed to explain why water is liquid.

Manufacture of ammonia (Haber process) from nitrogen and hydrogen.

(5) Good yields at low temperatures because reaction is exothermic (energy released) and yet process is conducted at quite high temperature (about 370°C);

(6) Ammonia (gas) sprayed beneath surface of soil to replace nitrogen removed by crops;

(7) Molecules of N_2 and H_2 react when they collide, catalysts encourage fruitful collisions, heat required to give molecules sufficient energy to react at acceptable rate;

(8) Mixture of three parts of hydrogen to one part of nitrogen heated to 370°C and compressed to 250 Atmospheres pressure, passed over catalyst of pure iron;

Behaviour of ammonia in water.

(9) Concentrated and dilute solutions available in labs., constant demand for demonstration of reactions of ammonia and for cleaning sinks and glassware;

(10) The properties of ammonia in water are quite different from these of ammonia in other solvents, e.g. covalent solvents like ethanol, clearly role of solvent is important;

(11) Solution is found to be alkaline to red litmus, to acids and to solutions of metal salts;

(12) NH_3 reacts with H^+ (aq) in water forming NH_4^+, OH^- are released and solution is alkaline, weakly alkaline though.

Tests for ammonia.

(13) Strip of filter paper dipped in a solution of a copper salt goes dark blue in ammonia gas;

(14) NH_3 (aq) displaces H_2O molecules which normally surround the pale blue copper ion (Cu^{2+});

(15) Phosphorus is in same Periodic Group as nitrogen but PH_3 will not do this test, phosphorus atom less able to give electrons away;

(16) More reliable test for ammonia than either red litmus or the 'white smoke test' with hydrogen chloride.

Thinking

in

Chemistry

2

In Part One ......

.... of this test you will meet five situations which are typical of the many situations which a chemist encounters. Each situation is described briefly, a question is posed and some statements are made. You are simply asked to write one letter in the appropriate box of the response sheet.

Put an A if the statement gives a SENSIBLE answer to the question and if it is also JUSTIFIED in relation to facts and observations which are described.

Put a B if the statement gives a SENSIBLE answer to the question, but it is NOT JUSTIFIED because it makes use of information not given.

Put a C if the statement gives only an INADEQUATE answer to the question because it IGNORES important parts of the information.

Put a D if the statement, though correct, is IRRELEVANT to the question and so does not answer it.

It is IMPORTANT ....

that you notice that all the science in this book is correct. You are NOT looking for incorrect statements.

SITUATION 1.

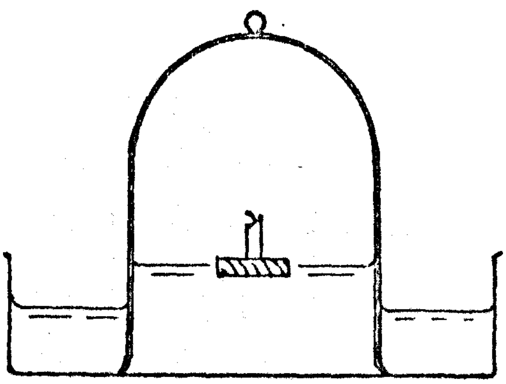
A bottle contains a red solution called Solution A. It has at lease one chemical dissolved in it. If .....

Solution A is cooled	a white solid appears in a red solution,
if this is then warmed	the solid dissolves.
If Solution A is treated with acid	a white solid appears,
and if excess alkali is added	the solid dissolves and the solution turns green.

What conclusions can we draw about the nature of Solution A?

1. The same precipitate is formed by either the addition of acid or by cooling.
2. A further sample of Solution A can be made in the lab.
3. Changes in temperature do not cause the same effects on Solution A as changes in pH.
4. Cooling causes Solution A to become saturated with respect to one chemical.
5. Solution A contains the chemicals which behave independently.
6. The addition of acid to Solution A causes a chemical change which leads to the formation of an insoluble chemical.
7. Solution A is a mixture.

SITUATION 2.



When the bell jar that is shown in the diagram is slowly lowered over the burning candle, the flame is extinguished and water is seen to enter the jar to a height of about one fifth of the total volume of the jar.

Analysis of the gas that remains in the jar shows that it contains 16% by volume of oxygen and 4% by volume of carbon dioxide with a little hydrocarbon vapour.

What conclusions can be drawn from these observations?

1. Hot gases are trapped when the jar is lowered over the flame and the water rises as the gases cool.

2. The flame is extinguished when there is insufficient oxygen to support combustion.

3. Similar observations are made when a candle with a long wick is used.

4. More than 4% by volume of carbon dioxide suffocates the flame.

5. The water level will rise to the same height if an alcohol flame is used.

6. The flame is extinguished when the proportion of oxygen in the air is reduced to a minimum value.

SITUATION 3.

A student observes that when chlorine gas is passed over solid iodine, an apparently black liquid is formed which slowly turns into an orange solid. If this orange solid is allowed to stand in the air, a black liquid is observed.

What conclusions may be drawn?

1. The 'apparently black liquid' contains both iodine and chlorine.

2. Chlorine reacts with iodine.

3. The iodine cannot easily be recovered afterwards.

4. Two reactions are occurring.

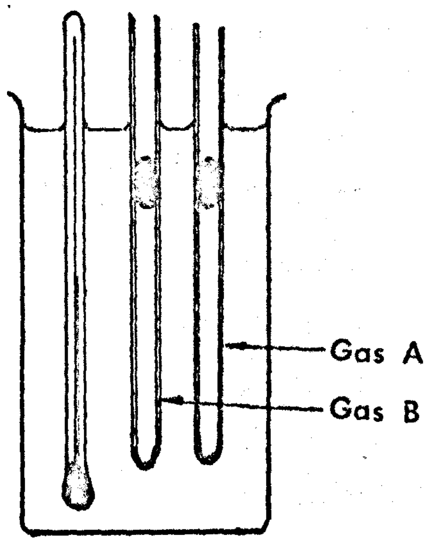
5. The orange solid can only be preserved in an atmosphere of chlorine.

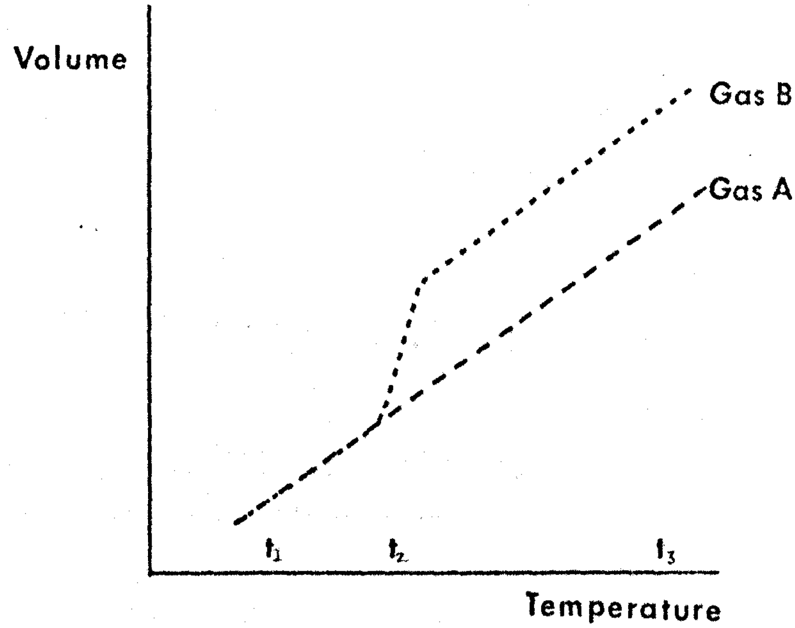
6. The black liquid which is formed when the iodine reacts with chlorine has the same composition as the black liquid which is formed when the orange solid is left in the air.

7. The black liquid resembles liquid bromine.

8. There is a greater proportion of chlorine in the orange solid than there is in the black liquid.

SITUATION 4.



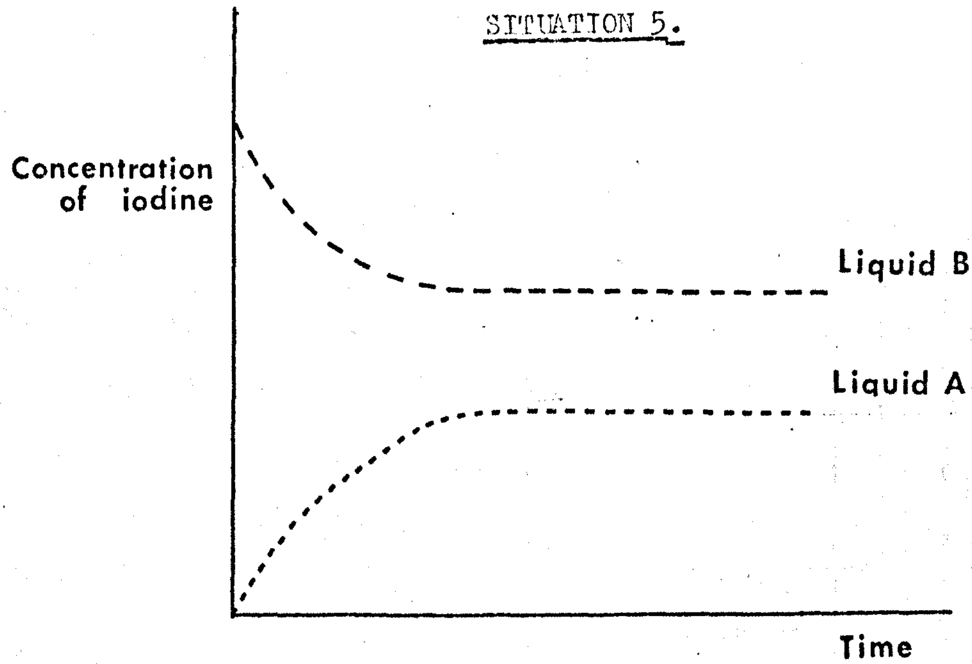


The diagram shows two tubes containing gas A and gas B, respectively, each trapped by a drop of mercury. The tubes are held in an oil bath so that the volume of the two gases at different temperatures can be measured. The graph shows these volumes.

What conclusions may be drawn about the behaviour of gas B?

1. Gas B may be turned into a liquid by cooling.
2. The volume which is effectively filled by each molecule of gas B contracts as the temperature is lowered.
3. At temperature t_3 there are more molecules of gas B than there are of gas A.
4. At temperature t_2 there are as many molecules of gas B as there are of gas A.
5. Gas B undergoes a reaction at temperature t_2.
6. At temperatures higher than t_1 the two gases behave differently.
7. When gas B is cooled from temperature t_3 to temperature t_1, it will have the same volume as gas A.

SITUATION 5.



Liquid A and Liquid B do not dissolve in one another but iodine is soluble in them both. A flask which contained a litre of Liquid B with iodine dissolved in it was shaken with a litre of Liquid A for five minutes. The flask was allowed to stand and then 10 cm^3 samples of each liquid were withdrawn with pipettes and the concentration of iodine in these small volumes was determined. Shaking was resumed and after each period of five minutes further 10 cm^3 samples of each liquid were withdrawn and analysed. The readings which were obtained are shown in the diagram.

<u>What conclusions can be drawn about the behaviour of the iodine?</u>

1. If all the iodine had been dissolved in Liquid A at the start of the experiment, the two curves would have crossed over.
2. The iodine is more soluble in one liquid than in the other.
3. Similar results would have been obtained if half-litre volumes of liquids had been used.
4. If the experiment were to be repeated at another temperature different readings would have been obtained.
5. The concentrations of iodine come to a state of equilibrium.
6. Bromine will behave similarly.
7. If all the iodine had been dissolved in Liquid A at the beginning of the experiment, the concentration of iodine in this liquid would have fallen.
8. The curves would have levelled off sooner if a mechanical stirrer had been used.

In Part Two.......

.... of this test you will meet five further situations. This time experiments are suggested which might be undertaken to shed light on the problem which is posed.

Put an E if the experiment will lead to information BEYOND that already given, AND this information is RELEVANT to the problem stated.

Put an F if the experiment will lead to information BEYOND that already given, BUT this information is IRRELEVANT to the problem stated.

Put a G if the experiment will merely give information which is ALREADY AVAILABLE.

SITUATION V.

Powder A was in a paper bag and Crystals B were in a large carton. Both were stored in a damp cupboard. After some days A was found to be wet and B was still fairly dry.

THE PROBLEM is to determine the cause of this difference in behaviour. You are to assume that the chemicals will be re-examined for change some time after the experiment has been set up.

1. Dry both chemicals before putting them away.
2. Store Powder A in the carton and Crystals B in the bag.
3. Grind B to a powder before storing A and B together as before.
4. Find the degree of humidity of the air inside the cupboard.
5. Store both the chemicals in a dry cupboard.
6. Store Powder A and Crystals B on separate watch glasses contained in the same paper bag in the cupboard.

SITUATION W.

A line of coloured iron poles was sticking out of a muddy beach between the low and high tide levels. The poles which were painted red were well covered with barnacles; the blue poles were fairly free of barnacles.

THE PROBLEM is to determine the cause of the uneven distribution of barnacles. Assume that a note will be made of the number of barnacles on each pole before the suggested experiment and another count will be made after a lapse of time.

1. Prepare a mixture of red and blue paint and then repaint some poles with the mixture.
2. Alter the colour of the paint by adding charcoal or other black pigment and repaint the poles.
3. Paint iron poles with yellow and green paint.
4. Paint wooden poles with the red and blue paint and stand them between the tide levels.
5. Move each of the existing poles to a new site.
6. Have the red and blue paints analysed to determine chemical composition.
7. Turn the poles upside down.
8. Clean the poles and then paint them in bands with alternating red and blue paint.
9. Study the size etc. of the barnacles on the red and blue poles.

SITUATION X.

When a solution of the blue dyestuff Q is freshly prepared in tap water it is coloured. The colour fades when the solution has been left to stand in sunlight in a beaker on a windowsill.

THE PROBLEM is to determine the cause of the fading. Assume that the colour of the solution will be re-examined some time after each experiment has been set up.

1. Prepare solutions of Q covering a range of concentrations.
2. Dissolve Q into water which has previously been boiled to remove the air dissolved in it and then stopper the bottle tightly.
3. Leave a fresh sample of Solution Q in a darkroom.
4. Prepare a solution of Q in alcohol and leave it in sunlight.
5. Add more tap water to a faded solution of Q.
6. Leave a fresh solution of Q in sunlight in a refrigerated thermostatic tank.
7. Examine the behaviour of other blue dyestuffs.

SITUATION Y.

Two bottles containing liquids were placed on a tray. After a while a fine layer of "smoke" was observed and this gradually settled on the tray as a white powder.

THE PROBLEM is to determine the cause of the "smoke". Assume that each experiment will be re-observed at a suitable time after setting it up.

1. Wash, dry and tighten the stoppers.
2. Place the tray and bottles in a fume cupboard or fume chamber to see whether the "smoke" disperses.
3. Test small portions of the two liquids in test tubes.
4. Leave the tray and bottles in a cabinet containing dry air.
5. Store the bottles separately.
6. Dip separate strips of filter paper into each liquid and hold the strips near one another.
7. Remove both stoppers and leave.
8. Examine the solubility of the powder in water.

SITUATION Z.

The painted iron posts which are used to support the netting round playgrounds and tennis courts rust far more rapidly near the ground (i.e. within 9 inches of the ground) than they rust elsewhere.

THE PROBLEM is to establish the cause of the uneven rusting. Assume that one or more poles are treated in each of the following ways and they will be carefully examined for rust at suitable intervals of time thereafter.

1. Scrape all the paint off.
2. Encase a pole in concrete.
3. Provide a shield to prevent rain water from splashing up from the ground on to the post.
4. Raise a pole and replant it with the buried section encased in polythene.
5. Replace an iron pole with another which is identical in all respects except in that it is made of aluminium.
6. Raise a pole and replant it the other way up.

Chemical Judgement

THIS IS A TEST OF CHEMICAL JUDGEMENT. It is designed for those who do not necessarily know a great deal of chemistry but have an understanding of the subject. You are given information and you are asked to say whether some further observations and experiments make sense. There are no catches, no mistakes, no false statements. The test is in THREE parts. Work carefully and sufficiently fast to finish the test in the time allowed (40 minutes).

INSTRUCTIONS for PART ONE.

In Part One there are five situations which are typical of the many situations which a chemist encounters. Each situation is described briefly, a question is posed and some statements are made. You are asked to put an 'X' in one (and one only) of the four boxes for that statement.

Put an 'X' in the A Box if the statement has a SENSIBLE bearing on the problem and if it is also JUSTIFIED and HELPFUL in relation to the facts and observations which are given.

Put an 'X' in the B Box if the statement has a SENSIBLE bearing on the problem but it is NOT JUSTIFIED because it makes use of important information about the situation which has not been given in the description.

Put an 'X' in the C Box if the statement has some bearing on the problem but contains information which is INADEQUATE AND UNHELPFUL because it merely restates the information which has been given already or fails to throw any light on the problem.

Put an 'X' in the D Box if the statement, though sensible, is IRRELEVANT to the problem given and so has no bearing on it.

REMEMBER Choose A if the statement is JUSTIFIED

	B	NOT JUSTIFIED
	C	INADEQUATE
	D	IRRELEVANT

THINK CAREFULLY about the DIFFERENCES between these choices BEFORE you start. When you have started, think about each statement carefully and quite independently of all the other statements.

Put ONE 'X' ONLY in EACH GROUP OF FOUR BOXES.

SITUATION 1.

A bottle contains a red liquid called Solution A. It has at least two chemicals dissolved in water. If ....

Solution A is cooled a white solid appears in a red solution,

if this is then warmed the solid dissolves.

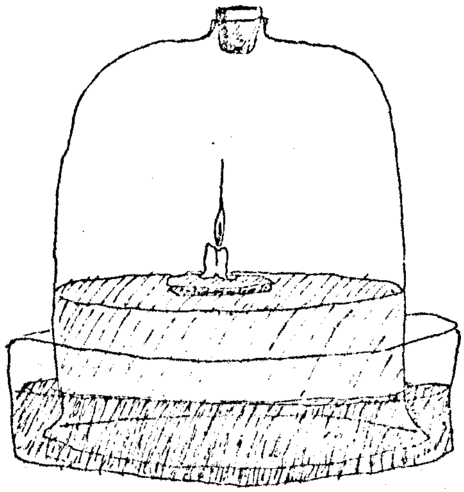
If Solution A is treated with acid a white solid appears,

and if excess alkali is added the solid dissolves and the solution turns green.

Solution A is unlike other solutions. Why?

1. The same precipitate is formed by either the addition of acid or by cooling.
2. A further sample of Solution A can be made in the lab.
3. The behaviour of one of the chemicals depends on pH.
4. Cooling causes Solution A to become saturated with respect to one chemical.
5. Solution A is a mixture.
6. Solution A contains an indicator and another chemical.
7. Solution A contains phenol (C_6H_5OH) which is soluble in alkali but not in acid.
8. The two chemicals in Solution A do not react with water.

SITUATION 2.



The bell jar is lowered over the burning candle and left for a few minutes. When the stopper is pushed into position the candle flame is slowly extinguished and water is seen to enter the jar.

Analysis of the gas that remains in the jar afterwards shows that there is 16% by volume of oxygen and 4% by volume of carbon dioxide.

The water rises to replace 1/5th of the air in the jar. Why?

9. While the candle is alight, it warms the air in the jar.
10. The flame is extinguished before all the oxygen in the jar has reacted.
11. Similar observations are made when a long, thin candle is used.
12. The water rises when the hot gases in the jar cool and contract.
13. The water level will rise to about the same height if an alcohol flame is used.
14. The flame is extinguished when the proportion of carbon dioxide in the gas which reaches the flame rises to a maximum value.
15. Air contains 1/5th oxygen.
16. Water does not rise so far up the jar if the stopper is pushed in immediately the jar is in place.

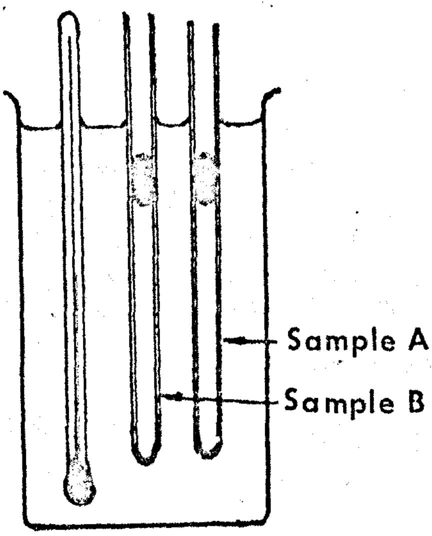
SITUATION 3.

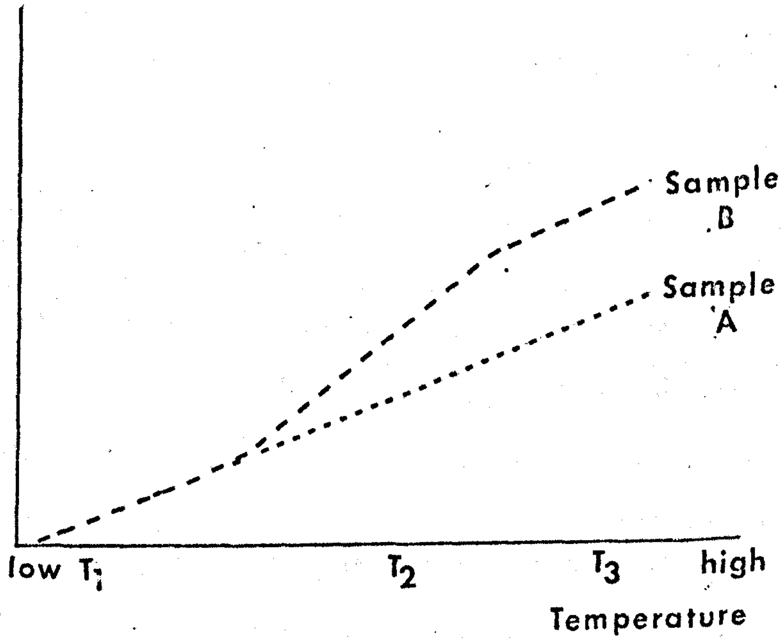
A student was experimenting with the elements in the VIIth Periodic Group, i.e. the Halogens. He observed that when chlorine gas is passed over solid iodine, first an apparently black liquid formed and there was a small increase in mass. Then the liquid slowly turned into an orange solid with a further increase in mass. When this orange solid was allowed to stand in the air, a black liquid was observed.

What has occurred? What do you think of these comments which the student made?

17. The 'apparently black liquid' contains both iodine and chlorine.
18. Two chemical reactions are occurring.
19. The orange solid is chemically unstable.
20. The black liquid which is formed when the iodine reacts with chlorine has the same composition as the black liquid which is formed when the orange solid is left in the air.
21. The black liquid resembles liquid bromine.
22. The proportion of chlorine in the orange solid is three times the proportion of chlorine in the black liquid.
23. Iron reacts with chlorine to give an orange solid.
24. Halogens react with one another.

SITUATION 4.





The diagram shows two tubes containing samples of gas trapped by a drop of mercury. The tubes are held in an oil bath which can be heated. The volume of the two gases at different temperatures can be measured. The graph shows the volumes of the samples over a range of temperatures.

The two gases behave differently. Why?

25. Sample B may be turned into a liquid by cooling.

26. At temperature T_3, there are more molecules of gas in Sample B than there are in Sample A.

27. Sample B undergoes a reaction when heated.

28. The graph shows that at temperatures higher than T_1, the two gases behave differently.

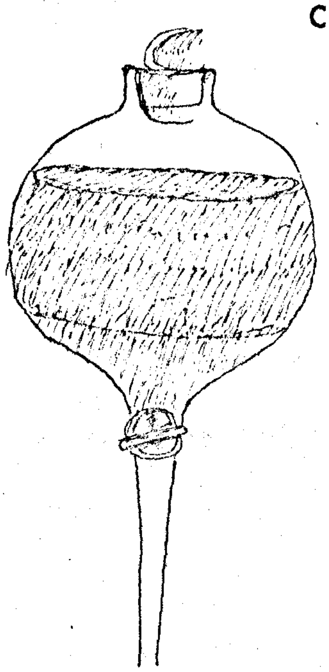
29. Sample B was dinitrogen tetroxide (N_2O_4).

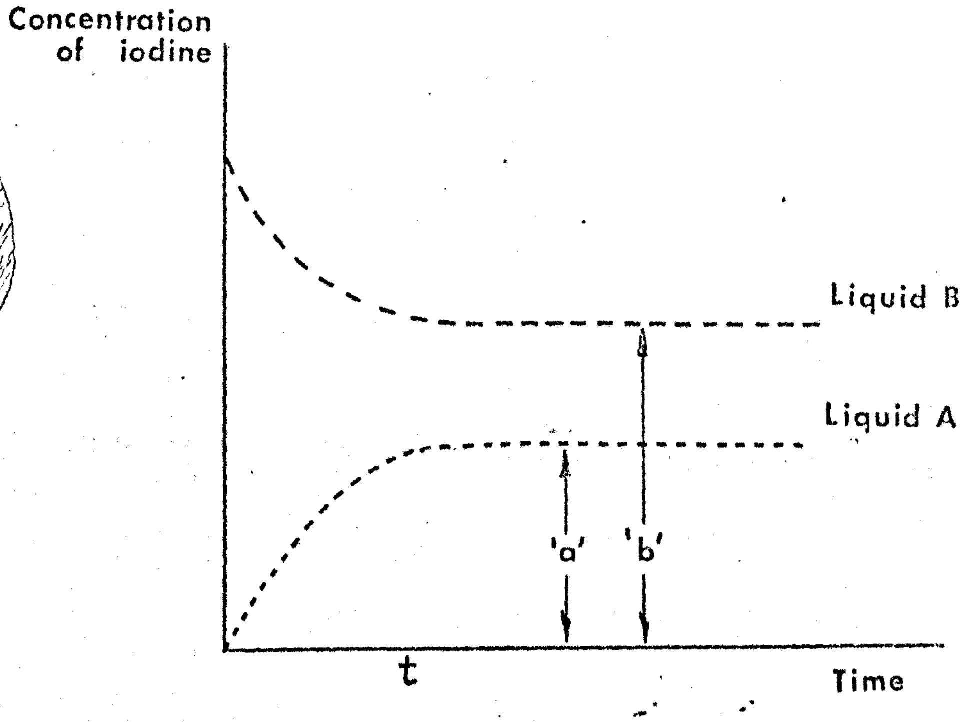
30. Both samples are gaseous at temperatures below T_2.

31. The gas hydrogen iodide would behave like Sample A at temperatures near 100°C and like Sample B at temperatures near 500°C.

32. The experiment was conducted at constant pressure.

SITUATION 5.





Liquid A and Liquid B do not dissolve in one another but iodine is soluble in them both. A flask which contained a litre of Liquid B with iodine dissolved in it was shaken with a litre of Liquid A for five minutes. The flask was allowed to stand and then 10 cm^3 samples of each liquid were withdrawn with pipettes and the concentration of iodine in these small volumes was determined. Shaking was resumed and after each period of five minutes further 10 cm^3 samples of each liquid were withdrawn and analysed. The readings which were obtained are shown in the diagram.

Some chemistry has occurred. What do you think of the following observations on it?.

33. The iodine is more soluble in one liquid than in the other.

34. The concentration of iodine in Liquid A increases until it reaches a maximum value.

35. If the experiment were to be repeated at a higher temperature the two curves would become level sooner.

36. The iodine comes to a state of equilibrium.

37. Bromine will behave similarly in the two liquids.

38. Some iodine is lost from the flask each time the 10 cm^3 samples are withdrawn.

39. The ratio of 'a' : 'b' is constant after time t.

40. When Liquid A is water and Liquid B is trichloromethane, the value of 'b' will be 40 times greater than 'a'.

In Part Two ....

.... of this test you will meet five further situations. This time experiments are suggested which might be undertaken to shed light on the problem which is posed.

Put an 'X' in the E column if the experiment which is suggested will lead to information BEYOND that already given, AND this information is RELEVANT to the problem stated.

Put an 'X' in the F column if the experiment which is suggested will lead to information BEYOND that already given, BUT this information is IRRELEVANT to the problem stated.

Put an 'X' in the G column if the experiment which is suggested will merely give information which is ALREADY AVAILABLE.

Put ONE 'X' ONLY in EACH GROUP of THREE boxes.

Do not assume that information from other experiments is available; consider each experiment according to its own merit.

SITUATION V.

Solution Q is made by dissolving 3g. of Solid Q in 100 cm^3 of water. When it is freshly prepared it is blue coloured. The colour fades when the solution has been left to stand in a 250 cm^3 conical flask on a windowsill.

The colour fades. Why? Assume that the colour of the solution will be re-examined some time after each experiment has been set up.

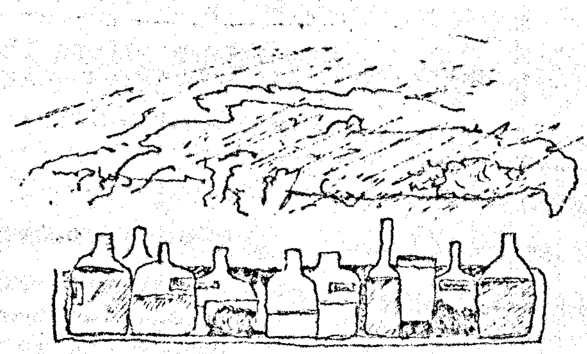
1. Dissolve Solid Q into water which has previously been boiled to remove the air which is normally dissolved in it and then stopper the bottle tightly.
2. Leave a fresh sample of Solution Q in a darkroom.
3. Prepare a solution of 3g. of Q in 100 cm^3 alcohol and leave it on the windowsill.
4. Dissolve 6g. of Solid Q in 200 cm^3 of water and leave in a 250 cm^3 conical flask on the windowsill.
5. Examine the behaviour of other blue chemicals.
6. Prepare Solution Q in a 250 cm^3 beaker and leave it on the windowsill.

SITUATION VI.

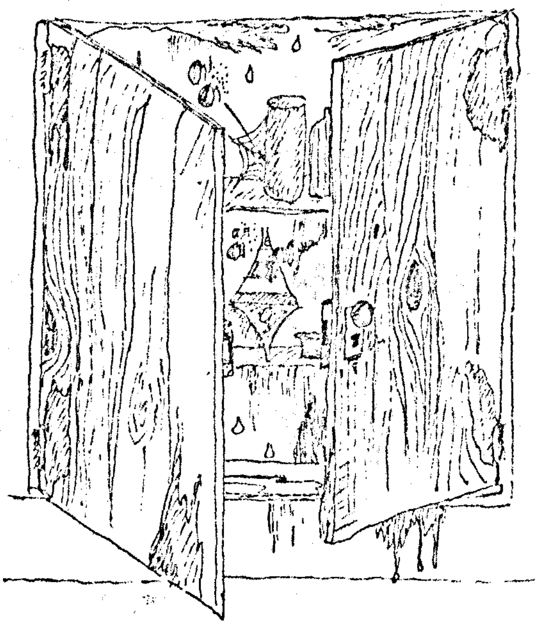
Two loosely stoppered bottles containing liquids in frequent use in the laboratory were placed in a rack. After a while a fine layer of smoke was observed and this gradually settled on the tray as a white powder.

Smoke is formed. Why? Assume that each experiment will be re-observed at a suitable time after setting it up.

7. Place the tray and bottles in a fume cupboard or fume chamber to see whether the smoke disperses.
8. Leave the tray and bottles in a cabinet containing dry air.
9. Stopper the bottles firmly and store them separately.
10. Remove both stoppers and leave.
11. Examine the solubility of the powder in water.
12. Empty the two bottles, then wash, dry and refill them. Stopper them loosely, shake and store them in the rack.



SITUATION X.

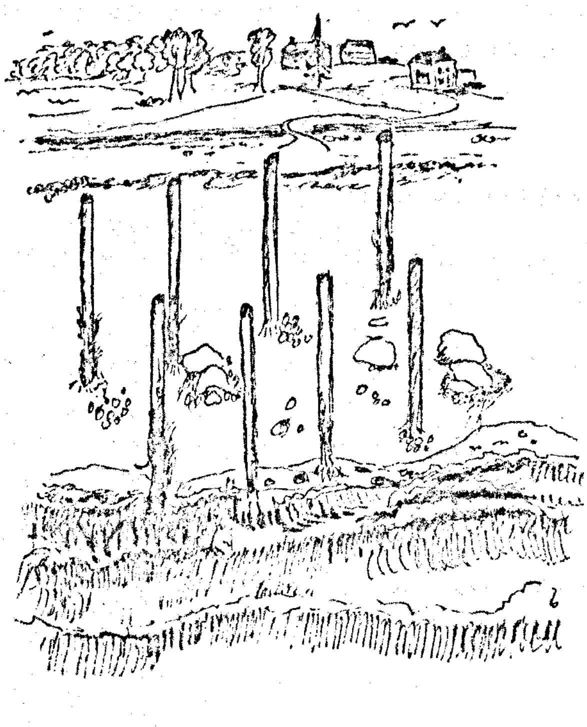


Two chemicals, A and B, were stored in a damp cupboard. A was a pure powder and it was in a paper bag. The crystals B were in a large open, waxed carton. After some days A was found to be fairly wet and B was still fairly dry, but the surface of the crystals were covered with white powder.

A and B have behaved differently. Why? You are to assume that the chemicals will be re-examined for change some time after the experiment has been set up.

13. Dry both chemicals before putting them away.
14. Store A in the carton and B in the bag.
15. Grind B to a powder before storing A and B together as before.
16. Measure the degree of humidity of the air inside the cupboard with a hygrometer.
17. Shake to remove the powder from the surface of B, dry A and store the chemicals as before.
18. Replace A with powdered sodium chloride and replace B with hydrated sodium carbonate ($Na_2CO_3.10H_2O$), and store as before.

SITUATION Y.

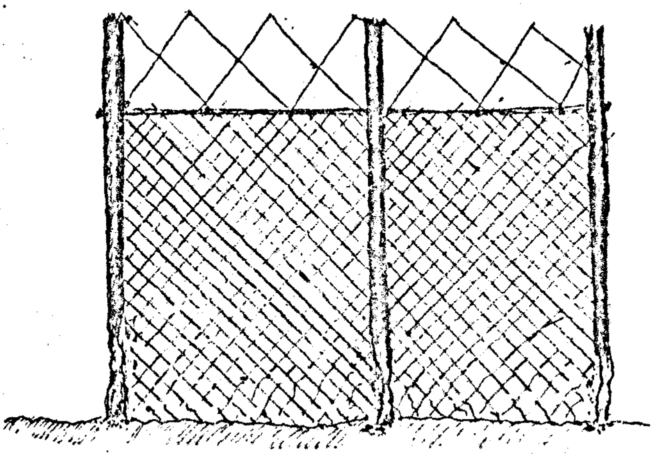


A line of iron poles was sticking out of a muddy beach between the low and high tide levels. Some poles were painted red and they were found to be well covered with barnacles; other poles were painted blue and they were found to be fairly free of barnacles.

The barnacles are unevenly distributed. Why?. Assume that a note will be made of the number of barnacles on each pole before the suggested experiment and another count will be made after a lapse of time.

19. Paint some iron poles with yellow paint and some with green and plant them between the tide levels.
20. Replace the iron poles with fresh ones which are identical with iron poles in all other aspects except that they are made of wood.
21. Paint some tall iron poles with the red paint and some with the blue and plant them below the low tide leve
22. Have the red and blue paints analysed to determine chemical composition.
23. Turn the poles upside down.
24. Plant a new iron pole close to each existing pole and paint the new poles to match the old.

SITUATION Z.



The painted iron posts which are used to support the netting round playgrounds and tennis courts rust far more rapidly near the ground (i.e. within 25 cm. of the ground) than they rust elsewhere.

The rusting is uneven. Why? Assume that one or more poles are treated in each of the following ways and they will be carefully examined for rust at suitable intervals of time thereafter.

25. Provide a shield to prevent rain water from splashing up from the ground on to the post.

26. Raise a pole and replant it with the buried section encased in polythene.

27. Replace an iron pole with another which is identical in all respects except in that it is made of aluminium.

28. Raise a pole and replant it the other way up.

29. Replant the poles with rust resistant paint.

30. Shorten a pole by cutting it through at 25 cm. above ground level and reset it in the ground.

TURN OVER for PART 111

Part THREE.

Here are FIVE more problems. On the piece of paper provided following on from Part II, devise TEN simple experiments similar to the ones given in Part TWO which you think would lead to RELEVANT information BEYOND that already given, i.e. experiments which you have given as 'E' in Part TWO. All your ten experiments may concern one problem only or they may be distributed between the five problems in any combination. Show which problem each of your experiments relate to by writing the appropriate letter in the margin. Outline experiments in as few words as possible and keep them simple.

Problem J.

You observe that all the bottles of a certain metal salt solution which have been on the shelves in your laboratory for a long time have a pale green material in the bottom.

What is the origin of this material?

Problem K.

You are supplied with some copper powder and you are surprised to see that it gives a gas when heated with dilute hydrochloric acid.

Why does this sample of copper react with dilute hydrochloric acid?

Problem L.

You notice that a block of ice in a metal can does not float to the surface as soon as the can is filled with water but does rise to the surface later.

Why is the ascent of the ice delayed?

Problem M.

You are told to prepare a solution of alcohol by mixing 50.0 cm^3 of water and 50.0 cm^3 alcohol. You do so carefully and find that you have only 96.0 cm^3 of solution.

Why is the total volume less than 100.0 cm^3?

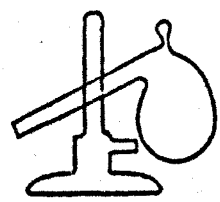
Problem N.

You notice that a drop of pink salt solution gives a steady supply of oxygen when added to some bleaching powder and the powder turns black.

What is the role played by the chemical in the pink salt solution?

Appendix G

Judgement Exercises - IV



CHEMICAL EXPERIMENTS

Inside this leaflet you will find a brief description of seven laboratory situations. Each one asks a question and then has some ideas about experiments which might be done to answer the question. If you think that the experiment would lead to information which is helpful in answering THAT question, please put a tick (✓) in the appropriate box on the answer sheet. If you think that the experiment is unlikely to be helpful please put a dash (-) in the box. You are not expected to be able to answer the question yourself; we only want to know if you think that the experiments would help you to discover the answer. Think particularly about the word which has been underlined.

Please assume that only one experiment is being done at a time; one thing only is being changed in each experiment.

Your answers to this exercise are going to be used to help in the design of teaching in Chemistry so do please take care. If you really have no idea, then you must make a guess, but do think carefully first. Work steadily and do not allow yourself to get too far behind.

Thank you very much for your help.

(A) You are told to mix 50 cm^3 of a certain solution of hydrochloric acid with 50 cm^3 of a solution of sodium hydroxide. You do as you are told and find that you have only 96 cm^3 of solution.

Why have you got less than 100 cm^3 of solution?

(1) Check on the accuracy of your measuring equipment.

(2) Vary the proportion of acid to alkali.

(3) Use a different acid.

(4) Use acid of different strength.

(5) Observe the temperature changes on mixing.

(6) Fit a condenser to ensure that there is no loss of water vapour.

(7) Weigh the two solutions before mixing and then again afterwards.

(8) Wait while the mixing flask cools.

(9) Repeat the experiment with alcohol as solvent instead of water.

(10) Mix 52 cm^3 of acid with 52 cm^3 of alkali.

(B) You take an ice cube from the refrigerator and put it into a stainless steel mug. You pour water into the mug and you find that the ice cube remains at the bottom of the mug for some seconds before it bobs up to the surface.

Why is the ice cube delayed at the bottom of the mug?

(11) Freeze the mug before putting the ice cube in it.

(12) Vary the time before pouring the water in.

(13) Use a jagged piece of ice.

(14) Use a glass.

(15) Roughen the inside of the mug.

(16) Paint the inside of the mug.

(17) Use a liquid other than water.

(18) Use a stainless steel mug of another shape.

(19) Allow the ice to stand in a warm room before using it.

(20) Fill the mug with water then add the ice.

(21) Use warm water.

(C) A small glass tube was hung in a horizontal position on a cotton thread. It was then filled with tiny crystals of one of the salts of the element manganese and both ends of the tube were plugged with wax. The tube was found to twist round on its thread when a magnet was moved up to it.

What is the cause of this twisting movement?

(22) Using magnets of different strengths.

(23) Use a tube made of iron.

(24) Use another salt of manganese.

(25) Use a salt of manganese in another oxidation state (valency).

(26) Use plugs made of cork.

(27) Hang the tube vertically.

(28) Dissolve the salt in water before filling the tube.

(29) Use salts of other metals.

(30) Use a single large crystal of the salt and no tube.

(D) If iron tubes are used to support the netting round a playground or tennis court, they appear to rust fairly evenly; if iron posts with an L-shaped section are used they rust first near the ground.

Why do the two types appear to behave differently?

(31) Protect both types of support from being splashed.

(32) Provide the posts with protective paint.

(33) Fix guy ropes to the posts to prevent them from bending in the wind.

(34) Examine the nature of the metal used in each type of support.

(35) Provide a protective layer round the buried part of all the supports.

(36) Ensure that dogs do not have access to the posts.

(37) Use plastic coated netting.

(38) Fix a tight cap on the tops of the tubes.

(39) Drill a hole in each tube at ground level to let rain water out.

(40) Examine the inside of the tubes.

(E) Some iron poles were planted between the low tide mark and the high tide mark on a sandy beach. The poles which had been painted blue were found to have many barnacles and the poles which were painted green were found to have few.

Why are the barnacles unevenly distributed?

(41) Vary the height of the poles.

(42) Exchange the coloured iron poles for coloured wooden ones.

(43) Examine the chemicals used in the paint.

(44) Exchange the poles for others painted in red and yellow.

(45) Mark the barnacles from the blue poles with a blue spot and those from the green poles with a green spot and release them.

(46) Study the changes in water temperature.

(47) Paint the poles with blue and green paint made by a different manufacturer.

(48) Turn the poles upside down, replant them and then repaint them.

(49) Coat the poles with blue and with green plastic.

(F) A student noticed that four of the bottles of manganese sulphate solution in the school laboratories contained samples of a soft and green material, and that a fifth bottle of the solution did not have the solid.

What is the origin of the green solid material?

(50) Filter the material and leave the solutions as before.

(51) Cork the bottles more tightly.

(52) Examine the paint on the shelf beneath the bottles.

(53) Examine the glass of the bottles.

(54) Discover how long each bottle had been in use.

(55) Leave bottles containing fresh solution in the dark.

(56) Dilute the contents of a contaminated bottle, shake and leave it as before.

(57) Store the uncontaminated bottle in a warm place.

(58) Examine bottles of other solutions.

(59) Ask a friend at another school to look at the bottles of manganese sulphate solution in his laboratories.

(G) A student removed the cork from an empty bottle and then cut off the bottom cleanly. He fixed a lighted candle to the bottom of a bowl, poured water round it and stood the bottle over the candle. He paused and then pushed the cork in firmly. The candle continued to burn inside the bottle for a short while and it then went out slowly. The water level inside the bottle rose until the bottle was about 1/5th filled with water. (About 15% of the gas which remained in the bottle was found to be oxygen.)

Why was the bottle filled to the 1/5th level?

(60) Pause for longer before pushing the cork into the neck of the bottle.

(61) Use a larger bottle.

(62) Use a longer candle.

(63) Use a longer wick.

(64) Use a candle made of a different wax.

(65) Put more water into the bowl.

(66) Use warm water in the bowl.

(67) Use a liquid other than water.

(68) Enrich the air in the bottle with oxygen.

(69) Enrich the air in the bottle with carbon dioxide.

(70) Experiment without the cork.

Responses to JE-IV

Code E : helpful — E_1 : one step) in hypothesis chain

F : unhelpful — E_2 : two steps) in hypothesis chain

G : already available — E_3 : three steps) in hypothesis chain

A	B	C	D	E	F	G
					50 G	60 E_3
1 G	11 E_2		31 G	41 G	51 G	61 G
2 G	12 E_2	22 G	32 F	42 F	52 F	62 E_3
3 F	13 E_2	23 F	33 E_2	43 E_2	53 F	63 E_3
4 E_2	14 E_2	24 E_1	34 G	44 F	54 E_1	64 E_3
5 E_2	15 E_2	25 E_3	35 G	45 E_2	55 E_1	65 G
6 F	16 E_3	26 F	36 F	46 F	56 F	66 G
7 F	17 F	27 F	37 F	47 E_2	57 E_1	67 E_1
8 G	18 G	28 E_2	38 E_1	48 G	58 E_1	68 E_3
9 E_2	19 E_3	29 E_3	39 E_1	49 G	59 G	69 E_3
10 G	20 G	30 E_2	40 E_1			70 F
	21 E_1					

A Here is an example:

A man broke a single piece of Cadbury's chocolate from a new bar and put it into his mouth. He bit on it hard and he then experienced a degree of pain which was sufficiently intense to cause him to remove the fragments of the chocolate from his mouth. The man thought that the pain was associated with his teeth even though he had had no dental trouble for sometime. Examination of the fragments revealed a very small piece of 'silver' wrapping foil.

<u>What was the cause of the pain?</u>

(a) examine his mouth for a piece of broken tooth, a cut or an ulcer.
(b) feed the man another piece of chocolate with no foil on it.
(c) examine the man's mouth for dental stoppings.
(d) place a clean nail on his tongue.
(e) give the man a piece of Fry's chocolate with a small piece of foil adhering to it.
(f) feed the man a piece of warm chocolate with a small piece of foil.
(g) try the man on a double sized piece of chocolate with a tiny piece of foil on it.

	Correct answer	
(a)	x	The man is obviously in good health.
(b)	✓	The foil is likely to be involved.
(c)	✓	The stoppings contain a base (less reactive) metal and are likely to be involved.
(d)	✓	A nail contains a more reactive metal (iron) than the metal in the dental stopping (mercury). Hence a little electric cell is established in his mouth.
(e)	x	The manufacturer is irrelevant.
(f)	x	The hardness of the chocolate is unlikely to be a factor.
(g)	x	The presence or absence of chocolate - and boiled sweet - is irrelevant.

When you have finished the test chose any one or more of the <u>problems</u> given (including the example) and suggest as many as ten more experiments which you think would be worth doing. The ten may all be concerned with the same experiment or with any combination of experiments. Write no more than a few words on each of your experiments (as was done in the test).

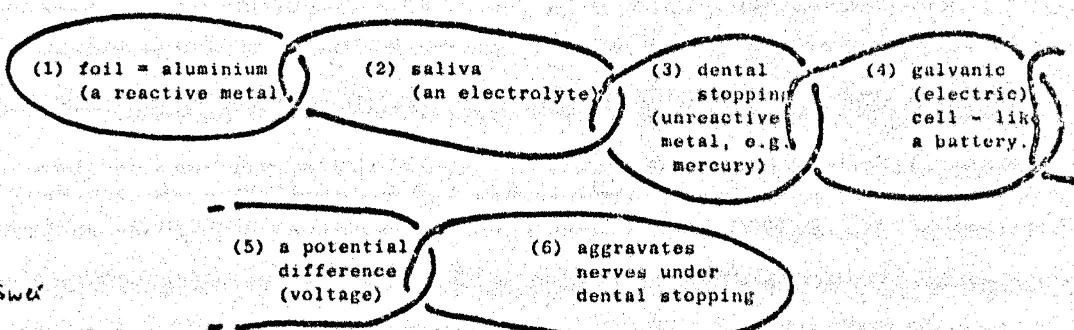
B & D Here is an example:

A man broke a single piece of Cadbury's chocolate from a new bar and put it into his mouth. He bit on it hard and he then experienced a degree of pain which was sufficiently intense to cause him to remove the fragments of the chocolate from his mouth. The man thought that the pain was associated with his teeth even though he had had no dental trouble for sometime. Examination of the fragments revealed a very small piece of 'silver' wrapping foil.

<u>What was the cause of the pain?</u>

(a) examine his mouth for a piece of broken tooth, a cut or an ulcer.
(b) feed the man another piece of chocolate with no foil on it.
(c) examine the man's mouth for dental stoppings.
(d) place a clean nail on his tongue.
(e) give the man a piece of Fry's chocolate with a small piece of foil adhering to it.
(f) feed the man a piece of warm chocolate with a small piece of foil.
(g) try the man on a double sized piece of chocolate with a tiny piece of foil on it.

Try to think of the problems in terms of chains of interconnected factors. Here is a possible chain for the chocolate problem:-



answer

Tick only the experiments which confirm the links of the chain <u>Thus</u>:

✓	(b)	is <u>good</u> because we are interested in whether aluminium has to be present <u>or</u> not (link (1)).
✓	(c)	is <u>good</u> because we want to know whether the stopping (unreactive metal) is important (link (3)).
✓	(d)	is <u>good</u> because if iron will cause pain regardless of whether chocolate is present or not our proposed chain is probably correct.
✗	(a)	, (e), (f) and (g) are irrelevant if our chain is correct.

<u>Before</u> you fill in your answers to each of the problems put in a chain (or more than one chain) with as many links in it as possible in the space below. Even two links may well be useful. <u>Then</u> put a tick or a cross in the boxes.

C & D

Here is an example:

A man broke a single piece of Cadbury's chocolate from a new bar and put it into his mouth. He bit on it hard and he then experienced a degree of pain which was sufficiently intense to cause him to remove the fragments of the chocolate from his mouth. The man thought that the pain was associated with his teeth even though he had had no dental trouble for sometime. Examination of the fragments revealed a very small piece of 'silver' wrapping foil.

What was the cause of the pain?

(a) examine his mouth for a piece of broken tooth, a cut or an ulcer.
(b) feed the man another piece of chocolate with no foil on it.
(c) examine the man's mouth for dental stoppings.
(d) place a clean nail on his tongue.
(e) give the man a piece of Fry's chocolate with a small piece of foil adhering to it.
(f) feed the man a piece of warm chocolate with a small piece of foil.
(g) try the man on a double sized piece of chocolate with a tiny piece of foil on it.

Try to decide whether each experiment is -

(a) relevant or not
and (b) useful in that it provides new information.

Help yourself to do this by assigning each experiment a letter from this matrix.

		the experiment is relevant	the experiment is not relevant
the experiment does	provide new information	E *	F
the experiment does	NOT provide new information	G	H

* An experiment is only worth doing if it has an 'E' classification.

Experiment (a) falls into category F - the information as to whether he has cuts and ulcers is new but it seems unlikely to be relevant.

" (b) " " " E - the foil is made of aluminium which is a reactive metal and may form the anode of a small cell (battery).

" (c) " " " E - the unreactive metal (mercury) in the stopping may form the cathode of this cell. (Saliva forms the electrolyte).

" (d) " " " E - the iron of the nail is sufficiently reactive to form a cell with the stopping.

" (e) " " " H - the manufacturer is irrelevant. Both Cadbury and Fry market a high quality product.

" (f) " " " F - the man will be able to bite the chocolate more easily if it is soft but he will experience the pain regardless of how hard the chocolate is.

" (g) " " " G - obviously no new information will be provided.

Assign one of these four letters to each of the experiments.

Appendix H

End-of-Year Chemistry Examination
IIIrd Forms

IIIrd FORM CHEMISTRY EXAMINATION. Your name & initials: . . . .

Summer 1976. Your form: . . . .

Do ALL the questions. Answer Part I on the question paper; answer Part II on lined paper.

PART I. Marks.

1. Two dyes can be separated by . . . . (1)

2. The percentage by mass of hydrogen in water is . . . . (2)
(H = 1, O = 16)

3. A catalyst is . . . .

. . . . (2)

4. How many times greater is the mass of a phosphine (PH_3) molecule than the mass of an ammonia molecule (NH_3) ? . . . . (2)
(H = 1, N = 14, P = 31)

5. - 10. The table below is a modified form of the Periodic Table for the first twenty elements. The letters X, Y, Z etc. are code letters and do not represent the actual chemical symbols of the elements. Use the code letters when answering questions 5. to 10. (inclusive).

I	II	III	IV	V	VI	VII	VIII
						X	U
		Y				W	
	Z						

5. The formula of the sulphate of Y is . . . . (2)
(The sulphate ion has the formula SO_4^{2-})

6. The electronic configuration of the W^- ion is . . . . (2)

7. The formula of the compound formed between X and Z is . . . . (2)

8. An important feature of the chemistry of U is that it is . . . .

. . . . (1)

9. Write an A in the table for a reactive metal with a valency of 1 . (1)

10. Write a B in the table for a reactive non metal with a valency of 2 . (1)

11. A green powder gave off a gas readily and in the cold when the powder was treated with dilute sulphuric acid. A blue solution remained.
The green powder could be . . . . (2)

12. A single covalent bond is formed between two atoms of chlorine in a chlorine molecule. This is done by each atom . . . .

. . . . (2)

13. Solid sodium chloride does not conduct electricity because . . . .

. . . . (2)

14. Bromine and chlorine are in the same group of the Periodic Table. They resemble one another in the following three respects:-

(i) . . . .

(ii) . . . . (3)

(iii) . . . .

15. When iodine crystals are heated in a dry tube, a purple vapour is observed and crystals of iodine appear near the neck of the tube.

This behaviour is described as . . . . (1)

16. A convenient reaction which could (or is) used for the preparation of hydrogen in the laboratory is the reaction between -

. . . . and . . . . (2)

17. A metal oxide which can easily be reduced to metal by heating in hydrogen is

. . . . (1)

18. A solid which loses its water of crystallisation to the atmosphere is said to -

. . . . (1)

19. When 4.0 g of calcium are burnt in chlorine to form calcium chloride according to the equation

$$Ca + Cl_2 = CaCl_2$$

the mass of calcium chloride which could be formed is . . . . (2)

(Cl = 35.5, Ca = 40)

20. Atoms of elements in Group II of the Periodic Table form ions by

. . . . (2)

21. The total number of atoms present in one molecule of propanone ($(CH_3)_2CO$) is . . . . (1)

22. The information needed to convert an empirical formula into a molecular formula is -

. . . . (1)

23. The relative molecular mass of carbon dioxide is 44. This means that

one molecule of carbon dioxide is . . . .

. . . . (2)

24. The best evidence that could be produced to show that a compound is ionic is -

. . . .

. . . . (2)

25. A good laboratory test for oxygen is . . . .

. . . . (2)

Appendix H

End-of-Year Chemistry Examination
IIIrd Forms

PART II

I. Show by means of a labelled diagram how you would prepare a few gas jars of chlorine in the laboratory. Suppose that you then wished to prepare gas jars of another gas which was light, soluble in water and alkaline. What changes would you make in your apparatus? (6) (3)

II. What would you see when you drop a piece of calcium into clean water containing litmus solution? What do your observations tell you about the hydroxide of calcium? (3) (2)
Can you think of one reason why phosphorus can be used in an experiment to determine the percentage of oxygen in air while sulphur cannot be used? (1)

III. Give two reactions in which hydrogen chloride (or dilute hydrochloric acid) and chlorine differ. Give the products of the reaction in each case. (This may be answered in the form of a small table.) (4)

IV. Describe and explain one experiment which you would use to show that one metal is a little higher in the electrochemical series than another metal. (3)

V. Why isn't sulphur included in the electrochemical series? (2)
The symbol Na stands for the element sodium and the symbol Na^+ for the sodium ion. One of these reacts with water and a gas is observed, the other dissolves in water and no gas is seen. Can you explain this difference in behaviour? (3)
Suggest three reasons why the metal mercury has so many more uses than the metal sodium. (3)

VI. How would you prepare a good crystalline sample of zinc sulphate from impure zinc. (6)

VII. Give brief outlines of separate experiments which would enable you to:-

(a) discover whether sodium sulphate was soluble if all the usual laboratory reagents were available except sodium sulphate. (2)

(b) distinguish a sample of hydrochloric acid from one of nitric acid. (2)

(c) establish which of sodium nitrate and sodium chloride was the more soluble at 40°C . (4)

(d) discover whether manganese dioxide or lead dioxide was the better catalyst for the preparation of oxygen from hydrogen peroxide. (4)

VIII. Give <u>three</u> important reasons why some metals are comparatively cheap to buy and others are expensive. (3)

IX. Chemicals Q and V are formed when Chemical R is heated in a dry test tube. Chemical Q turns out to be a white, insoluble solid. Chemical V is a coloured, poisonous gas. Can we predict that Chemical R will be -

(a) insoluble, (b) a compound, (c) coloured and (d) poisonous? (3)

(Simply answer 'yes' or 'no' without giving reasons.)

X. Tin (II) chloride ($SnCl_2$) is a solid which will dissolve in water. The solution will conduct electricity. Tin (IV) chloride ($SnCl_4$) is a liquid which will dissolve in tetrachloromethane (CCl_4) and this solution does not conduct electricity. How do we explain these differences in behaviour of these two compounds? (4)

21. When a force acts on an object the effect it has can be one or more of the following:

(i) The object is made to start moving or increase its speed.

(ii) The object is made to slow down or stop moving.

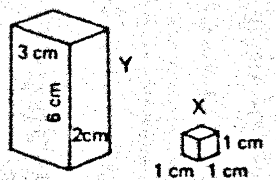
(iii) The object remains at rest or moves steadily because the force balances other forces acting on it.

(iv) The size or shape (or both) of the object is changed.

Which two are involved when a stone tied to the end of a piece of string is whirled round in a circle?

(A) i) + ii), (B) i) + iii), (C) i) + iv), (D) ii) + iii)

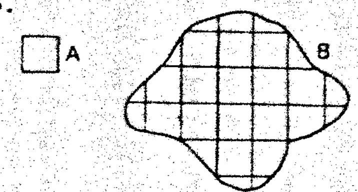
22. X is a wooden building brick which is a cube 1cm long, 1cm wide and 1cm high.



How many of these would it take to build Y, a solid tower 6 centimetres high, 3 centimetres long and 2 centimetres wide?

(A) 36 (B) 12 (C) 48 (D) 18

23.



If the area of the square A is called one unit, how many units of area are enclosed by the shape B?

(A) 9 (B) 17 (C) 26 (D) 20

24. A food chain shows how one kind of animal depends on others for its food. For example, A shows how thrushes feed on snails, which in turn feed on green plants.

A
Thrushes ⟶ snails ⟶ green plants

B is another food chain in which there is a gap.

B
Blue-tits ⟶ ? ⟶ greenflies ⟶ plants

Which of these is most likely to be the missing link in chain B?

(A) butterflies (C) ladybirds

(B) worms (D) none of these

25. A food chain shows how one kind of animal depends on others for its food. Why do all food chains end in plants? It is because:

(A) more animals eat plants than eat other animals

(B) plants are not living things as animals are

(C) plants make their food from non-living things

(D) plants grow more quickly than animals

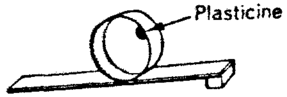
26. When a thin shaving of cork is examined under a microscope a regular pattern of cell walls can be seen, rather as in the diagram.

This means that cork:

(A) was once living (C) is very old

(B) none of these (D) was man-made

27. A light cardboard ring has a piece of plasticine, heavier than the ring, attached to it. The ring is held on a rough sloping surface in this way.



When it is released, it will:

(A) remain still (C) roll down the slope

(B) roll up the slope (D) spin round but not roll

28 Some of a yellow powder was added to a colourless liquid and shaken up well. A cloudy liquid was produced, and after it had been allowed to stand for a while a yellow sediment formed at the bottom and the liquid became clear but slightly yellow in colour. This meant that:

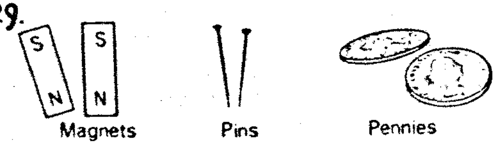
(A) the solid did not dissolve in the liquid

(B) the liquid must have been water

(C) the solid would dissolve in water but not in the liquid

(D) a little of the solid dissolved in the liquid

29.



Suppose someone tries to magnetise one of the pins. In order to tell whether it had become a magnet or not, which one of these things would you test it on?

(A) one of the magnets (C) any of these

(B) the other pin (D) one of the pennies

(30 → (see Answer Sheet))

Appendix J

Science Orientation Questionnaire

Please will you find out by answering the following questions. Give yourself a rating like this:-

Give +2 for much pleasure

+1 for some pleasure

0 fair or no strong views

-1 for not much pleasure

-2 no pleasure at all

1. Do you look back on your science classes with pleasure?

2. Do scientific pursuits (model railways, slot cars, practical botany, geology, wine making, electronics, etc.) give you pleasure?

3. Do you get pleasure from working with tools (as in carpentry, soldering, sewing, weaving, in the kitchen, etc.)

4. Would you (do you) get pleasure from scientific magazines (The New Scientist, Scientific American, The Ecologist, The Dietician, The Food Technologist)

5. Do you get pleasure from practical classes for their own sake, i.e. not for reasons which relate to whether or not they are easier or make a change etc.?

6. Do you enjoy science programmes on television?

7. Do you think that you would get pleasure from a career which is scientific (medicine, nursing, engineering, veterinary, agriculture, etc.)?

Appendix K

Rotated Factor Analysis of Correlation Coefficients of Inter Item Scores (N = 45/IV/1976)

Factor No.	I	II	III	IV	I	II	III	IV	I	II	III	IV		I	II	III	IV	
Item Number: 1	59						36					54		52				
2	54				(49)	58			(35)		-50					49		
3		56						57				72		51				
4			71		50							26	(weak)		(45)	49		
5	69						79		46							50		
6	76				(49)	57			41					64				
7	62				(47)				64							56		
8	68					65							(weak)	82				
9	56				66				63					(47)		51		
10				59				66	70	(44)						47		
11				56	77						47			54				
12						40			67							(43)	52	
13	49	(48)			(34)	(36)	36	(31)			74						37	
14	87				73							46		45				
15		42					46				39				38			
16	34						49				42					(24)		(weak)
17	64					48				72				65				
18		77			72					65				50			(41)	
19		50			66				46		(44)						(31)	(weak)
20	44		(39)					43	53					50	(49)			
21	81				54						60						65	
22	(34)	39					70			(-35)		52			76			
23			53		60			55	36		69				53			
24				63										57				
Eigen values	8.40	2.46	1.94	1.53	6.82	2.51	2.42	1.70	5.58	2.32	2.19	2.0		5.73	2.82	2.17	1.68	
Cumul. %age	35.0	45.2	53.3	59.7	28.4	38.8	48.9	56.0	23.3	32.9	42.0	50.3		23.9	35.6	44.6	51.7	