The Nature of Cognitive Style and its Relationship to Educational Performance

by

Timothy John Lowe

*****

A thesis submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy (Education) in the Faculty of Social Sciences

***

University of Aston in Birmingham

November, 1973

ACKNOWLEDGEMENT

To Professor A. G. Joselin, whose advice, encouragement and criticism have ensured the completion of this thesis, the author is deeply indebted.

He would also like to express thanks to the staff of the Department of Education, and especially to Dr. N. C. Graham and Miss M.A. Small for their co-operation in the administration of of the empirical studies.

Also to his wife, for her tolerance in matters emotional, and for her help in matters intellectual and secretarial.

Finally, the author is very grateful to the students of the Department of Education, and in the Complementary Studies Programme, who were the "subjects" of the investigation.

****

SUMMARY OF THESIS

A survey is made of the literature relating to a number of dimensions of cognitive style, from which it is concluded that cognitive style has a strong theoretical potential as a predictor of academic performance. It is also noted that there have been few attempts to relate cognitive style to academic performance, and that these have met with limited success. On the assumption that theories of individual differences should be congruent with theories of general functioning, an examination is made of the model of cognition presupposed by dimensions of cognitive style. A central feature of this model is the distinction between cognitive content and cognitive structure. The origins of this distinction are traced back to the normative and experimental or quasi-experimental characteristics of research in psychology. The validity of the distinction is examined with reference to modern research findings, and the conclusion is drawn that the normative experimental method is an increasingly inappropriate tool of research when applied to higher levels of cognitive functioning, as it cannot handle subject idiosyncracy or patterns of interaction. An examination of the presuppositions of educational research leads to the complementary conclusion that the research methods imply an over-simplified model of the educational situation. Two empirical studies are reported : (1) An experiment using conventional cognitive style dimensions as predictors of performance under two teaching methods (2) An attempt to predict individual differences in overall academic performance by means of a research technique which uses a questionnaire, intra-individual scoring, and an analysis of patterns of responses, and which

attempts to take some account of subject idiosyncracy. The implications of these studies for further research are noted.

CONTENTS

....

CHAPTER I

THE SCOPE OF THE STUDY

The study of cognitive style is of relatively recent origin, developing out of the American cognitive tradition of the late 1940's and early 1950's : it was one symptom of the resurgence of interest in "the mind", "cognition" and other intervening variables, concepts which had been almost ignored for thirty years as a result of the influence of Behaviourism and Positivism.

As with other topic areas in psychology, it is not possible to define cognitive style in any categorical way, and the borders with neighbouring phenomena are blurred or imprecise, and change with fashions in theory and definition. Nevertheless, cognitive style is clearly something to do with cognition, and it is clearly something to do with individual differences. Let us therefore take as a point of departure the following definition:

> "Cognitive styles are individual differences in cognitive functioning."

Were this definition subjected to public discussion it would undoubtedly be criticised for over-inclusion rather than over-exclusion. This definition would classify the large volume of research into intelligence and ability as aspects of cognitive style, whereas researchers into cognitive style would not consider this to be their domain. Warr (1970) suggests that studies of cognitive abilities are concerned with how well a person can think, whereas studies of cognitive styles are concerned with how a person habitually does think. This seems a valid distinction when applied to the scaling and measurement of individual cognitive abilities, but is perhaps less convincing when applied to research into the nature of cognitive abilities, and into biases between one ability and another within the same person. Thus studies of the nature of intelligent functioning (as compared with instinct or habit) are equally a part of "cognitive ability" and "cognitive style". The same can be said for the study of intelligent functioning vis-a-vis creative ability : it would seem inappropriate to classify studies of creativity and intelligence as "cognitive ability", but studies of the conjunction of the two in one person (convergent vs. divergent thinking) as "cognitive style".

Furthermore, the definition of "cognitive style", above, would lead one to include some aspects of personality under the rubric of "Cognitive style". Indeed, Warr (1970) names his useful collection of readings in cognitive style "Thought and Personality". Personality is at least partly the study of individualy differences in response to the social environment, and, if a cognitive theorist, one would assume that such responses are mediated by knowledge, models, or cognitive structures, representing the social environment. The relationship is clear; Warr (1970) admits this relationship, but points out the distinction between content and structure, a distinction which partly separates the two topic areas. Much of the work on personality has dealt with <u>what</u> a person thinks, wants or feels, whereas cognitive style is usually concerned with <u>how</u> a person thinks, wants, or feels. This, Warr points out, would distinguish between authoritarianism, which deals with content and is thus a personality variable, and dogmatism, which is structural and hence a cognitive style variable. This distinction is only partial though: other views of personality, such as Allport's (1961) concept of the personality as an organic totality, and the organising factor which makes the pattern of habit, experience, and inherited characteristics a unique but holistic person; this type of concept is clearly structure as much as content. Furthermore, although the writer recognises that cognitive style research has been primarily into the way a person thinks, Chapter 4 will argue that the content/structure distinction is inappropriate and misleading when applied to cognition. Without attempting to prejudge that argument here, therefore, the distinction must be rejected as part of the definition of cognitive style.

Warr (1970), himself, offers no definition of cognitive style, perhaps wisely. And indeed, the writer has only found two earlier definitions. Cropley (1967) defines it as "the characteristic way in which an individual goes about taking in information from the world". This definition is clearly based on the information processing tradition, and in its parent the empiricist tradition. It strongly, perhaps even exclusively, emphasises the perceptual basis of cognitive style. As such it fits the early studies into perceptual style (field dependence) and categorising style (cognitive complexity) very well. It is less appropriate to styles which are more at the motor end of things, such as reflection-vs-impulsivity, which deals with speed of cognitive response, and convergent-vs-divergent thinking, which deals with type of processing. For this reason, it will not be adopted here.

Gardner, Holzman, Klein, Linton and Spence (1959) define the earlier and similar concept of "cognitive control" as "dimensions of individual difference in cognitive structures that mediate the expression of particular intentions when a person is confronted with particular classes of stimulus conditions." This definition clearly involves "individual differences in cognitive functioning", but makes additional restrictions as to the set of phenomena considered. In particular, the phrase about "particular intention......... with particular classes of stimulus condition" appears to imply that cognitive controls are at least partly context-bound. The writer would not argue with this implication, which, however, is apparently dissonant with the maintenance of the structure-content distinction in the definition. In any case, the appropriateness of the content-structure distinction has not yet been argued here, and it would be unwise to include in the definition an unnecessary assumption. A second restriction is that cognitive controls are "dimensions of individual difference". Here again, it would seem unwise to specify that the tools of analysis be dimensions until the predictive power of other forms of hypothetical construct have been investigated. The "type", the "schema" and "habit Hierarchy" are three such constructs which spring to mind, and which have not been used in the context of modern taxonomic techniques.

The provisional definition in terms of "individual differences in cognitive functioning" will therefore be left standing. It is true that it is over-inclusive, but in the present state of theoretical ignorance, the plea is made that over-exclusion is a greater sin than over-inclusion. Over-exclusion is liable to cause distortion of theoretical concepts and interpretations of results, whereas over-inclusion merely makes it difficult to delineate a thesis or study.

In the absence of theoretical distinction, the present thesis will rely on empirical delineation. The first aim of the study was to examine the predictive potential of cognitive style variables in the higher education situation. The initial task was therefore to build up a battery of cognitive style tests. The battery should obviously include the most well-known and heavily researched tests, but should also be varied in its sampling of cognition and should avoid duplication of similar tests. It should have a high a priori chance of obtaining significance, as insignificant results in the social sciences may be equally due to unwanted and uncontrolled variance as to lack of experimental variance. Finally, the battery should not be too time-consuming, as it is usually impractical to test students for much more than three hours. Fortunately, cognitive style tests do not tend to be long, so a three-hour battery can include eight to ten tests.

Thus a preliminary survey of the literature was made, and a battery of tests was selected as follows:

1. Cognitive complexity, cognitive differentiation and authoritarianism have all been widely researched, and are of major importance as cognitive style dimensions. Measures of these dimensions were therefore included.

2. In an attempt to sample a wide variety of cognitive phenomena, the lesser known dimensions of strong-vs-weak automatisation, and convergent-vs-divergent thinking were included. An additional factor in the selection of the latter was that I.Q. is one of the component scores: as the predictive

power of intelligence in higher education has been well researched, this measure provides a good point on which to anchor any results. Because of the need for variety, and its converse, the need to avoid duplication of similar tests, a measure of dogmatism, which is similar to authoritarianism, was not included.

3. Tests which lacked a priori predictive potential, or theoretical clarity were excluded. Hence equivalence range and category width, although reviewed briefly, have not been used as predictors. Hence also repression-vs-sensitisation, although mentioned in the text in view of its relevance to cognitive defences, was excluded from the battery.

The review of cognitive style literature has also followed this rationale, and aims to be an exploration of the theoretical nature, and the predictive potential to education, of the above cognitive style dimensions. The attempt has been made to cover the width of cognitive style dimensions while omitting, for reasons of space, the plethora of little known scales that have accumulated. Many of these attempt to measure qualities which are tangential to other qualities reviewed here.

A second aim of the thesis was to examine the nature of cognitive style. This was a vague aim, which has been interpreted in terms of the nature of the analytical concepts of cognitive style. An examination of these has led back to the analytic concepts of general cognitive functioning, and to the research methods used in investigating these phenomena. Such an examination clearly cannot be both deep and broad, and it was thought that in this case deep is better than broad. The attempt is therefore made to examine some of the presuppositions of research into cognition and cognitive style, and to take this to the maximum depth encompassed by psychology: no attempt is made to go into the philosophical issues which arise. In lieu of a broad approach, the analytic concepts are examined in the context of three "isolated" topics: perception, memory, and higher cognitive functioning. The literature survey here is obviously even more arbitrary, and dependent on the biases and limited experience of the writer. It is to be hoped that the resulting arguments and the resulting conclusions are not jeopardised as a result.

Finally, the empirical section gives a report of the two empirical studies performed, and discusses their implications for the prior theoretical analysis, and for further research.

CHAPTER 2

COGNITIVE COMPLEXITY AND COGNITIVE DIFFERENTIATION

The complexity of the cognitive system is a concept which underlies a variety of cognitive style dimensions. Besides a dimension which is usually called cognitive complexity, and which is measured by Bieri's (1955) Modified Role Construct Repertory Test (Mod RCRT), the dimension of concreteness vs.abstractness (Harvey, Hunt, and Schroder, 1961) is theoretically almost identical. The concept is also embodied in Witkin's dimension of field-dependence vs.field independence (later generalised and renamed articulated vs.global functioning), and is related to authoritarianism (Adorno, Frenkel-Brunswick, Levinson and Sanford, 1950), dogmatism (Rokeach, 1960), equivalence range (Gardner, 1953) and category width (Pettigrew, 1958), although these styles originated in the study of other phenomena. The two approaches to cognitive complexity, and the field dependence dimension, are the styles which have been most thoroughly researched, and they are considered in this chapter.

2:1 The Presupposed Cognitive System

The dimensions discussed in this chapter have been developed from a variety of views of cognition. However, they all originate in the American cognitive tradition of the early 1950's. Lewin, 1936, Kelley, 1955, Werner, 1940). And despite minor theoretical differences, there is clear agreement on some aspects of cognition:

1. <u>It is possible</u> to make a valid distinction between the structure of cognition and the content of cognition. Bieri, Atkins, Lobeck, Miller and Tripodi (1966) state:

"First cognitive structures refer to organised systems whose properties are dependent upon the inter-relations of the various elements in a given system. Second, knowledge of cognitive structure implies that predictions can be made of the way in which the person copes with his environment."

Scott (1963) further amplifies this, and suggests that the concern for cognitive structure is based on a number of assumptions:

1. The content of experience is organised into structural assemblies from which any element of content derives its significance.

2. The way in which any new experience is reviewed, processed, and interpreted depends on the capabilities and characteristics of the pre-existing cognitive structure into which it is read.

3. While the contents of cognition may be endlessly varied, structural properties can be described in a limited number of genotypic terms, thereby permitting a more parsimonious formulation of psychological processes.

4. The contents of cognition develop from social norms and other fortuitous experience which cannot be well predicted from personality theory: they may be widely shared by different individuals for different reasons and they may fluctuate markedly within a single individual over time. By contrast the structure of cognition is regarded as more enduring, organism-specific, and invariant over situations; hence structural variables provide a better description of the person as conceived in most psychological theories".

Scott also points out that the structure/content distinction is relative, and depends on the level of analysis of the research question. Thus although a cognitive structure is an association of lower order elements, it is also an element within a higher order structure. Ethnocentrism is a structure vis-a-vis particular beliefs about race and racial characteristics, but it is also the content within a higher order structure of authoritarianism.

Cognitive style dimensions are intended to be individual differences in the structural properties of cognition. Theoretically, at least, they are intended to generalise across different content areas, although in practice, their range of application is recognised as being limited. In the writer's view, this discrepancy between the aspiration level and the achievement level of cognitive style dimensions is of central theoretical importance, and reflects on the validity of the structure/content distinction. The distinction itself permeates much of psychological thinking and is a basic and usually unstated assumption in psychological theories and models. The validity of the distinction, and its origins in research methodology are examined in Chapter 4.

2. Cognitive Systems can be distinguished in terms of the degree of differentiation. The term "differentiation" was first used by French (1947, 1948) in a discussion of sentiments. It refers to the distinctiveness of an element, or the amount of separation between one element and another. Krech and Crutchfield (1948) discussed the "precision" of an attitude in terms of its clarity and differentiation, and the concept is now solidly embedded in cognitive style theory. Whilst "differentiation" is generally agreed to refer to the extent to which an individual's cognitive structuring can distinguish between elements in the "objective" environment,

there are different operational definitions. Bieri et al (1966) point out that their measure of cognitive complexity is a measure of differentiation, but of a particular sort of differentiation:

> "....we are concerned with the differentiation of dimensions of judgement, rather than with categories, concepts or regions".

Bieri's measure of complexity involves a matching procedure which scores the extent to which the dimensions of judgement in Kelley's (1955) RCRT are used differentially. Other measure of complexity include the strength of the first factor in a non-parametric factor analysis of Kelley's RCRT (Campbell, 1960), an index of multidimensional information yield developed from information theory (Ulehla, 1961) and another index of information yield based on an object sorting task (Scott, 1962, 1965). Furthermore, these normative indices of differentiation should be distinguished from the internally-referenced measures used by the phenomenological approaches (Zajonc, 1960).
Research by Wyer (1964) makes the situation more confusing. He applied two measures of differentiation to the same data, in the domain of person perception. They were Scott's (1962) index of information yield and a matching procedure similar to Bieri's (1955). He found that there was little correlation between them, and that the two measures appeared to be tapping different things.
In the absence of systematic research on these different measures of cognitive differentiation, the only conclusion which is fairly certain is that the concept of differentiation is insufficiently differentiated. The corollary is that any

research in this area gives conclusions which are specific to the particular measures used.

The term "articulation", in the cognitive style literature, is often used synonymously with differentiation (see Scott, 1963).

3. Cognitive Systems can be distinguished in terms of the extent of integration.

In cognitive theorising, the term integration is used in much the same way as it is in general language, and like much of general language, it remains vague. Everybody knows what is meant by integration, but it is extraordinarily difficult to define it in precise terms. Perhaps the most commonly accepted definition refers it to the connectedness between parts of a structure, or the degree to which a person can move from one part of the structure to another, as appropriate. Zajonc (1960) uses a phenomenological method to measure the dependency of one concept on another, and defines a concept of 'cognitive unity' as being equal to the sum of the dependencies in relation to the maximum number of dependencies possible for a given number of cognitive elements. To the extent that cognitive elements are considered to be connected by dependencies, this is as precise a definition of integration as will be found. Harvey, Hunt and Schroder (1961) present the idea of 'Compartmentalisation vs. inter-relatedness' of structures, a dimension expressing the degree of 'essentialism' of a concept: this dimension is obviously similar to the concept of integration. The degree of integration would also be reflected in the 'centrality' or 'peripherality' of a concept in a structure, and might be manifested ideationally by ego-involvement in beliefs.

What is meant by integration must depend to a large extent on what is meant by relatedness. Zajonc (1960) means by relatedness, dependency. Other possibilities are similarities, cognitive consistency, psycho-logic, and indeed any criterion which might be specified in a rating scale. It is to be hoped that different rating criteria reflect purely operational differences, and will be reducible to more basic forms of integration. Until these can be found, the concept of integration is inevitably vague and unsatisfactory. There is not yet any good objective measure of integration. Zajonc's (1960) procedure, using phenomenological judgements, is independent of the experimenter's judgement, but makes heavy demands on the ability of the subject to comprehend what he is supposed to do. In the cognitive style dimensions below, integration is usually estimated by trained judges on the basis of projective material: although the reported inter-judge rating reliabilities are quite high, such experiments demand particular procedures to avoid any possibility of experimenter effect, and the writer has not been able to find evidence that such procedures have been used in many cases.

However, Wyer (1964) has produced some encouraging findings: he used two measures of integration in addition to measures of differentiation, and a number of other variables. One measure involved a count of the number of sets of attributes which could be used to infer a concept, and then summing across the domain. His other measure was based on subject judgements of dependency. He found that the two measures were significantly and highly correlated (0.66, n = 40). He also found that they were positively correlated with the frequency with which a change of judgement on one dimension led to a change of judgement in the second, a confirmation of the construct validity of the measures.

2:2 Cognitive Complexity

Bieri's dimension of cognitive complexity, and Harvey et al's concreteness/abstractness use different empirical measures, and embody slight theoretical differences. It is not possible to consider them separately, however, as the predictions that can be deduced from them are almost identical, and there is insufficient empirical evidence to define them independently.

2:2:1 Cognitive Complexity-Abstractness/concreteness

The core theory of this approach was elaborated by Harvey, Hunt and Schroder (1961). They began with an analysis of the structural variables of cognition, from which they deduced that the growth of cognition followed certain stages. Depending on the end point of the growth process, a person was said to have one of the four basic types of cognitive system, Systems 1, 2, 3, and 4, which varied continuously from concreteness to abstractness. System 4 is the most abstract. The four different systems were not intended to be a typology in the strict sense of the word, but rather an operational categorising of a continuous dimension. Although differentiation is acknowledged as being a determinant of the concreteness/abstractness continuum, Harvey et al's main emphasis is on the degree of integration.

System 1 is the most concrete cognitive system, and its structures are characterised by compartmentalisation, and by a rigid hierarchical integration of parts. Each element is linked to other elements in a restricted way, with few alternative paths. This type of system is depicted in figure 2:1. Developmentally it is produced by child-rearing practices which closely approximate the conditioning model of S-R psychology, with the generation of concepts by trial and error, the internalisation of values without insight, and the ritualistic adherence to rules without understanding. It is characterised by few degrees of freedom and has a low potential for generating conflict or ambiguity or for resolving ambiguity by means other than exclusion. Dimensions tend to be dichotomous rather than finely graded, leading to black/white classification of the world. One can infer general behavioural characteristics of repression or denial of conflict-inducing stimuli, anchoring of behaviour in external conditions, and a greater generalisation of rules over a certain range (and greater change when that range is exceeded). Harvey (1967) suggests that more specific symptoms of this system are "high superstition, high religiosity, high absolutism and closedness of beliefs, high evaluativeness, high positive ties with and dependence on representatives of institutional authority......., high identification with social roles and status positions, high conventionality and high ethnocentrism".

System 2 is slightly more abstract than System 1, and the integration index is only moderately low. There are alternative paths between items (see figure 2:2) allowing given stimuli to be classified in different ways according to which rules are applied, but there is no higher order apparatus for organising this variation in interpretations. Developmentally, the system is a result of blind conditioning applied inconsistently; the

Aston University

Illustration removed for copyright restrictions

Diagram 2:1. System 1 - low integration index

Aston University

Illustration removed for copyright restrictions

Diagram 2;2. System 2 - moderately low integration index child suffers from a surplus of unpredictable variety which leaves him with a disorganised diversity of functioning. This treatment provokes a greater abstractness because of rejection of authority and forced independence, an ability to appreciate ambiguity and divergent viewpoints but no capacity to organise these and choose between them. This system is a psychological vacuum, propelled along by kicking against the pricks, rather than by purposeful locomotion. The behavioural characteristics are a movement away from absolutism instability, and non-commitment, and there is a certain rigidity in that when a choice is made, it is held to stubbornly in the face of environmental pressures.

System 3 has a moderately high integration index (see figure 2:3) allowing a combination of schemata and a comparison between choices. Where the moderately low structure is characterised by the emergence of rules for combining alternatives, the moderately high structure is characterised by the emergence of rules for identifying more complex relationships than alteration, and for organising schemata in relatively independent ways. However, these rules themselves are still delineated and relatively rigid. It follows that the system is less deterministic; even when the individual has made a choice, he is still open to pressures to change that choice. Abstraction has become a formal rule of the system. The system is thought to be the result of over-indulgence by one or more of the parents, encouraging the child to manipulate the parents rather than to explore the world. System 3 individuals, more than the representatives of any other system, are concerned with establishing friendships, intra-group concensus and

Diagram 2:3   System 3 - moderately high integration index

Diagram 2:4   System 4 - high integration index

Diagrams 2:1 - 2:4 are taken from Schroder, Driver, and Streufert (1967).

dependence relations in order to avert feelings of helplessness and social isolation.

System 4 is the most abstract of all the systems (see fig. 2:4). At this level comparison rules can be further integrated and varied, allowing for the generation of more complex structures, and highly abstract inter-relationships. The implications are (1) an increased degree of diversity, (2) greater discrimination between stimuli within a dimension, (3) an increased potential for generating patterns of interactions. This system is said to result from child-rearing techniques as they are practised in the All-American home.

This dimension bears strong similarities to authoritarianism and dogmatism. Both of these are found to correlate with concreteness/abstractness, and the discrimination of systems is improved if both measures are used together rather than independently (White, Alter, and Rardin , 1965: White and Harvey, 1965).

Concreteness/abstractness is inferrable from behaviour symptoms as well as cognitive performance, but there are no reliable objective tests. The tests most usually used in the research are semi-projective measures, where the subject's output is rated by the examiners as belonging to one of the four systems. The "This I believe Test" (TIB) is aimed at eliciting material about the "individual's more central concepts" (Harvey, 1967), by requiring the subject to complete a number of sentences which begin "This I believe about.......". In fact concepts are provided to fill the blanks, rather than obtaining the individual's central concepts: typically they include friendship, the American way of life, guilt, marriage, myself, religion, sin, majority opinion, people, compromise, the future, and the past. This material is scored in terms of their positive and negative orientation toward the referents, their absolutism, evaluativeness, multiplicity of alternatives, triteness, and normativeness. Also used are two other semi-projective measures, the Sentence Completion Test, and the Paragraph Completion Test. Harvey (1966) reports that the "This I believe" test is fairly reliable, and that the interjudge rating reliabilities for trained judges was at the level of 0.9 and above on 12 different samples. Also Vannoy (1965) reports that all of the cognitive complexity tests for which reliabilities have been published are in the 70's or above. It is regrettable that the writer can only report reliabilities second-hand, but much of the relevant information appears to circulate the American universities in unpublished form. Many of the studies are reported only in summary, with experimental details remaining unpublished. Alternatively, the information is recorded in restricted access documents. Thus the only reference to the Princeton Objective Test of Conceptual Systems, an objective test of concreteness/abstraction based on completions to the TIB, is "Schroder, H.M.(1964): An Objective test of conceptual systems. Social Psychology Laboratory Library, Princeton University, Princeton, New Jersey". Inter-library loan requests and personal enquiries to the authors produced only a further reference to Schroder, Driver and Streufert (1967), in which there is a brief note to the effect that objective tests are inferior to the projective tests normally used (pp.203-204). This reporting practice undoubtedly reduces what is often tedious detail, but is a great hindrance to the development of reasoned criticism.

2:2:2 Cognitive Complexity

This approach was developed originally by Bieri (1955), and its empirical correlates have been demonstrated almost exclusively in the field of interpersonal perception. It is grounded in Werner's (1957) developmental psychology, which described a growing cognitive system involving increasing differentiation and articulation of elements, together with an increasing inter-dependence of these elements. The approach follows closely Kelley's (1955) and Lewin's (1951) conceptualising of cognition. Bieri suggests that the over-all cognitive complexity was two distinct sub-components, differentiation and integration. These are sometimes estimated independently, but more often combined into the single measure. The two sub-components are assumed to correlate positively. The test most often used is a modification of Kelley's (1955) RCRT in which the subject had to rate ten individuals according to 10 provided constructs on a 6-point Likert scale. Evidence from Tripodi and Bieri (1963) and Kieferle and Sechrest (1961) suggest that provided constructs are functionally equivalent to elicited constructs. The ratings of different individuals are compared using a matching procedure, (described in Bieri et al. 1966 and in Chapter 7 below), which assesses the extent to which dimensions have been used differentially. The most simple subject would apply all dimensions identically to all individuals and would obtain the maximum score of 450. A score as low as 100 would indicate a relatively complex individual.

Estimated separately, differentiation is obtained from the number of dimensions used to judge a cognitive domain: the greater the number, the greater the differentiation. Integration is obtained from multi-dimensional scaling techniques. These measures are often used inter-changeably with each other, and with Harvey's measures, an unjustifiable practice in the light of evidence from Vannoy (1965) which suggests there is little correlation between them.

2:2:3 Empirical Evidence

There has accumulated a vast amount of research on cognitive complexity. As there is at least one major volume devoted to the subject, it is obviously impossible to perform the same function here. The evidence is reviewed comprehensively in Bieri et al (1966), Bonarius (1965), Crockett (1965), and Wiggins (1965). The present review attempts to cover the major theoretical points, and subsidiary findings which have a bearing on the relevance of cognitive complexity to education. A major question is the degree of generality of the dimension. Is a person who is cognitively complex where interpreting the weather is concerned, also complex in the domain of sport? A priori, it seems unlikely. Unless a dedicated nativist, one would assume that an individual would be more complex in those areas in which he was experienced. Both approaches to cognitive complexity claim only limited generality within particular domains, but even the truth of that claim is doubtful. Also, field dependence, as will be seen, has demonstrated considerable generality and is to some extent related to cognitive complexity. The question is therefore not an empty one.

The empirical evidence is scarce and equivocal. Bieri and Blacker (1956) measured cognitive complexity in responses to the Rorschach, and to the modified RCRT. They found a moderate positive correlation between the two, a result which was not replicated by Sechrest and Jackson (1961). Ulehla (1961) measured the cognitive complexity of nurses, in dealing with schizophrenic symptoms, and in defining the nurses' role, and found a near zero correlation. Scott (unpublished, but mentioned in Scott, 1963) found no correlation between estimates of cognitive complexity from concepts relevant to nations and concepts relevant to people. However, Allard and Carlson (1963) used three measures of complexity, all on the Mod RCRT format. The measures varied in the objects of the rating procedure, which were either (i) friends, (ii) famous people, or (iii) geometrical designs. They obtained inter-correlations of between 0.57 and 0.67.

Vannoy (1965) made an extensive factor-analytic study of 20 tests which measured cognitive complexity or related theoretical dimensions. He hypothesised that a unitary cognitive complexity variable should result in a first factor which extracted a large part of the variance. Using a varimax factor analysis based on data from 113 male subjects, 8 factors were extracted, the largest accounting for only 24.3 per cent of the variance. Furthermore all of the inter-correlations were at very low levels, as can be seen from the sample of the inter-correlation matrix reproduced in table 2:1. Of particular interest for present purposes, is the correlation of 0.05 between Bieri's Mod RCRT, and

Harvey's Sentence Completion Test. Vannoy concluded that there is little evidence for a dimension of cognitive complexity, even within the domain of interpersonal perception, although he noted that different results may be obtained with a more heterogeneous sample. Vannoy's conclusions echoed the view of Gardner and Schoen (1962), who suggested that cognitive complexity is "grossly overgeneral".

	1	2	3	4	5	6
1						
2	-.29					
3	-.20	.18				
4	.11	-.06	-.10			
5	.05	.04	.01	-.10		
6	.06	-.01	-.04	-.18	.15	

Table 2:1 A sample of the inter-correlation matrix reported by Vannoy (1965)

1. Mod RCRT - Bieri (1955)
2. Intolerance of trait inconsistency - Steiner (1954)
3. Authoritarianism - ten statements from the original F scale (Adorno et al. 1950), together with ten other negatively-valenced statements.
4. Category Width - Pettigrew (1958)
5. Sentence completion test - Schroder, Driver, and Streufert (1967)
6. No. of groups. Modified Scott test - Scott (1962)

The generality of cognitive complexity is also tested negatively in several experiments designed to investigate a 'frequency of interaction' hypothesis, i.e. that people who have a greater experience of social interactions are likely to be more complex with regard to interpersonal perception. Mayo (unpublished, but reported in Crockett, 1965) measured cognitive complexity in two groups of college students, and found that fraternity members were more complex than non-fraternity members. It was reasoned that those students who were selected for the fraternities were likely to be more oriented to social interactions. Bieri and Messerly (1957) found a positive correlation between cognitive complexity in person perception, and extroversion, and Supnick (1964), in a factorial experiment, found that subjects used more constructs in judging peers and same-sex colleagues than in judging older and opposite-sex colleagues. There was also a significant interaction effect between the two dimensions, which also supported a frequency of interaction hypothesis. These results, while not conclusive, suggest that the generality of cognitive complexity is likely to be highly restricted. The difficulty is that as one cannot know how near or how far apart are two domains or sub-domains, one cannot measure the degree of generality. These results also cast doubt on the assumption that the tests are measuring "structural variables".

One inference from cognitive complexity is that persons high on the dimension should use a higher number of categories in judging people and things. They should also be able to draw a greater number of inferences from a given amount of information. This hypothesis has been confirmed by Campbell (1960), who showed that less complex judges were more likely to classify people into a good/bad dischotomy, than were more complex judges. They were also more likely to perceive relations amongst their colleagues as balanced i.e. mutual likes and dislikes. The hypothesis was further supported in experiments by Carr (1965), and Tripodi and Bieri (1966). Carr found that abstract subjects as measured by the Sentence Completion Test, made finer interpersonal discriminations on Kelley's RCRT. Tripodi and Bieri (1966) used a task similar to the Thematic Apperception Test (TAT), in which subjects were required to tell stories about imaginary persons in three different social situations. The more complex subjects were found to perceive a significantly greater amount of inter-personal conflict in the situations than did the less complex subjects. These results could be interpreted as showing that complex subjects are simply more pathological, although if one considers the experiment in conjunction with the other reported research, a conclusion involving the usage of a greater number of stimulus dimensions is perhaps more plausible.

Finally, Supnick (1964), in the experiment reported in the above script, found that for both males and females the more complex subjects provided significantly less univalent descriptions than did the less complex.

Unfortunately, length of description was not controlled or taken into account, and it may be that more complex subjects are simply more verbose: this could account for the experimental results without recourse to any postulates about the dimensionality of the cognitive system.

In general, the evidence appears to support the hypothesis that persons high in cognitive complexity use a higher number of categories, and finer dimensions of judgement. One can also deduce from cognitive complexity theory that persons will be more accurate in their perceptions of other people.

This hypothesis was tested by Bieri (1955), who had his subjects fill in the modified RCRT as they thought their classmates would. He found a low but significant positive correlation between cognitive complexity and accuracy of prediction to their classmates actual responses. Further analysis of results showed that this greater predictive accuracy was attributable to the perception of points of difference rather than points of similarity, the respective correlates with differences and similarities being 0.35 and 0.02. Furthermore, the more complex subjects recognised where their classmates would differ from them, whereas the less complex subjects were more likely to assimilate, and believe friends to be similar to them.

This result was only partially confirmed by Leventhal (1957). Leventhal varied the cognitive complexity of both the judges and the judged, and the amount of stimulus information available. While there was no significant correlation between cognitive complexity and accuracy, the data demonstrated that the less complex subject tended to see the other as similar to himself to a greater extent than the more complex subject. He also demonstrated that more complex judges differentiated more between the judged. However, a study by Sechrest and Jackson (1961), using a large number of measures of both accuracy of prediction and of cognitive complexity, failed to obtain significant differences, Finally, a study by Plotnick (unpublished but reported in Bieri et al. 1966) supported the hypothesis. He asked social work students to judge the attitude of three patients towards authority, and found that more complex judges predicted the mean attitude toward authority scores of the three patients in the correct rank order, while less complex judges could not discriminate between two of the three patients.

The evidence on the accuracy of interpersonal perception is equivocal. Some encouraging results have been found but final judgement must await research into the defining conditions of the non-supporting experiments. It is likely that the hypothesis will need to be further differentiated.

A further hypothesis which has been tested is that persons high in cognitive complexity will be better able to deal with cognitive dissonance. This statement is vague and necessarily so: methodological difficulties in this area of study mean that it is not exactly known what is being tested. For instance, in an experiment where dissonant information is presented, it might be found that more complex subjects take the original information and the dissonant information into account while the less complex subjects only take into account the latter. There would be several possible explanations for this: firstly, more complex subjects might be more able to use abstract structures in order to resolve any apparent inconsistency and combine both types of information into a single interpretation. Secondly, it might be that more complex subjects do not perceive any dissonance as they already possess this higher order structure. Thirdly, more complex individuals might be subject to the same amount of dissonance as less complex individuals, but be constitutionally more able to tolerate it. Dissonance is a hypothetical construct, which cannot be directly measured, but which must be inferred from the behaviour which can be construed as reducing it. There have been attempts to find physiological correlates of dissonance, such as GSR's (Burdick and Burnes, 1958; Brehm Back and

Bogdonoff, 1964; Steiner, 1964; Zimbardo and Dworkin, 1964, Glass, 1968), but none of these are yet reliable. Furthermore, the basis on which cognitions can be congruent or dissonant is not at all well defined. In Aronson's (1968) words, the situation has become "if you want to be sure, ask Leon". Dissonance in the experimental situation is specified by the experimenter, and the subjects may or may not experience it. If they experience it, they may choose to resolve it in the way provided for in the experiment, or they may use another way. As dissonance can only be inferred from attempts to reduce it, the measurement of dissonance is a chancy business, and the testing of the above hypothesis difficult.

Bearing these methodological problems in mind, there is some empirical evidence on the subject. Nidorf (1961) showed that more complex subjects are significantly more likely to integrate positive and negative valenced attributes of an unknown person in a single description. The less complex subjects either fail to integrate the attributes or ignore the attributes of one valence. Crano (1967) found that more abstract subjects were more diverse and less internally consistent in reducing conflict created by counternorm messages. These results are supported in an experiment by Mayo and Crockett (1964) which investigated primacy/recency effects in attitude change. The subjects were presented with a list of behaviours, manifested by an unknown other and were asked to write an impression of that other. This was followed by a second list and a second written impression. Both lists were univalent, each containing behaviour of an opposite valence: positive and negative were counter-balanced for order effects. It was found that less complex subjects reacted with a substantial recency effect, ignoring the earlier information. More complex subjects, on the other hand, tended to use the information from both lists. It should be noted that Mayo and Crockett's index of cognitive complexity was the number of verbal constructs produced rather than the more usual Mod. RCRT.

Leventhal and Singer (1964) used the Mod. RCRT to predict the resolution of inconsistent information, and got different results. They presented information about an unknown other that was either positive, negative, or intermediate in valence, and took a number of judgements from the subjects. They then presented information of a contrary valence and repeated the judgements. In view of the large number of judgements taken, the findings are too complex to report in detail, but with respect to the main hypothesis, they were non-significant. An interesting side result was that more complex subjects appeared to search for information related to inner states such as maladjustment, and less complex subjects responded more to the surface qualities of behaviour. If persons of differing cognitive complexity do use different information as the basis for their judgements, then this factor must obviously be taken into account in dissonance resolution experiments. The apparent contradiction between this experiment and the experiment of Mayo and Crockett might be the result of differences in stimulus structure. Neither of these experiments placed a tight control on the type of information contained in the vignettes about the other persons. Tripodi and Bieri (1964) have performed an experiment which shows the importance of this variable. They specified three properties of stimuli that might relate to the cognitive complexity of the judges. These were dimensionality of the stimulus, univalent or ambivalent combinations of stimulus dimensions, and quantity of information. These three properties were varied in a task where subjects had to make judgements of pathology based on combinations of symptoms. The results showed that as the stimulus dimensionality increased, there was an increase in information transmission, and this increment was greatest for judges who were low in cognitive complexity. This suggests that less complex judges gain more discriminability with an increase in dimensionality, whereas the more complex judges can rely on internal structures rather than stimulus dimensions. This conclusion is congruent with the results of Leventhal and Singer (1964), who found that less complex judges, as compared with more complex judges, increased the accuracy of their predictions with more information. For most stimulus conditions, Tripodi and Bieri (1964) found that more complex judges showed more accurate discrimination than less complex judges. These results would be congruent with those of Mayo and Crockett.

Tripodi and Bieri's results seem to provide a rapprochement between Leventhal and Singer's findings and the findings of Mayo and Crockett. If this is so, one can say that the general results appear to support the hypothesis that more complex judges are better at resolving cognitive dissonance. However, the situation appears to be considerably more complex than was originally thought, possibly requiring parameters of the content of the information to be taken into account.

Some interesting correlates of cognitive complexity are revealed in a factor analysis of the Princeton Objective Test of Conceptual Systems (POT), as used by Harvey (see Harvey, 1967). The statements in the POT were derived from completions to the TIB, and ratings of these statements were subjected to factor analysis which revealed seven independent factors. These factors are as follows:

1. Divine Fate Control, manifested in such items as "there are some things which God will never permit man to know".
2. Need for consistency.
3. Need for structure order. This is reflected in such items as "I do not like to work on a problem unless there is a possibility of coming out with a clear-cut, definite answer".
4. Distrust of social authority.
5. Friendship absolutism.
6. Moral absolutism.
7. General pessimism.

Harvey (1966) reports a number of studies which demonstrate the construct validity of the four conceptual systems. The method of testing hypotheses is slightly different from many of the studies reported above. Correlations cannot be used as the TIB results only in a categorisation into one of the four systems. The customary practice is therefore to test differences in performance between independent groups. Furthermore, subjects are only included in a study when judges are unanimous in labelling of the subjects' functioning. This results in the elimination of approximately 30 per cent of the subjects. Harvey notes that the remaining 1400 subjects showed the following distribution of scores:

System 1	Most concrete	30 per cent
System 2		15 per cent
System 3		20 per cent
System 4	Most abstract	7 per cent
Mixtures of two or more systems		28 per cent

As noted above, Harvey's (1966) report of findings is in summary form only, and the writer has been unable to trace original research reports. This places undue strain on the syntax of the summary report: in particular, Harvey sometimes mentions that observed differences are significant, sometimes mentions that observed differences in a way which implies that they are significant, and sometimes mentions observed differences without reference to statistical significance. In view of the second category, it is difficult to know whether Harvey is only reporting significant findings but omitting tedious repetition, or whether he is reporting all findings. The writer can therefore only repeat Harvey's report, complete with ambiguities.

Harvey notes that differences were found between conceptual systems on the following variables:

1. Intelligence, as measured by the WAIS. The only differences were on the verbal intelligence and vocabulary subjects, where systems 2 and 4 scored higher. Implication that differences are significant.

2. Cognitive complexity, measured by a modification of Kelley's RCRT devised by Campbell (1960). System 4 was the most complex, followed by systems 3, 2, 1, in descending order. Significance unmentioned.

3. Religion. Response to items on religion suggested that system 1 was most religious, followed by systems 3, 4, 2, in that order. System 2 individuals, much more than system 4, were likely to be actually against religion. Significance unmentioned.

4. Authoritarianism. F.Scale. System 1 most authoritarian, followed by systems, 3, 2, 4, respectively. Significance unmentioned.

5. Dogmatism. D-Scale (Rokeach, 1960). System 1 scored highest, followed by 2, 3 and 4, in order. Significance unmentioned.

6. Rigidity. Scale on Rigidity (Gough and Sanford, 1952). Scores were as follows System 1 *System 2 System 3 *System 4. The * marks significant differences.

Harvey also reports (1966) the following experimental findings:

1. System 4 performed significantly better in a concept formation experiment.

2. In a concept formation experiment by Felkner and Harvey (1964), subjects were rated on the redundancy of information requested and on the basis for guessing the solution. They found that System 1 individuals requested significantly more redundant information than System 2 and System 4 individuals, and that System 4 were significantly better and System 1 significantly worse, at using attributes in devising hypotheses.

3. On the EFT, System 4 individuals scored highest, followed in order, by Systems 3, 2, 1. "The attainment of the first three systems was close together and widely separated from System 4" (Harvey, 1966, p.54). Significance was not mentioned.

4. An impression formation experiment, in which subjects wrote an impression of: first, an individual described by three positive-valenced adjectives; second, an individual described by three negative-valenced adjectives; and third an individual described by all six adjectives. Scored in terms of ability to present an integrated impression, System 4 individuals scored better significantly. The other Systems scored in the hypothesised order, but the differences were not significant.

5. The same subjects were asked to list three wishes and how they might attain them. They were then asked to combine these into a super-ordinate wish and say how they would attain it. Again, scored in terms of degree of integration, the System's scores were in the predicted order, but the only significant difference was between System 4 and other Systems.

6. Using a concept attainment task, based on Bruner, Goodnow, and Austin's (1956) materials, the presentation method was initially ordered, but was then changed to random, thus requiring a change of set on the part of the subjects. On the first trial under the random presentation condition, "System 1 individuals took longer, tested more cards, and had a significantly higher total number of steps in attaining the concept". (Harvey, 1966: p.55).

Harvey (1967) also reported the following correlates of concreteness:

1. A poorer capacity to role-play (Harvey, 1963)
2. The holding of opinions with greater strength and certainty that the opinions will not change over time (Hoffmeister, 1965)
3. A lower score on the factor of task orientation

(Harvey, Reich, and Wyer, 1966)

4. A greater likelihood of accepting false information if it is attributed to a high status rather than to a low status communicator (Harvey, 1964)

Taken together, this list of research findings from Harvey and his students presents impressive evidence for the generality of cognitive complexity, with predictions being verified in the more centrally cognitive processes of concept attainment and problem solving in addition to person perception. The findings are doubly impressive when compared with research on cognitive complexity, as measured by the Mod. RCRT, where, as detailed above, there is doubtful generality even within the domain of person perception. With this comparison in mind, and with the increased possibility of experimenter effects which occurs if experimenters score both the predictor and predicted variables subjectively, it becomes more than usually desirable to trace the findings back to the original experimental reports. Unfortunately, with the exception of Harvey (1964), all references reported here from Harvey's (1966) and (1967) articles, are unpublished Technical Reports, from the University of Colorado. In the absence of the original experimental reports, it is one's duty to remain sceptical.

There has been some recent research using concreteness-abstractness as a predictor of information search behaviour. The situation typically used is the Tactical War Game, in which the subject has to make economic, political and military decisions about the island of Shambe, which is an undeveloped island involved in civil strife. The game is computer controlled and the subject is fed with a certain amount of information on which to base his decisions. He can also request further information. The computer also feeds back the results of his decisions. The game is extremely complex, and is described in detail in Streufert, Kliger, Castore, and Driver, (1967). A number of experiments (Sieber and Lanzetta, 1964; Suedfeld and Streufert, 1966; Streufert, Suedfeld and Driver, 1965) have varied the information load and have measured both the active, self-initiated information search and delegated information search in group situations. The type of question was also classified as to whether it was a monitoring question, which dealt with past and ongoing aspects of the game, or an integrative information question, which dealt with situations not asked about previously, or future decisions growing out of previous information. The results of these experiments showed that there was no difference between subjects of differing cognitive complexity in the total information search. However, as the information load was increased, the information search was decreased, and complex subjects were less affected than simple subjects. They also found that more complex subjects showed a greater amount of self-initiated search and a greater number of integrative information questions than less complex subjects. Karlins and Lamm (1967) used a variation of this task called the Community Development Task. The game is similar except for the fact that all questions involve integrative information. They therefore measured the total amount of information search and found that more complex subjects produced more activity than less complex subjects.

Streufert and Castore (1971) differentiated the information presented according to three variables:

1. The information load, as used in previous experiments
2. Noxity: this was defined as the proportion of information which reported failure.
3. Eucity: defined as the proportion of information indicating success

All three variables were manipulated in their experiment. Dyads were formed that were homogeneous in terms of cognitive complexity and were requested to play the Tactical War Game. Seven messages were delivered in a thirty-minute period of playing and the game consisted of six periods. In the first period, one message in seven indicated failure, and the rest indicated success. The proportion of failure messages rose monotonically to 6/7 in the sixth period. Streufert and Castore hypothesised that noxity and information load were both stresses, and as such should have equal effects on information search. The results were confusing; they found that cognitive complexity did discriminate successfully between the performance of high complex and low complex subjects, but while more complex subjects used less delegated search throughout the range of noxity as predicted, there were not significant differences in self-initiated search. These results are different from Streufert et al's (1965) findings, suggesting that information load and noxity cannot be considered equivalent stresses in determining information search. Streufert and Castore also measured information itilisation frequency and information utilisation efficiency, and predicted that all subjects should increase and then decrease on these measures as noxity increased, but that the curve for more complex subjects should be higher than the curve for less complex subjects. These results were confirmed as can be seen in figures 2:5 and 2:6.

The results from these information search experiments might prove interesting to educational researchers. The problem on which the subjects are tested is exceedingly complex, and may have a high external validity. In particular, it appears to be similar to the type of behaviour which students manifest when they perform independent study. McLeish (1968) has suggested that some experiments investigating individual differences in reaction to teaching methods have obtained zero results because any differences in performance produced by the experimental variable are entirely overlaid by independent study, library work, and homework, all of them outside the experimental situation. If cognitive complexity can predict information search, and if information search is analagous to independent study as involved in tertiary education, then it is possible that cognitive complexity will prove an effective screening device to be used as an extra control variable in such educational experiments.

The results also have general theoretical importance: much of university education is self-initiated rather than directive, and if cognitive complexity is predicting the self-initiating aspect rather than the information search aspect, as indeed is suggested by the results seen above, then cognitive complexity would seem to be a potentially important variable in our understanding of the educational situation.

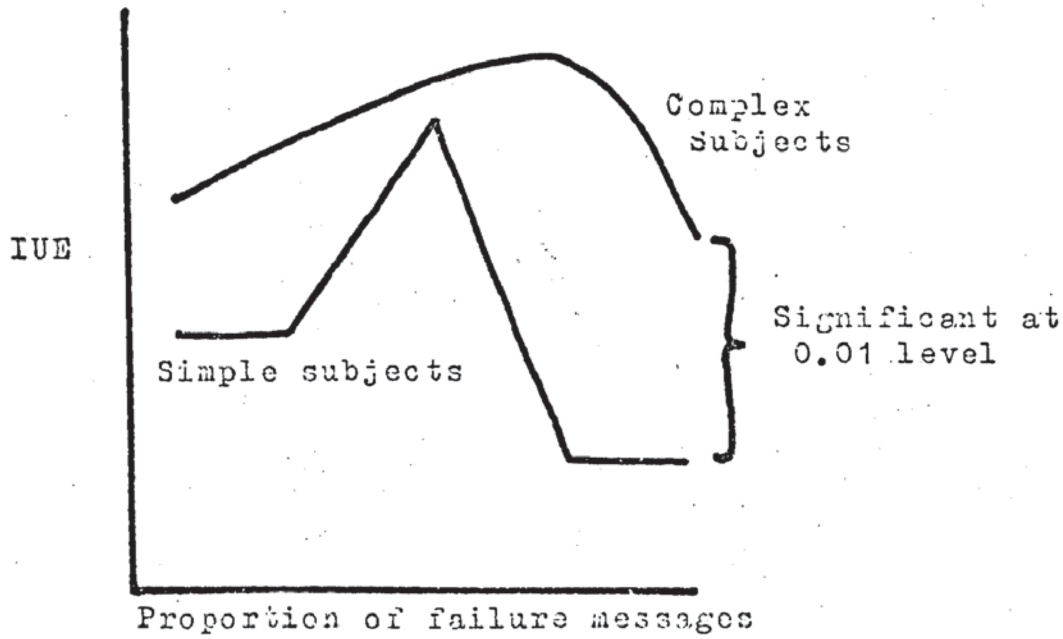


Figure 2:5. The function resulting when information utilisation efficiency (IUE) is plotted against proportion of failure messages.

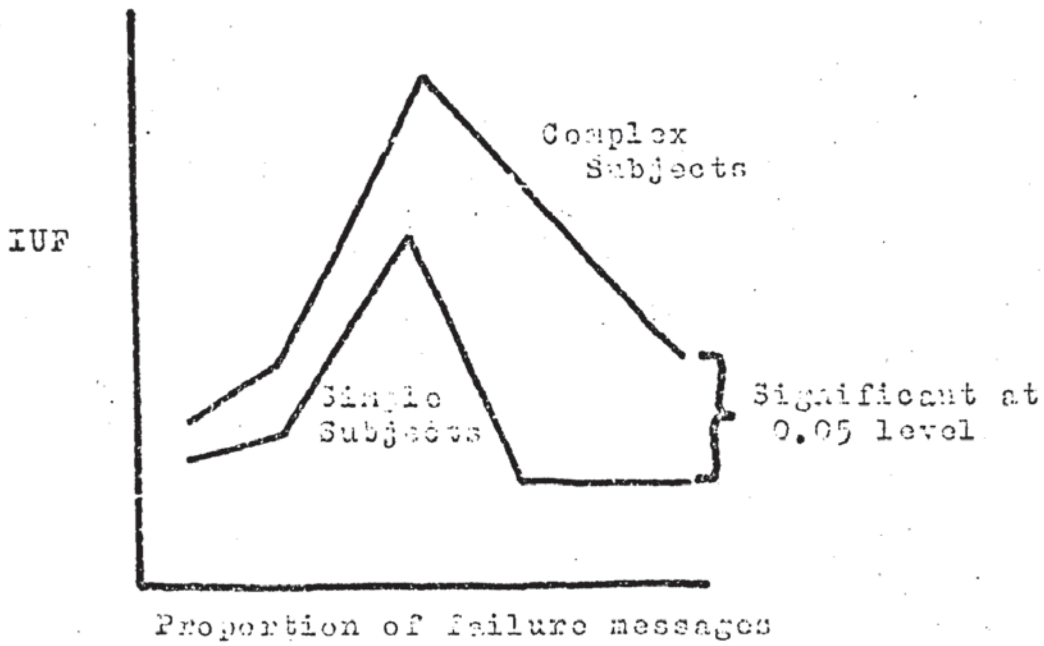


Figure 2:6. The function resulting when information utilisation frequency (IUF) is plotted against proportion of failure messages.

2: 3: 1 Cognitive Differentiation vs. Global Functioning

A dimension which bears a remarkable theoretical similarity to cognitive complexity, and which is possibly centred on the same underlying variable, is Witkin's differentiated vs. global functioning. The dimension was originally called field dependence vs. field independence (Witkin, Lewis, Hertzman, Machover, Meissner, and Wapner, 1954), and developed from the interest in individual differences in perception, which had been stimulated by the New Look theories of the late '40s. Originally conceived of as the extent of dependence on the perceptual frame of reference, field dependence was quickly found to have other personality and intellectual correlates. Elliot (1961) suggested that the focus of research should be shifted from personality to intellectual functioning, and Witkin, Dyk, Faterson, Goodenough, and Karp (1962) renamed the dimension as psychological differentiation, and broadened the underlying theory. There has probably been more research concerned with field dependence than with any two of the other styles discussed in this thesis. Consequently, what follows can only be a summary of the more important findings.

The field-independent individual, or the individual with differentiated functioning, is said to have an "active-coping" approach to the environment. He attempts to master and re-organise those aspects of the environment which do not fit in with his needs, he has a striving for independence, competence and leadership, is concerned with inner-life and internal frames of reference and shows control over undesirable impulses such as aggression. The field dependent person has a passive-submissive approach to the environment, and is described as having little initiative in challenging the structure of the status quo, being conventional and submissive to authority, being blandly unaware of inner feelings, being dependent on external and social frames of reference, and fearing and denying, as well as having poor control over impulses, such as sex and aggression. As with other personality dimensions, these two descriptions represent the two extremes of the normal distribution.

2: 3: 2 Measures

The original test used was the Tilting-Room-Tilting-Chair Test (TRTC). The apparatus consisted of a chair which could be tilted from side to side within a room which could also be tilted in the same axis. Subjects are asked to (a) adjust their body until it is vertical, and (b) adjust the room until it is vertical. It was found that some subjects can perform these tasks irrespective of initial degrees of tilt: these subjects are independent of the perceptual field. Other subjects, when asked to adjust themselves to the vertical, adjusted themselves with the room, even if the room was tilted by as much as 35°. These subjects are dependent on the perceptual field. In a second test, the Rod and Frame Test (RFT), the subject is asked to adjust a rod until it is vertical. The rod, and the square frame behind it, are luminous, and are used in a darkroom. Both can be tilted. Field independent subjects can adjust the rod until it is vertical, depending only on internal, bodily, information, whereas the field dependent subjects rely heavily on the perceptual field, in this case the frame.

The Embedded Figures Test (EFT) is based on the Gottschaldt figures. The subject is presented with a simple figure for a short time, and is then given a complex figure in which the simple figure is embedded. He is required to trace out the simple figure as soon as he sees it, the score being the time taken. With field independent subjects, the figure simply "pops out" of the background, whereas field dependent subjects may not find it in the five minute search time allowed.

Witkin et al. (1962) report very high intercorrelations between these three tests, a finding which did not replicate Gruen's (1955) study: out of fourteen intercorrelations, only four were significant. Elliott (1961) and Young (1959), also found significant intercorrelations, but they were lower than Witkin's.

A few studies have used Thurstone's Closure Flexibility (Concealed figures) Test (CFT), which is similar to the EFT, except that the figures used are simpler, monochromatic, and the test is suitable for group administration. The subject sees a simple figure which may or may not be embedded in complex figures, and has to put a tick or zero accordingly. The test is timed, and errors are subtracted from correct responses to allow for guessing. Elliott (1961) reports that it correlates with the EFT at 0.55, and with the RFT at 0.30 (both significant at p 0.001, n = 128). He also notes that it has a similar pattern to the EFT in its intercorrelations with other measures.

There are also two other embedded figures tests, one by Jackson (1956), which correlates in the mid-90's with

Witkin's EFT, and a later group version by Jackson, Messick and Myers (1964), which is available in both chromatic and achromatic forms. This produces correlations with the EFT at between 0.56 and 0.84. Jackson et al also compared the Witkin EFT with the group versions and concluded that the latter may justifiably be substituted for the more time-consuming individually-administered test. The more recent research from the Witkin laboratory has made much use of the Figure Drawing Test (FDT). This test, usually used on children, requires the subjects to draw a picture of a person, and then to draw a person of the opposite sex. The sophistication of the drawing is rated on a 5-point scale. The rating is based, not on the usual projective qualities, but rather on the direct characteristics of the figures drawn. Scores from the FDT have been shown to be significantly related to scores on the three other main tests of cognitive differentiation. (TRTC, RFT, EFT) Witkin et al, 1962). Witkin describes the drawings of different children as follows:

"In the drawings of field dependent children, we find very little detail and unrealistic representation of proportioning and body parts. Sexual characteristics are shown minimally or not at all, to the extent that in some pairs of drawings it is difficult to tell which is male and which is female. In most cases there is no attempt at role representation. On the other hand, in the drawings of children whose performance is at the field independence extreme, we find the body drawn in realistic proportion.

Parts of the body are represented in some detail and realistically. There is clear representation of sex and sex differences. Aside from representation of the sex by body characteristics, the sex of the figure is also indicated by such externals as clothing. We also find attempts at role representation, suggesting a sense of the uses to which the body may be put". (From Witkin, 1965). The test-retest reliabilities for the RFT and the FDT, over varying periods of time are reported by Witkin (1965), and are reproduced below in table 2:2. This level of reliability was also replicated by Bauman (1951) -

Ages	RFT	FDT
10 - 14	0.75	0.84
10 - 17	0.80	0.79
10 - 24	0.66	0.64
14 - 17	0.93	0.67
14 - 24	0.86	0.67
17 - 24	0.93	0.68

Table 2:2. Test-Retest reliabilities for the Rod and Frame test and the Figure Drawing Test. (RFT and FDT).

on adults tested over a four year period. The reliabilities reported above from Witkin are stunning: the possibility that a test-retest over a fourteen year period, during adolescence - a time of life in which so many other bodily and personality factors are changing rapidly, can still allow correlations 0.8 and 0.7 (a level of reliability not achieved by some tests over weeks), is very strong evidence that cognitive differentiation is a most basic variable of cognition. If one takes Witkin's figures on the correlations between such different tests as the TRCT, EFT, and FDT, it is also/a remarkably general dimension. However, these figures have been disputed.

Gruen (1955), as mentioned above, obtained much lower correlations, particularly between the EFT and the other tests. Postman (1955) criticised Witkin's research on methodological grounds, as also did Gruen (1957), who made the following points:

1) The scoring of the tests obscures the fact that different types of behaviour can occur in the test situation. Because these other behaviours are ignored, one is led to an over-simplified theoretical explanation.

2) There is a wide variability within a series of trials on the tests. Field-dependent subjects thus sometimes show field-independent behaviour.

3) Remarking on the correlations found by Gruen (1955), Green (1957) suggested that the difference might result from different samples, and that consequently the tests might mean different things to different sectors of the population. He also suggested that, as the inter-correlations with the EFT were very low, the claim for the generality of the dimension of field independence rested solely on the correlation between the TRTC and the RFT. Accepting these criticisms, and bearing in mind the higher test intercorrelations which were found by Elliott (1961) and Young (1959), one must conclude that Witkin's claims for the dimension of field-independence or cognitive differentiation are substantially upheld. Green's criticisms are not able to cast much doubt on the very impressive empirical findings; they merely suggest caution in the theoretical interpretations.

2: 3: 3: Empirical Evidence

Sex Differences.

A consistent finding in research into field dependence is that men are considerably more field-independent than women. This was originally reported by Witkin (1950), and Witkin et al (1962) report fourteen studies which have found sex differences, beginning at the age of 8, and becoming most pronounced in adulthood. Three other studies did not find sex differences (Gruen, 1955; Bieri, 1960; Jackson, Menick and Myers, 1964).

Witkin et al (1954) suggested that the differences originated in biological roles, an interesting suggestion, which however, has little research evidence to its credit. A more plausible suggestion, (or at least, a more fashionable suggestion) was that field dependence differences are attributable to differences in social roles (Bieri, 1960). It was argued that women are encouraged to be passive and dependent, whilst men are encouraged to be assertive, and to change the environment when the environment does not fit, Inferences from this were that people who identified with their fathers should be more field-independent than people who identified with their mothers, irrespective of their own sex, and secondly that acceptance of authority should be linked with field dependence. Bieri's (1960) study provided only partial confirmation of these hypotheses; the hypothesised relationships were found, but the former relationship was more important in women, whilst the latter relationship was more important in men.

Bieri, Bradburn and Galinsky (1958) had earlier suggested that sex differences were related to spatial ability. Their study found that men performed better on the EFT and on mathematical aptitude tests, and that these two scores were correlated. They reasoned that there were two factors in sex differences in field dependence:

(1) the superior aptitude of males for mathematics

(2) males are more able to combine this aptitude effectively with a conceptual approach to social and objective stimuli, thus facilitating EFT performance.

Fiebert (1967) reviewed a number of previous studies which examined the relationship between field dependence and the dimension of masculinity-femininity. There were only a few studies, but they had benerally found a weak relationship. Certain sub-tests, those containing items related to pity and disgust, correlated with field-independence, particularly with women. He concluded that it is a particular dimension of masculinity-femininity which is related to field-independence.

In summary, the finding of sex differences in field dependence is well-attested. The reason for such differences is less clear. Suggestions of differential mathematics aptitude fits in with the idea of differential social roles, but it is difficult to see how that might relate to Fieberts' conclusion that correlations with the masculinity-femininity dimension are carried by items relating to pity and disgust. This latter finding is not a strong one, as it is possible that the correlation with field dependence is carried by a suppressor variable in the masculinity-femininity dimension or by error variance in the masculinity-femininity scale. In addition, the reliabilities of the field dependence tests which suggested the basic nature of the field-dependence, might also suggest the basic nature of sex differences, making a link with biological and/or social roles, and/or body concept differences, more likely.

Intellectual Functioning

Several studies have found that field-dependence correlates only moderately with intelligence. Thus, Jackson (1957) reported a correlation of -0.53 between EFT scores and ACE scores. Witkin et al (1962) correlated the battery of field dependence with IQ scores from the Stanford-Binet and found that r = 0.59 for boys, and r = 0.76 for girls. Cohen (1959) found a moderate degree of correlation between the FDT and the WISC, and noted that the main burden of the correlation was carried by the block design, picture completion, and object assembly sub-tests: there were only low and insignificant correlations with the vocabulary, information and comprehension sub-tests.

These results were confirmed by Goodenough and Karp (1961) who performed a factor analytic study of the TRTC, RFT, and the WISC. The three factors which emerged were (1) verbal comprehension (2) attention-concentration and (3) the three field-dependence tests together with the picture completion, block design, and object assembly sub-tests of the WISC. The hypothesis that field dependence is correlated with particular aspects of intelligence was also confirmed in studies of Bieri et al (1958); Elliott (1961) and Spotts and Mackler (1967).

Creativity

Spotts and Mackler (1967) compared subjects on the EFT and on four creativity tests from the Torrance and Guilford batteries. They found that field independent subjects performed better on the creativity tests, both when the groups were matched and unmatched for intelligence. This result was partially confirmed by Bieri et al (1958) who found a significant correlation between EFT scores and originality and elaboration. There was no correlation with flexibility or with fluency.

Bloomberg (1967) reviewed studies on the relationship between field dependence and creativity. He noted that both tests have a similar pattern of intercorrelations with other variables: in particular, both field independence and high creative ability have been found to be associated with low conformity, high level of incidental learning, relative lack of repression, risk-taking, and low identification with the mother (among males).

Bloomberg also noted Crutchfield's argument that field independence might to some extent hinder creativity:

"Analytic perception is sometimes the enemy of creative insight. What may be needed is a free spontaneous look at the phenomenon, a child-like apprehension of what is there, an attitude of what may be called disciplined naïveté".

Bloomberg also noted Witkin's (1965) observation that some field-independent subjects function consistently at a highly differentiated level, whereas others shift from one mode of functioning to another according to the situation. These had been labelled, respectively, as "fixity of functioning", and "mobility of functioning". Bloomberg therefore hypothesised a curvi-linear relationship such that field-independence is necessary but not sufficient for high creativity: field independence and mobility of functioning are both necessary and sufficient.

Personality

Field dependence was originally conceived of primarily as a personality dimension, and its relationship to other aspects of personality have been extensively investigated. In one of the earlier studies, Witkin et al (1954) correlated field dependence with a large number of personality indices based on Rorschach, TAT and interview data, and concluded that there was strong supporting evidence for his theoretical hypotheses. However, this study has been criticised on methodological grounds and theoretical ground by Gruen (1957). Gruen noted that Witkin's scoring of projective data simply involved a summation of pluses and minuses, which indicated the presence or absence of particular signs, and that these scores had been used in the correlations. He pointed out that the scores were probably no more than nominal data, in that a given additional "point" might refer to one of a number of personality characteristics, and that they were therefore inadequate input to correlation analyses.

Gruen also noted that the direction of causality between perceptual processing and personality could not be inferred from correlations, as a full support of Witkin's position would require. He suggested that a person-centred approach was required, in order to investigate the full complexity of personality-perception inter-relationships. Attempts to replicate Witkin's results have produced data at lower levels of significance.

The main personality dimensions which have been investigated in connection with field dependence are as follows:

1. Activity vs. passivity. This was originally hypothesised as one of the main dimensions in field dependence. Partial confirmation of a relationship was found by Young (1959). However, League and Jackson (1961) found that activity in perceptual tasks was not related to activity in other measures of personality.

2. Orientation towards inner life. This was originally the other main dimension. Again partial confirmation of a relationship was found by Young (1959); he found that introspectiveness correlated at a significant level with the RFT, but not with the EFT. Evans (1967) found that extroversion as measured by the Maudsley Personality Inventory (MPI), was associated with field dependence on the EFT, a result which confirmed Marlowe's (1958) demonstration of a relationship between intraception and EFT scores. However, Elliott (1961) found no relationship between EFT/RFT scores and "psychological mindedness" or "self-concept differentiation" and Bieri and Messerly (1957) found a relationship (contrary to hypothesis) between field dependence and introversiveness. In conclusion, a link with introversion and inner vs. outer orientation seems to be demonstrated, although the usual proliferation of non-significant results and the occasional contrary result suggest caution in accepting an oversimple interpretation.

3. Independence. Witkin et al (1962) reported a number of findings confirming an association between field-independence and independence. Linton (1955), using the converse of independence, conformity, found that field dependent subjects showed more conformity in studies of attitude change and the autokinetic phenomenon. Elliott (1961) measured independence of judgement, and obtained a significant correlation with Thurstone's CFT but not with the EFT or RFT. Marlowe (1958) correlated EFT scores with dimensions from the Edwards Personal Preference Schedule (EPPS): he obtained partial support for theoretical hypotheses in a significant correlation with Succorance, but did not find any relationship with Autonomy and Dominance. Marlowe's study was replicated by Dana and Goocher (1959), who found no significant correlations at all. Ohnmacht (1968) and Johnson, Neville and Workman (1959) correlated independence as measured in the 16PF with field dependence, with only partial success: Ohnmacht found no relationship, while Johnson et al obtained a significant correlation only for females. Here again, it is difficult to draw any general conclusions: the numbers of studies who have found a relationship on the one hand, or who have found no relationship on the other, are almost perfectly balanced.

Sense of Identity

Field dependence/independence is also said to predict the sense of individual identity. Witkin (1965) describes the correlations as follows:

> "Persons with a more articulated or more global mode of cognitive functioning also differ in an important aspect of the self, namely sense of separate identity.

Persons with an articulated cognitive style give evidence of a developed sense of separate identity - that is to say they have an awareness of needs; feelings; attributes which they recognise as their own and which they separate from those of others. Sense of separate identity implies experience of the self as segregated. It also implies experience of the self as structured: internal frames of reference have been developed and are available as guides for the definition of the self. The less developed sense of separate identity of persons with a global cognitive style manifests itself in reliance on external sources for definition of their attitudes, judgements, sentiments, and of their views of themselves".

This description was substantiated by a number of experiments. For instance Konstadt and Forman (1965) observed that children of a global cognitive style who were taking tests under conditions of stress looked up at the face of the examiner twice as often as children of a differentiated cognitive style. Crutchfield, Woodworth and Albrecht (1958) found that people with global functioning were relatively better at recognising and recalling faces of people they had been with earlier. Both these results suggest emphasis on external sources of definition of self and ideas. Messick and Damarin (1964) found that field dependent subjects showed greater incidental learning in a learning task where the material was human faces; but this relation was found to be reversed when non-human incidental material was used (Witkin et al 1962). This reliance on external and incidental material for definition of own ideas was also demonstrated in Linton's study (above) which found that field dependent subjects were more susceptible to conformity pressures in an autokinetic situation than were field independent subjects. The studies of the association between field dependence independence are also relevant here.

A study by Winestein (1964) correlated the twinning reaction with field dependence/independence. The strength of the twinning reaction is assessed by evaluating the degree to which the attitudes, feelings and beliefs of twins are congruent: the extent to which one twin perceives himself an integral part of the twinship rather than as an individual. Twins rated high on the reaction were found to be field dependent and global in body concept as measured by the Figure Drawing Test. These two measures were also predictive of whether or not the twins dressed alike. Lewis et al (1966) noted a relationship between field dependence/independence and dreaming in a study of dreaming using rapid eye movements (REM) and EEG stage 1 sleep. Field dependent subjects were found to dream significantly more often of the laboratory situation. They also showed a greater number of dreams to be concerned with their relationship to the experimenter.

Cognitive Defences.

Witkin also hypothesised that where cognitive defences are concerned, individuals with global functioning should exhibit equally global defences, such as denial or massive repression. On the other hand, differentiated individuals should show more specific defences, such as isolation, sensitisation, and rationalisation. Differentiated individuals should find it relatively easy to split thoughts and feelings, and maintain themselves intact from external emotional disturbances. This hypothesis has been supported in a number of experiments. For instance, Minard (See Witkin, 1965) performed a perceptual defence experiment in which the subject was presented with emotional words in a tachistoscopic recognition threshold task. The perceptual defence effect was only demonstrated in field dependent subjects: field independent subjects produced no significant differences between recognition thresholds for neutral and emotionally loaded words.

Other experiments have supported the cognitive defence hypothesis in demonstrating that field dependent subjects are more likely to repress or forget dreams. (Eagle, Linton - both reported in Witkin, 1965; Schonbar, 1964). Also, Witkin (1965) reports analysing data from Lewis et al's (1966) experiment in which the subject had to keep a home diary of dreaming. He found that 7/8 of the field dependent subjects, and only 1/8 of the field independent subjects were non-reporters, and consequently repressers. (Recent studies on REM and desynchronised EEG activity in sleep strongly suggest that everybody dreams every night, - Dement and Kleitman, 1957).

Type of Pathology

Further evidence that cognitive defences are related to articulated bs. global functioning can be inferred from research on clearly pathological cases. If global styles are associated with non-specific defences like repression and denial, it might be probable that they will also be

associated with non-specific pathological features.
Alcoholics have been found to be markedly field dependent, whether they are heavy drinkers, full-scale alcoholics or reformed alcoholics. Alcoholics also show a global body concept on the draw-a-figure test (Bailey, Hastmeyer, and Kristofferson, 1961; Karp and Konstadt, 1965; Karp, Poster and Goodman, 1963). Alcoholism is often characterised as a dependency problem, in which an individual is withdrawing from emotional stresses in a clearly undifferentiated and non-specific way. Other dependency problems, such as obesity, ulcers and asthma in children, have also been found to be associated with field dependent or global functioning, (respectively, Gordon, 1953; Pardes and Karp, 1965; Fishbein, 1963). Taylor (1956) has noted that hallucinators tend to be field dependent whereas patients who are subject to delusions tend to be field independent. Hallucinations are known to be relatively non-specific, and imply dissolution of ego boundaries, whereas delusions have a greater logical structure, and do not represent as great a fusion between self and non-self.
On the other hand, articulated cognitive styles are found among many paranoids (Witkin et al 1965), obsessive-compulsives (Zukmann, 1957) and ambulatory schizophrenics with well-developed defensive structures (Korchin, reported in Witkin, 1965). These maladies are obviously all highly differentiated and specific in nature, and dissimilar from the type of illness which appears to fit field dependent patients.

2: 3: 4 Psychological Differentiation and Cognitive Complexity

Above is an impressive array of studies demonstrating the construct validity and reliability of the dimension of differentiated vs global functioning, or field dependence vs field independence as it is more often called. Despite some ambiguities and contrary findings, one is drawn to the conclusion that it has a considerably greater generality and predictive power than is usual with hypothetical constructs in the domain of individual differences. Like cognitive complexity, it also appears to have an a priori potential in the prediction of academic performance, and other educational criteria. One can make a good argument that education largely consists of improving the differentiation and integration of the students' cognitive model of a sector of reality, and such a process would obviously be affected by the capacity for differentiation, or the habitual level of differentiation, of the cognitive system which is its object. The reported test-retest reliabilities of the RFT and FDT (Table 2:2) unfortunately suggest that this base level of differentiation is resistant to educational processes, although further research is required before this suggestion could be called a conclusion. The personality and other correlates of psychological differentiation would also seem to be a priori predictors of educational processes: if a University education attempts to stimulate abilities of critical judgement, and confidence in that judgement, personality characteristics of independence, inner orientation and inner self-definition are obviously relevant.

Thus psychological differentiation would seem to be potentially useful as a predictor of academic performance. Indeed, at least one study (Spotts and Mackler, 1967) has found a correlation between RFT scores and academic aptitude. However, academic aptitude is not the same as academic performance, and the results are not a foregone conclusion.

As regards psychological differentiation and cognitive complexity, both concepts originate in the theories of cognition propounded by Lewin and Werner. Both concepts explicitly rely on the hypothetical construct of differentiation, and to some extent on integration. A direct comparison of Witkin's position and Harvey, Hunt and Schroder's (1961) formulation was made by Haronian and Sugarman (1967), who suggested the following correspondence:

Harvey, Hunt and Schroder		Witkin
Stage 1	Most concrete	Field dependent
Stage 2		Field independent-fixed functioning

Despite this theoretical similarity, however, there is little direct evidence of correlation, and the two lines of research tend to have focussed on different cognitive domains. Whereas field dependence has been investigated in connection with perceptual processing, personality, and sense of identity, the dimensions of cognitive complexity have been heavily restricted to the domain of interpersonal perception and impression formation. The writer has not been able to find a study which has correlated the two sets of measures, although the lack of such a study seems remarkable considering the prominence and similarity of the dimensions. However, other intercorrelations of cognitive styles, and inter-correlations of different tests of cognitive complexity, suggest that a significant association with measures of field-dependence is unlikely.

CHAPTER 3

OTHER COGNITIVE STYLE DIMENSIONS

3: 1 Authoritarianism and Dogmatism

The concept of the authoritarian personality was first proposed in the context of a study of prejudice against Negroes and against Jews (Adams et al 1950). However, the study was soon widened to include political attitudes and numerous other projective and objective indices. Ultimately, the authors collected sufficient evidence to suggest a common syndrome of personality characteristics which they labelled "the authoritarian personality". Stugner (1961) lists the chief attributes of this type of person as below:

1. Conventionalism. There is a rigid adherence to established middle-class values.

2. Authoritarian aggression. A readiness to condemn, reject and punish anyone who violates these conventional values.

3. Authoritarian submission. Submissive and uncritical acceptance of these values and authority.

4. Anti-intraception. Hostility to imaginative or tender-minded attitudes.

5. Superstition and stereotypy. Beliefs in mystical and all-powerful agents and rigid categorised thinking.

6. Power and toughness: a focussing on weak-strong, leader-follower relationships and an over emphasis on the importance of domination and power. A corollary of this is that the authoritarian indulges in exaggerated assertion of his own power and toughness.

7. Destructiveness and cynicism. A generally high

hostility level, a belief in the inherently evil nature of man and a general pessimism about the worthwhileness of the human race.

8. Projectivity. A projection of inner hostility onto the environment, leading to the belief in the "law of the jungle", and in the world as a wild and dangerous place.

In the development of the F scale, the authors inter-correlated 435 items relating to the above characteristics, obtaining an average correlation of -0.13, and a range of correlations of -0.05 to +0.44, thus giving some indication that the characteristics hang together as a syndrome.

While the theory of the authoritarian personality was based on psycho-analytic concepts and on the displacement theory of prejudice (put simply, that prejudice is a displacement of inner hostility onto a 'scapegoat' minority group), and while much of the research is concerned with the emotional and developmental determinants of authoritarianism, the dimension clearly bears some resemblance to cognitive complexity. Both high scores on authoritarianism and low scores on complexity are thought to relate to 'defective' child-rearing practices ('cold' unemotional parents, harsh discipline, etc.), in both are hypothesised to have a moderate negative correlation with intelligence, and both are believed to indicate a rigidity in the construing of the world. Furthermore, Rokeach (1954) systematically 'translated' the theory into American cognitive-theoretical terms and extended it to the structural dimension of dogmatism. The two dimensions should therefore be correlated at a moderate level: high levels of correlation are theoretically precluded by the difference of emphasis of the tests, in the case of cognitive complexity on a purely structural variable, and in the case of authoritarianism on particular domains of content.

The F scale consists of a number of statements with which respondents are asked to agree or disagree. The statements have been described as "cliche-ridden" (Lee and Warr, 1969), and as "written in a vague and pompous style" and including "glittering over-generalisations" (Stagner, 1961). These are not necessarily criticisms, as they might be hypothesised to be the type of statements with which authoritarians find themselves in agreement. A more serious criticism of the F scale is that it is univalent: a high score is always obtained by agreement with the items, and could thus result from acquiescence as much as from authoritarian beliefs. To demonstrate this, Jackson, Messick and Solley (1957) prepared a set of negative items which involved reversals of the usual positive items: they found that the positive scale correlated with its negative counterpart at +0.35. Bass (1955) showed statistically that at least 25% of the variance in F scale scores was attributable to the acquiescence response set, and Stagner (1961) writes that "many psychologists suspect that the 25% figure is low" (p.254). Perhaps an equally serious criticism is that the items of the F scale are now highly dated: as a measure of cognitive content, which was moreover developed in the years just after World War II, the Fascist characteristics of the totalitarian regimes of Hitler and Mussolini naturally had a disproportionate effect in the theory of the authoritarian personality and in the development of the F scale. Attitudes have changed to such an extent in the 20 years since the scale was developed that modern authoritarians, perhaps with left-wing beliefs, are unlikely to be detected by the scale.

In response to these criticisms, Lee and Warr (1969) have developed a modern and balanced version of the F scale (Bal.F) which is more suited to use on English samples. They claimed that sentence reversals used by Jackson et al (1957) were inadequate, and that negative statements invited different response biases from positive statements. They therefore developed a 30 item questionnaire which involved a positive statement of beliefs, but where half of the beliefs are negatively-valenced with regard to authoritarianism. They report a test-retest reliability of 0.82 over 6 weeks, and split-half reliabilities for a variety of categorisations of the items, most of which are in the 80's. When split between the positive - and negative - valenced items, the reliability was 0.56.

It is necessary at this point also to mention the D scale (Rokeach, 1960), although dogmatism has not been researched in detail for this thesis. Rokeach (1960) criticised the F scale, and the conceptualisation of authoritarianism, as overemphasising particular domains of content, and being particularly geared to Fascism and right-wing viewpoints. He postulated dogmatism as a structural variable of cognitive systems, and hypothesised that dogmatic belief systems would be as prevalent in the extreme left-wing of politics as on the right. This was confirmed by Barker (1958), who showed that 'authoritarianism' can be found in conservatives, radicals, and middle-of-the-roaders. He also found that right wing dogmatics and left wing dogmatics were equally in favour of censorship: they differed in who they wanted to censor.

One would hypothesise that there would be a moderate correlation between the D scale and the F scale. This has been confirmed by Pettigrew (1958), and Kerlinger and Rokeach (1966): reported correlations are usually in the 40's. Lee and Warr (1969) found only an insignificant correlation between the D scale and the Bal.F, although a significant correlation of 0.41 was obtained with the positive-valenced items only. Lee and Warr note that this may be due to differences in subject samples.

There have been investigations of the predictive validity of authoritarianism to a wide number of situations, amongst which are the following:

1) Parental discipline: a feature of the original theory was that authoritarianism develops in response to the child-rearing style or type of discipline. Indirect evidence was published by Hart (1957), who found that mothers who score high on the F scale tend to use more physical punishment, ridicule and threat. To the extent that one can assume that these mothers are exhibiting behaviour which they learned from their own parents, this can be taken as supporting evidence. Mussen and Kagan (1958) also reported indirect supporting evidence. Conformist and non-conformist children from an Asch-type conformity experiment were questioned about their relations with their parents.

They found that conformists, more often than non-conformists, perceived their parents as cold, punitive and rejecting. Evidence is quoted below to show that conformity is related to authoritarianism.

2) Social class. Roberts and Rokeach (1956) found that authoritarianism (F scale) correlated at -0.24 with income, and at -0.45 with educational level. Similarly MacKinnon and Centers (1956) found that manual workers were more authoritarian than non-manual workers. One suspects that this result may be partly due to the "glittering over-generalisations" of the F scale, which are likely to be rejected by the highly educated, regardless of their degree of authoritarianism.

3) Conformity. One of the most extensive studies of the personalities of "conformists" was that of Crutchfield (1955). He presented 100 Air Force Officers with a variety of personality tests and a number of situations based on Asch's (1952) paradigm. Even though these men would presumably have good leadership abilities, Crutchfield found no difficulty in inducing conforming behaviour to incorrect or unlikely stimuli. An examination by a staff of psychologists over several days suggested that conforming subjects were "submissive to authority, narrow of interests, over-controlled, vacillating, confused under stress, and lacking in self insight". Such a personality description bears strong similarities to the authoritarian personality, and indeed Linton (1955) showed directly that high F scorers do tend to conform more than others. However, with regard to Crutchfield's

study, the element of subjectivity in the personality assessments renders the results particularly prone to experimenter effects, and at a time when experimenter effects were less discussed and less compensated for: the conclusions should therefore be treated with caution.

4) Judgement. DeWit (1955) used the semantic differential technique to check the dimensionality of authoritarian attitudes. He predicted that authoritarians, more than non-authoritarians, should use dimensions of conventionalism, anti-intraception, and power as equivalent to the good - bad dimensions. These four dimensions were used to judge various stimulus objects, and a measure of the relationship between the dimensions was calculated. DeWit confirmed his hypothesis that "conventional" is equivalent to "good", when used by authoritarians: dimensions of power and anti-intraception, however, did not differ with degree of authoritarianism. Although it is usually assumed that more authoritarian individuals should be less accurate in their judgements, Nebel (1967) - reported in Lee and Warr, 1969) reasoned that they should be more accurate in a highly formalised judging situation where there is restricted information about the stimulus-objects. Nebel studies the Princeton "Bicker" interaction, which is a "somewhat stereotyped selection process whereby undergraduate students find their way into one of the Princeton dining clubs" (Lee and Warr, 1969, p.123). Using the Bal.F, he predicted that more authoritarian interviewers would be able to predict the candidates' choice of club more accurately. The observed correlation between authoritarianism and predictive

accuracy was +0.40 ($N = 34$, p 0.025).
Another study using the Bal.F (Lee and Warr, 1967, p.124) analysed the dimensionality of judgement about 26 stimulus persons. An index was calculated from factor analysis results to show the amount of emphasis the judges placed on the evaluative dimension. The prediction that more authoritarian individuals should put greater emphasis on the evaluative dimension was confirmed by a significant observed correlation of +0.43.

5) Resolution of Inconsistency. Predictions from the authoritarian personality to ways of handling dissonance or inconsistency are the same as the predictions for individuals with concrete or simple cognitive systems. In the section on cognitive complexity, it was noted that research has tended to find that less complex subjects show a greater recency effect, ignoring information in the critical communication and favouring the later dissonant information. On the other hand, more complex individuals use and integrate the information from both communications in their final judgements, while more complex subjects indulge, in effect, in repression or denial. Similar results using differentiated bs. global functioning, and authoritarianism and dogmatism, as predictors, underline the relatedness of the different formulations. Thus Foulkes and Foulkes (1965) found that more dogmatic subjects as measured by Rokeach's D scale, produced either a primacy or a recency effect. On the other hand, less dogmatic subjects compromised, and took account of both types of information. A similar result was found by Steiner and

Johnson (1963) when a confederate of the experimenter was critically presented in a favourable light, and then produced behaviour designed to antagonise the subjects. The criterion predictor of primacy-recency effects was authoritarianism as measured by the F scale.

A more direct demonstration of denial as predicted by F scale scores was given by Harvey (1962). In this experiment, high and low F subjects were presented with fictitious ratings of themselves, one by a friend and one by a stranger. The ratings involved such characteristics as friendliness, sincerity and considerateness. Some of the subjects were given ratings which were lower than their self ratings and others were given ratings which were identical to their self ratings. All subjects then filled in a post-test ratings and the attitude change was measured. Harvey found a negative correlation between authoritarianism and the amount of change in the self ratings. Authoritarianism also correlated positively with (a) the estimate of how well the source knew the subject, (b) denial that the source was angry with the subject, and (c) denial that the source had made the negative ratings.

6) Extremity of response. It has long been believed that subjects vary systematically in the extremity of their responses to questionnaire and other rating scales, and it has been suspected that this response style is linked with a personality type or trait. O'Donovan (1965) made a comprehensive review of the research but concluded that there was no strong evidence supporting either belief. However, little research had been done. If there is such a response

style, it would seem likely that it would be associated with authoritarianism, and with cognitive complexity. Early studies showed conflicting results. Magar (1960) found a significant correlation between F scale scores and extremity of semantic differential response. However, the observed correlation was only 0.22, with an N of 315, suggesting a very weak relationship. Peak, Muney and Clay (1960), found no relationship between these variables in a similar study.

A more recent study (Warr and Coffman, 1970) used the Bal.F, the D scale and the PCT (measuring abstractness vs. concreteness) to predict 10 measures of extremity of response. Their results led them to conclude that (a) extreme responding is a general response style, spanning the 10 different measures, and (b) there was no relationship between the personality prediction and the extreme response. Warr and Coffman then postulated a model of performance, which included an intervening variable of involvement in the task. Their previous results were rewarded, and another experiment was performed, both of which showed that all three personality measures predicted extremity of response in condition of high involvement.

7) Authoritarianism and other cognitive styles. The research reviewed above suggests that authoritarianism is related in some way to cognitive complexity, differentiated vs. global functioning, and dogmatism. This relationship is both theoretical and at the level of experimental predictions to objective situations. The question remains as to how well the actual measures inter-correlate. As usual, the evidence is equivocal.

Vannoy, (1965) found that authoritarianism correlated with the Sentence Completion Test, measuring abstractness vs. concreteness, at 0.01, and with the Modified Scott test, measuring cognitive differentiation, at -0.04 (see Table 2:1). However, it should be noted that the measure of authoritarianism consisted of 10 statements from the California F scale, together with 10 other negative-valenced statements: Lee and Warr (1969) criticise such statements as not being equivalent to positive statements <u>for</u> a negative-valenced issue. Lee and Warr (1969) found correlations between the Bal.F and Paragraph Completion Test, and Princeton Objective Test (both measuring abstractness vs. concreteness) of respectively -0.32 and -0.16. The former correlation only is significant.

Studies of the relationship between authoritarianism (F scale) and field independence have generally found a negative correlation (Linton, 1952); (Jackson, 1955); (Rudin and Stagner, 1958); (Pollack, Kahn, Kasp and Fink, 1960), although one unpublished study (Mednick, reported in Witkin, 1965) purports to show a curvilinear function in which extremes of field dependence and independence are both related to high authoritarianism. However, Messick and Fredericksen (1958) found no relationship between authoritarianism and concealed figures detection: similarly, Lee and Warr (1969) report a correlation of 0.01 between the Bal.F and the EFT. Furthermore, Hellkamp and Marr (1965) found no relationship between RFT scores and dogmatism.

A variety of studies have confirmed that there is a low negative correlation between F scale scores and intelligence

(Adorno et al 1950); (Cohn, 1952); (Messick and Fredericksen, 1958), and this is also the case with the Bal.F and intelligence (Lee and Warr, 1969). Lee and Warr (1969) also report correlations of the Bal.F with the Gough-Sanford Rigidity Scale, and with the Tomkins Conservatism Scale, of, respectively, +0.39 and +0.51: these results give further evidence of the construct validity of that instrument.

8) Authoritarianism and Educational Indices. On the assumption that authoritarian individuals have a less open cognitive system, are more rigid and have a lower tolerance for ambiguity, one could predict that they will have different reactions to education. They should prefer courses and subject matter which have a clearer structure and less ambiguity, and so also with teaching methods and learning situations. Their performance, also, should be lower than less authoritarian individuals, except possibly where the material is very highly structured.

There is little empirical evidence against which to test these predictions. However, one unpublished study, (Pohl, 1967) (reported in Lee and Warr, 1969) examined the academic preferences and performance of 152 students at Princeton University. With regard to subject performance, Pohl assigned scores of 3, 2, and 1, respectively to the Faculties of (a) science, engineering and maths., (b) social science, and (c) humanities, and found a highly significant correlation of +0.33 with Bal.F scores. He also correlated the Bal.F scores with academic performance on "core" studies. These "core" studies are subjects taken by all students. The correlations for the core studies, classified according to the Faculty which administered the study, were:

Humanities	Social Science
(N = 40)	(N = 101
r = -0.38 (p 0.01)	r = -0.36 (p 0.01)

Natural Science

(N = 65)

r = -0.18 (n.s.)

Here again the correlations are as predicted. In two faculties, authoritarian individuals perform less well than non-authoritarians. And if one takes the trend across the three faculties, the decrement in performance increases as the course structure decreases, until there is little difference in the highly structured natural sciences course. Studies examining the relationship between dogmatism and academic performance have found conflicting results, with variations across subject areas. Ehrlich (1961) observed correlations of -0.30 to -0.54 with objective test performance in sociology, at the Ohio State University. However, Christensen (1963) found no correlation between dogmatism and academic aptitude or psychology test marks. Costin (1965) repeated Christensen's observations, with the same conclusions. The largest study was by White and Alter (1967), who calculated correlations between dogmatism and performance in psychology for 2099 students in 14 classes at the University of Utah. The observed correlations were consistent in sign, but varied in size from 0.13 to -0.52. The weighted oversize correlation was -0.18, which was significant.

Finally, Smithers and Batcock (1970), in Great Britain, found that dogmatism was significantly related to the performance of Social Science students, but unrelated to the performance of Health students.

The trend of the results appear to suggest that a weak relationship exists between authoritarianism and dogmatism, and academic performance, and that other intervening variables need to be taken into account. Furthermore, it would be unwise to assume a direction of causality: it would seem likely that authoritarianism and cognitive complexity would both be changed by University education. The subjects of study would serve to make the cognitive system more differentiated and integrated, increasing the sophistication of students and discouraging black and white or absolutistic judgements. University environments are also more left wing than the general population, and are likely to present strong social pressures against authoritarians. The cognitive style/education interaction is thus likely to be particularly complex in this case.

3: 2 Repression vs. Sensitisation

Considerations of the different ways in which subjects can resolve cognitive dissonance have led to suggestions that individuals may prefer, or habitually use, one mode rather than another. It is argued there is a basic difference between those individuals who sensitise themselves to dissonant information, in order to develop superordinate structures capable of handling it, and those individuals who repress the information or deny it. Byrne's (1961) Repression/Sensitisation Scale is one attempt to measure this hypothetical style. It consists of a 156-item tests, with material largely drawn from the MMPI. The scale is generally agreed to be fairly reliable (Byrne, 1964), although its construct validity is doubtful (Christie and Lindauer, 1963). Very little research has been done using this scale (for a review, see Glass, 1968), and it would not appear, a priori, to have much relevance to education. It is mentioned here solely because of its bearing in dissonance resolution and on cognitive defences, as predicted by the dimensions of cognitive complexity and authoritarianism.

3: 3 Conceptual Differentiation and Category Width

There are a number of dimensions which have emerged from the recent study of concepts and categorisation (Bruner, Goodnow and Austin, 1956) which hold the possibility of coming together with the more global dimensions above. Typically these dimensions are very closely linked to the tests used to measure them, and their investigators have been more concerned to catalogue the individual differences involved in the psychology of judgement than in postulating more global principles of cognitive functioning. For this reason, most of the research has been involved in defining the dimensions, and extracting further dimensions which might be involved in a particular test performance, and there has been practically no research on any predictive validity to other situations. Consequently, there are few inferences one can make from this research to educational situations. It is felt necessary to at least mention them here, however, as they hold considerable promise for the future.

One of the most prolific sources of research in the area is the Menninger Foundation studies (Gardner, 1953; Gardner et al 1959; Holzman, 1954; Holzman and Klein, 1951). They took

Wulf's (1938) early concept of levelling and sharpening, which referred to the way a memory trace could be modified over time, and suggested that it would also explain individual differences in perceptual and memory organisation. Its operational definition is in terms of the Schematising Test, a modification of Hollingworth's (1913) Squares test, in which subjects have to judge the absolute size of squares projected onto a screen. A number of groups of stimuli are presented and judged, a group being changed into the next group by omitting the smallest of the squares, adding a larger square, and changing the order of presentation. Measures of the 'lag' of the subjects in catching up with the progressive change, and of 'ranking accuracy' are used interchangeably as an expression of the levelling-sharpening dimension, despite little evidence that they correlate. An explanation of the subject's performance is given in terms of assimilation and contrast effects. There is some supporting evidence for the dimension in terms of time error assimilation effects (Gardner et al, 1959), but much non-supportive evidence. Vick and Jackson (1967) found little internal consistency between different measures of levelling-sharpening, and little relation to other factors: they concluded that the dimension probably needs 'drastic redefinition'.

Further work at the Menninger Foundation has revealed a dimension originally called equivalence range (Gardner,1953) and later conceptual differentiation (Gardner and Schoen, 1962). This refers to the number of categories used in sub-dividing a meaning domain, and should be distinguished from the more global 'cognitive differentiation'. It is measured by tests such as the Object Sorting Test, and the Photo Sorting Test, in which the subject is given approximately 50 stimuli, and asked to group them into categories "in the way that seems most logical, most natural and most comfortable" (Gardner, 1953). Subjects also have to give their reasons for the categories used* and the degree of abstraction used in the categorisation.

In a factorial study, Gardner and Schoen (1962) have established three independent factors:

1. differentiation
2. spontaneous abstraction
3. capacity to abstract.

Differentiation has further been found to correlate with Witkin's field dependence/independence (Gardner, Jackson and Messick, 1960), although it has no relationship with authoritarianism (Vannoy, 1965).

While these dimensions appear conceptually similar to cognitive differentiation and cognitive integration or abstraction, a consideration of the tasks used in their measurement will show that the latter are much more complex, and in fact there is a zero correlation between conceptual differentiation and Bieri's cognitive complexity (Gardner and Schoen, 1962), a result confirmed by Vannoy (1965). It is perhaps worth mentioning that

---

*This is not true of the work of some investigators, e.g. Sloane, 1959; Marrs, 1955. The importance in this lies in the requirement that the experimenter detect 'waste-paper basket' categories, consisting of all items that fit nowhere else. Gardner counts each item in such categories as a separate category.

Glixman (1965) has found considerable consistency in measures of conceptual differentiation across different meaning domains.

These studies of conceptual differentiation can also be compared with the studies of category width, or bandwidth. The dimension involved here, narrow vs. broad categorisers, is measured by the Dots Test (Bruner and Tajfel, (1961) and the Lines Test (Tajfel, Richardson and Everstine, 1964) as well as a number of similar tests. Here the subject is presented with a number of slides, each of which contains a number of dots. The number of dots ranges from 20 - 28, and the subject merely has to say whether there are 20, 21 or 22 dots, or whether there are more than 22 dots. The score is the percentage of responses to category 1 (20, 21 or 22 dots). The consistency between different tests is moderate to poor, and predictions to other tests are also poor (generally the results are in the predicted direction, but non-significant). Here again a process explanation in terms of assimilation and contrast effects is

The dimension is as yet too ill-defined to be useful in predicting judgemental behaviour of the complexity likely to occur in educational situations, but may become viable if future research can control the cognitive and motivational 'moderator' variables which may be masking the underlying dimension.

3: 4 Strong vs. Weak Automatisation, and Conceptual vs. Perceptual Motor Thinking.

Lazarus et al (1957) and Broverman and Lazarus (1958) isolated a cognitive style dimension of perceptual-motor dominant vs. conceptually dominant thinking. There was little other supporting evidence for this dimension (but see Hurwitz, 1954; Phillips and Rabinovitch, 1958), and unpublished research (alluded to in Broverman, 1960) suggested that the dimension only emerged with difficult concentration-demanding tasks. With highly practised tasks, Broverman found a second dimension of strong vs. Weak Automatisation. Factorially this dimension is defined by performance on the Stroop Colour Word Interference Test (SCWIT) at one end of the dimension, and the Embedded Figures Test at the other. Both of these tests involve responding to a situation against some sort of interference: SCWIT involves reading out the names of colours, themselves printed in other colours, and the EFT involved identifying a simple figure which is concealed within a more complex one. However, the person who has his responses well automatised will be helped when reading out colour names in the SCWIT, but will be hindered in the EFT. In an experiment, Broverman, (1960), successfully predicted that strong automatisers would be less distracted in automatised tasks, and that conceptually dominant (perceptual-motor dominant) subjects would be less distracted on conceptual (perceptual motor)tasks. It is important to note that this experiment demonstrated the two dimensions only in conditions of distraction, and not in neutral non-distraction conditions, by the use of ipsative scores i.e. the subjects' deviations of scores on particular tasks from their general levels of performance on a heterogeneous group of tasks. These scores serve to give the within-individual biases in ability. His hypothesis was supported, and the dimensions of weak/strong automatisation and conceptual/perceptual-motor thinking were demonstrated in neutral conditions. This demonstration has implications for general personality and cognitive style experiments, where within-individual variation is seldom taken into account (but see convergence/divergence, below).

It is arguable that one of the tasks of education is to get students to use limited capacity processing in novel, concentration-demanding situations, and so 'master' the situation so that rules, concepts, techniques, etc., can be applied automatically. Presumably the strong automatiser would be at an advantage in those subject areas and those situations which embody a higher proportion of 'routine', and those situations which have subsidiary skills on which the cognitive skills are based. Thus subjects like statistics, engineering, architecture, and possibley mathematics, where 'routine' calculations and procedures are a pre-requisite to higher order "thinking" skills, would favour the strong automatiser: in other subjects in the arts and social sciences faculties, where essays, projects, and other high information content products are the examination criteria, the strong automatiser would have no advantage. Similarly with teaching-learning situations, the strong automatiser would excel where there are subsidiary and routine tasks. For instance, in lectures and practicals, but not in project work and library work.

It is interesting to note how this style may overlap with cognitive complexity. It is arguably that the "automatising" of cognitive and motor skills involves the development of structures and representations in that cognitive domain.

If this is so, the strong automatiser should also be a high scorer on cognitive complexity. Very little research has been done on automatisation, and the writer knows of none which has a bearing on this point.

The dimension of strong vs. weak automatisation does appear to have some potential in predicting educational indices: theoretically the inferences are clear. In practice, it may turn out that the dimension is restricted to the relatively simple tasks in which it has been investigated, and it may be inapplicable to the complex types of cognitive skill involved in the field situation of higher education. If it is applicable and indeed is related to cognitive complexity, it may turn out, like cognitive complexity, to have restricted generalisation across different domains of content. These questions need further investigation.

It is difficult to see how conceptual vs. perceptual-motor thinking is related to education, except possibly to the conceptual vs. perceptual-motor emphasis of different subject areas. As the present study is omitting the variable of subject area, the dimension would appear to be of little use for present purposes.

3: 5 Reflection vs. Impulsivity

The line of enquiry was initiated by the observation (see Kagan, 1966) that people differed in the way that they classified visual stimuli. While some subjects categorised on the basis of thematic rules, others used features comprising an individual aspect of the picture. This tendency was measured by the Conceptual Systems Test (CST), in which the subject is asked to state which twom of three stimuli presented, are alike. A typical group of stimuli are

1. a picture of a house with smoke coming out of the chimney,

2. a book of matches,

3. a pipe with smoke curling up from the bowl.

The thematic classification is that 'the matches are used to light the pipe', and is contrasted with the analytic description 'the pipe and the house both have smoke coming from them'. This dimension was originally observed with adults, but most of the subsequent work was performed on schoolchildren. A number of researches (Kagan, Moss and Sigal, 1963; Kagan et al 1964; Lee, Kagan and Rabson, 1963) pin-pointed some of the salient characteristics of the child who produces a large number of analytic concepts, and these are summarised in Kagan (1966):

> An impressionistically consistent cluster of characteristics was possessed by seven to ten-year-old boys who reported many analytic concepts. These boys were less distractible in the classroom, less likely to display task-irrelevant gross motor behaviour on the playground or in a restrictive laboratory situation, and less likely to report many incorrect solutions. They were more likely to become involved in sedentary tasks requiring long periods of concentration, more likely to prefer intellectual vocations that required motoric passivity (e.g. scientist, writer), and typically produce more complex drawings of objects. One of the most objective demonstrations of the relationship between motoric restlessness and analytic attitude was contained in the positive association between analytic concepts and regular, non-variable respiratory rhythms during episodes of rest, and episodes when the boy was attending to simple visual or auditory stimuli (Kagan and Rossman, 1964). Moreover, boys with many analytic concepts demonstrated greater cardiac deceleration when asked to attend to external stimuli....... It appears, therefore, that young boys who prefer analytic concepts are more capable of sustained attention to visual unputs than less analytic youngsters".

This and other research seemd to suggest that there were two factors underlying the production of analytic concepts: a dimension of 'conceptual tempo', labelled reflection/impulsivity, and a dimension of visual analysis.

The former dimension has since been fairly well researched, Operationally, it is the response time in matching a standard stimulus to a set of comparison stimuli, and has been measured by at least three tests. In the Delayed Recall of Designs Test, a simple design is presented for 5 seconds, removed, and replaced by a set of 10 stimuli 15 seconds later. The subject has to select which one is identical to the standard, Measures are taken of response time and errors. The Matching Familiar Figures test is similar to the DRDT, except that all the stimuli, standard and comparison, are presented simultaneously. In the Haptic Visual Matching Test, the child explores a wooden form with his fingers, and then has to select that stimulus from a visual array of 5 stimuli. Measures of errors, response time and palpation time are taken. There is a fairly high consistency of scores across this broad spread of tests (typically r's of 0.3 to 0.6), and a higher consistency of response times. The measures are also fairly stable over 9 week and 17 month periods, with response times more reliable than the other two scores. For this reason, the response times are usually taken as the measure of reflection vs. impulsivity. There was a negative correlation between errors and response time (0.3 - 0.6) and a slightly lower negative relationship between errors and verbal ability as measured by the three verbal subtests of the WISC. Response time was orthogonal to verbal ability. Reflection vs. impulsivity has also been found to be unrelated to field dependence, as measured by the EFT (Kagan, Rossman, Day, Albert and Phillips, 1964), although impulsive children offered far more hypotheses before gaining the correct solution.

Thus it appears that the reflection/impulsivity dimension is a fairly reliable parameter of behaviour with some generality. There is also some evidence which suggests, by inference, that it might have a constitutional bias. Schaeffer and Bayley (1963) found that extremely active one-year-old infants were minimally attentive to intellectual problems at five and six years of age. This result has been confirmed by the Fels longitudinal Study. Kagan and Moss (1962) found that ratings of hyperkinesis at ages 3 - 6 were inversely correlated with ratings of involvement in intellectual activity during adolescence and adulthood. Furthermore, Kagan (1966) suggests that "excessive motor restlessness and distractability at age 8 have their anlage in congenital deficit resulting from minimal and subtle brain damage during the peri-natal and early post-natal period". As hyperkinesis, restlessness, lack of concentration are key characteristics mentioned in the summary of the analytic personality above (and it must be remembered that the reflection/impulsivity dimension is an important determinant of performance in the Conceptual Systems Test), these results suggest that it can in theory be tapped extremely early, and is at least to some extent constitutional.

Thus the reflection/impulsivity dimension has been shown to predict some aspects of performance in the primary school situation. Furthermore, error scores on the MFF, HVMT, and DRDT have been found to correlate with orthographic reading errors in young children (Kagan, 1966). If the dimension proves significant at later developmental stages, it should prove a valid predictor in the tertiary educational situation. In particular, one might hypothesise that the impulsive individual would be more motivated in the project situation, where he is allowed free rein, but more productive in the lecture situation, where his listlessness will be dampened down, and he is forced to listen. Similarly on over-all or examination performance, he should do less well than the reflective individual, who is better able to put in the "solid" rote learning and mastering of detail which are a pre-requisite of good performance. There are obvious inferences to subjects requiring more or less "flair" or "drudgery".

3: 6 Convergent/Divergent Thinking

The distinction between convergent and divergent reasoning was initially introduced by Guilford (1950), although one can trace it further back to Kretschmer's A-type and R-type thinking. It was hypothesised as an individual difference dimension by Hudson (1966), although his work is largely based on Getzels and Jackson's (1962) distinction between the 'high creative' and 'high IQ' schoolchild. Hudson's work involved a large sample of English schoolboys who were tested on the AH5 intelligence test, and a number of American creativity tests, including the Uses of Objects, Meanings of Words, and Drawing Test. Other tests used were Controversial Statements, in which the subject has to write comments to as many or as few of the statements as he chooses, and the Personal Qualities Questionnaire. A boy was classed as a converger, diverger or all-rounder by finding the bias in his abilities towards the IQ test, or towards the open-ended tests. Thus it is feasible for a diverger to have a higher IQ score than a converger, either by virtue of the fact that his open-ended score was so high, or because the convergers over-all ability level was so low. This bias score is functionally much the same as Broverman's ipsative scores (see automatisation dimension, above) although mathematically somewhat different and convergence/divergence can therefore be classified as an intra-individual dimension.

Hudson notes that the three open-ended tests have very little consistency between them, with inter-correlations of around 0.3. Similarly, the different parts of the intelligence test intercorrelate only slightly ($r = 0.4$) and intelligence and open-ended tests correlate even less ($r = 0.2$). Despite this lack of logical consistency within the dimension, the final score for convergence/divergence has some striking behavioural correlates. On a sample of 267 schoolboys, it predicted Arts vs. Science choice in the VIth form at a massive level of significance. "Between three and four divergers go into arts subjects, like history, English literature, and modern languages, for every one that goes into physical science. And vice versa, between three and four convergers do mathematics, physics and chemistry, for every one that goes into Arts" (Hudson, 1966, p.56). This result was confirmed in Australian schoolboys (Cropley and Field, 1968).

What this result means is not exactly clear. As Hudson points out, the open-ended tests do not appear to measure originality or achievement: in selecting the most outstanding personal projects, including a computer to optimise the speed of model racing cars, at a grammar school, the convergence/divergence bias was not an effective predictor*. It is possible that the trivial nature of tests like the Uses of Objects simply does not appeal to the more practically orientated converger. More light can be cast on the dimension by observing some of the other correlates which Hudson found. The dimension not only predicted the quantity of answers to the Uses of Objects and the Meanings of Words, it also predicted qualitative differences. Divergers tended to produce far more rare responses, whereas convergers tended to produce usual and stereotyped answers. In the Drawing Test, (where the subject is asked to "draw a picture in the space below to illustrate the title 'Zebra Crossing'. You can draw whatever you like as long as it appears appropriate".) divergers drew much more unusual themes, and convergers were much more likely to omit people from the scene. Hudson hypothesised that scientists and convergers would be much more likely to frame answers to the Uses of Objects in terms of general properties. In fact, the contrary result was found: artists and divergers excelled in this. Further features of answers to the Uses of Objects were that convergers were much less likely to give responses including violence, but when voilence was included by these people it was of a 'ghoulish' and 'morbid' sort. On the Controversial Statements questionnaire, divergers were more likely to hold opinions shared by only a few of the sample. Hudson notes that this was not because they had 'way out' or 'odd-ball' opinions, but because they thought the issue out more deeply, rejecting the socially stereotyped response. This was only true, however, with statements about human issues: divergers lapsed back into social stereotype where logical puzzles were concerned. Furthermore, divergers tended to express their opinion far more emphatically, both on the Controversial Statements questionnaire (equalitative judgement) and on the Personal Qualities questionnaire. This latter contained items which were thought to load on five dimensions - authoritarianism, rigidity of attitude, social conformity, freedom of emotional expression, and defensiveness. Only the first three of these were correlated with convergence/divergence, the converger taking the higher score on each.

A number of studies have used tests of divergent thinking ability as predictions of educational performance. In Great Britain, Hasan and Butcher (1966) obtained a correlation of 0.62 with performance in English, and a correlation of 0.76 with performance in arithmetic. However, a similar study by Richards and Bolton (1971) concluded that divergent thinking is only slightly related to ability at maths. Studies in the U.S.A. have been similarly conflicting, although significant correlations with science (Cline Richards and Needham, 1963), maths, (Petersen, Guilford and Hoepfner,1963), and general achievement (Ohnmacht, 1966), have been found. Bennett (1973), using 331 British children, used 10 tests of semantic divergent thinking, a convergent verbal reasoning test, and measures of formal English ability, and imaginative story writing. The divergence tests correlated significantly and evenly with both English criteria, and the convergence test predicted formal English and story writing at respectively 0.90 and 0.50. A factor analysis gave supporting evidence as to the separation of the two sorts of abilities.

It is difficult to know what underlies the convergence/divergence dimension. It is almost certainly something other than convergent and divergent thinking. There is very little evidence that the open-ended tests predict creative performance in everyday life (Hudson - projects in grammar schools - above), and some that they do not. In view of this it seems likely that the tests do not tap the converger's flow of thought, either because he thinks them trivial, or because he is not used to expressing ideas in language. One hypothesis that seems to fit the data fairly well is that the convergence/divergence tests are tapping a sort of 'practicality' variable - the old distinction between thinkers and doers. The thinker is fluent at manipulating ideas, following up any suggestion, no matter how nonsensical, merely for the fun of it. The doer, on the other hand only thinks as a means to the end in hand. He cannot afford to allow his imagination full play, or to question the accepted body of knowledge, because if he does, he is left with no time to do anything. According to this argument, politicians should turn out to be overwhelmingly convergent. A concomitant of this reasoning is that convergers should have less tolerance for ambiguity of dissonance, a fact which seems to emerge from responses to the Controversial Statements. This hypothesis is basically the one that Hudson elaborates to a much greater depth of psycho-analytic complexity. Final comment must await further research.

Convergence has also been found to be a predictor in tertiary education. Cropley (1967) has shown that at the senior undergraduate level in Australian universities, those who are outstanding in science tend to be divergers. This result is the reverse of the bias noted in schoolboys in the original work in England, and in Cropley and Field's study in Australia. A longitudinal study being undertaken in Australian universities (Field and Poole, 1971) promises to throw light on this paradox. So far results have come in for the 1st and 2nd year examinations. It has been found that at the university level the entry into science faculties is predominantly convergent, and the entry into arts faculties is primarily divergent - in agreement with the VIth form findings. In the first year examinations, convergers got much better results in the science faculties, as expected, and also, contrary to expectations, in the arts faculties. Field and Poole reasoned that this was the result of first year courses, which consist mainly in the mastering of a body of information, a task much more congenial to the convergent mind. The second year examination results show no bias of achievement however: the diverger attains equally as good grades as the converger. If the trend is continued, the apparent paradox will be solved. In short, as a student proceeds through University, the academic demands made on him change: qualities of application, and devotion to a set and unquestioned body of knowledge, are initially required, to be replaced with a requirement for imaginative and open-ended thinking, in which the fundamentals newly learned by the student are subjected to a fresh look and rigorous questioning. It is tempting to say that initially the converger succeeds, and then he is overtaken by the diverger; however, the possibility that the same student is initially a converger, who changes his style of thinking as the requirements change, has not yet been tested. This possibility is perhaps unlikely, and the results justify Hudson's worst fears, that school VIth forms are presenting a picture of science that tends to repel the very person who will later become the best scientist.

With regard to the social science subjects, one might perhaps postulate that the diverger would perform better there also. However, the converger should be relatively better (i.e. at his optimum, if lower, overall, performance) in well structured teaching-learning situations, such as lectures and assigned problems or practicals, rather in open-ended projects and essays, where non-stereotyped solutions are more valuable. The uncertain nature of the convergence/divergence dimension makes these inferences highly tenuous, however, although still worthy of investigation.

3 : 7 Cognitive Styles and Education : Conclusions

Whilst education is clearly concerned with the development of personality, emotional responses, moral judgements, attitudes, and social behaviour, there would surely be few people to deny that the main developmental responses to education are cognitive. The transmission of a body of knowledge - knowledge of, and knowledge how to - is the matrix which social and emotional changes occur. This being so, it would seem that the relationship between educational performance, and individual differences in, or styles of, cognitive functioning is self-evident. The a priori nature of this general proposition leads to the empirical questions of which particular cognitive styles predict which particular educational indices: from the general preposition one can surmise that the answers to these empirical questions would be an important factor in unravelling and explaining the cognitive processes which occur in teaching, learning, and the growth of knowledge in tertiary education. Cognitive styles, then, would seem to be a better research proposition than personality variables, in the prediction of educational indices, despite the strong research effort on the latter, and the paucity of research into the former.

General logic of a different kind, however, suggests the opposite. General research experience over the past half century suggests that things simply are not as clear-cut as is suggested by the analysis above. The reasoning could validly be called naive and unsophisticated. Cognitive style dimensions, as also personality dimensions, are at best poor approximations to what goes on in the mind. Theoretically similar dimensions are found not to correlate, and diverse dimensions are found to overlap. In addition, intervening variables are so complex, numerous, and invisible, that experimental control is in practice unachievable. The result is that similar, or indeed "identical" studies frequently come up with opposite conclusions. As will be seen in Chapter 5, personality predictors of educational performance rarely achieve correlations higher than 0.3, and even tests designed to predict performance are struck at 0.5: the scant research on cognitive styles have scarcely reached the lower of these two levels. Such predictions account for less than 25% of the variance in educational performance, and are in practice useless for helping individual students or understanding individual processes.

Further research will, of course, help to clear up the theoretical imprecision and empirical confusion associated with many cognitive style variables. There are, however, grounds for believing that the basic concepts used in the analysis of cognitive processes are inappropriate, and that the research approach itself embodies assumptions which conflict with the nature of cognition and of education. It is these possibilities which must be examined next.

CHAPTER 4

PRESUPPOSITIONS OF RESEARCH INTO COGNITIVE FUNCTIONING

4: 1 Introduction

The preceding two chapters have reviewed research on a number of postulated cognitive style dimensions. A number of these dimensions are theoretically related, being clustered around a concept of cognitive complexity, but do not correlate well with one another. The evidence reviewed gave further doubts and contradiction as to the construct validity of these dimensions. Others of the dimensions (for instance convergent vs. divergent thinking, reflection vs. impulsivity) lie in a theoretical vacuum as to their significance in general cognitive functioning. The generally unsatisfactory state of research into cognitive style may be attributable to one or both of the following causes:

1. Measurement error. The tests used do not give an adequate measure of the theoretical dimensions.
2. The theoretical dimensions of cognitive style do not give an adequate representation of individual differences in cognitive functioning.

The present chapter will attempt to explore the second possibility, by explicating the model which research into cognition and cognitive style presupposes, and assessing its

All theories presuppose other theories. The cognitive style dimensions are presented in the context of a model of cognitive functioning, without which they become ad hoc. Even personality research, much of which is empirical and taxonomic in the extreme, makes presuppositions as to the nature of social functioning by using the statistical methods which it does. And research on the general aspects of cognition themselves make presuppositions about cognition by the type of research methods used. Such presuppositions may be explicit or implicit, and veridical or non-veridical. They only become dangerous when implicit, for then their veridicality cannot be assessed. The present chapter will deal with some of the implicit presuppositions made by research into cognitive functioning and cognitive styles, and with explicit presuppositions insofar as their more general implications have not been realised. The author cannot claim to deal with all the implicit presuppositions relevant to the research area, as they ultimately involve questions as to the nature of knowledge itself, questions which are still the subject of active philosophical debate (Polanyi, 1967; Piaget, 1971; Popper, 1972). These questions are beyond the scope of this monograph when empirical data cast no further light.

The chapter will assess the veridicality of the pre-supposed model. Veridicality will be examined against the criterion of the field situation, of the human being in everyday life, and, in particular, in higher education. The more usual criterion of veridicality in psychology is the empirical observation in the research laboratory. This is a useful and worthwhile criterion only insofar as the behavioural mechanisms seen are the same as the behavioural mechanisms seen in the field situation of everyday life. This chapter will argue that the laboratory experiment in psychology has involved a systematic selection from the type of functioning that can be observed in life, and that our model of the organism is distorted in consequence. Whether this argument is accepted or not, most psychologists would surely agree that the ultimate criterion of truth in psychology is life itself.

Finally, the term "element" will be used extensively in the following discussion of cognition. Some theorists (e.g. Festinger, 1957; Peak, 1958) have imbued "cognitive elements" with certain characteristics. Indeed, the present author will infer certain characteristics of elements later in the chapter, but until that point, no pre-suppositions should be made. Initially, "element" will be used as an erudite synonym of "thing".

4 : 2 The Nature of the Experimental Method

The significance of factors extracted in a factor analysis naturally depends on the input variables submitted, and on the assumptions made by the factor analysis model. In the same way, theories of cognitive, and indeed psychological, functioning depend logically on the types of behaviour observed, and on the assumptions or pre-suppositions made by the research method. If the types of behaviour observed are a biased sample from the population of observable behaviour, or if the research method makes unjustified assumptions, the predictive validity of theories in psychology is likely to suffer accordingly. In explicating the model of cognition pre-supposed by cognitive style theories, it will therefore be wise to go back to first principles and examine the research methods of psychology. The paradigm research method in psychology in general, and in cognitive psychology in particular, is the experiment.

The experiment in psychological science is based on the scientific method of the physical sciences, which is, in turn, a refinement of natural logic stating the conditions under which evidence may be said to disprove or support a hypothesis. Briefly stated, and admitting oversimplification, the situation in the pysical sciences is that a hypothesis is deduced from a theory, or built up on the basis of a series of observations, and is then subjected to test. These are respectively the hypothetical-deductive, and inductive methods. The hypothesis will usually involve a relationship between determining variable, and another variable that is determined. When embedded in the experimental situation, these become respectively the independent, or predictor, variables and the dependent, or criterion, variables. The process of testing the hypothesis involves manipulating the predictor variable and observing the change in the criterion variable, whilst maintaining possible confounding variables constant. The variance in the criterion variable can be attributed to two main sources:

1. It is causally linked to the predictor variable.
2. It is causally linked to one or more confounding variables.+

The success of the physical sciences depends on the efficiency of the experiment, and this depends on the efficient control of confounding variance. In psychology, together with the other life sciences, it is often not possible to control the

---

+ There is the third possibility of random variance in the criterion variable. Such a possibility raises questions which are beyond the scope of the present thesis: questions, the answers to which do not substantially affect the present argument. Other sources of variance, such as measurement error, are also consciously ignored here.

confounding variables in the laboratory to any appreciable extent. The development of sampling theory and statistics is an attempt to substitute control by randomisation for physical control. Control by randomisation involves the replication of the observation on a sample drawn randomly from the population of possible observations. The tools of statistics then allow the separation of the experimentally manipulated variance from the randomised confounding variance. Before proceeding further with this analysis, it should be noted that the experimental method already pre-supposes that the determining variables of behaviour can be experimentally manipulated i.e. that they can be predictor variables. This implies the assumption that behaviour is externally referenced: that behaviour is dependent on environmental causes, and consequently can be defined in terms of those environmental causes. If behaviour is determined by forces acting within the organism, and must therefore be defined internally, there is little point in manipulating external variables. The most explicit embodiment of this assumption is, of course, S-R theory. Although S-R theory has now been modified to S-O-R, the external referencing of behaviour is still assumed. In fact, there is little doubt that much of behaviour is in practice externally referenced and objective: to the extent that we all have common needs and operate in a common environment, we must have a common model of that environment which is externally referenced and objective. However, it is equally beyond doubt that not <u>all</u> behaviour is externally referenced. The experiments associated with the New Look theories in Perception illustrate the importance of motivation, expectancy and personality in perception, variables which are not subject to complete experimental manipulation. However, the external referencing assumption is a relatively superficial presupposition in psychological theories, and it is embodied in a more subtle theoretical distinction which is discussed below. The empirical evidence relevant to the veridicality of both these assumptions will be discussed in the latter context.

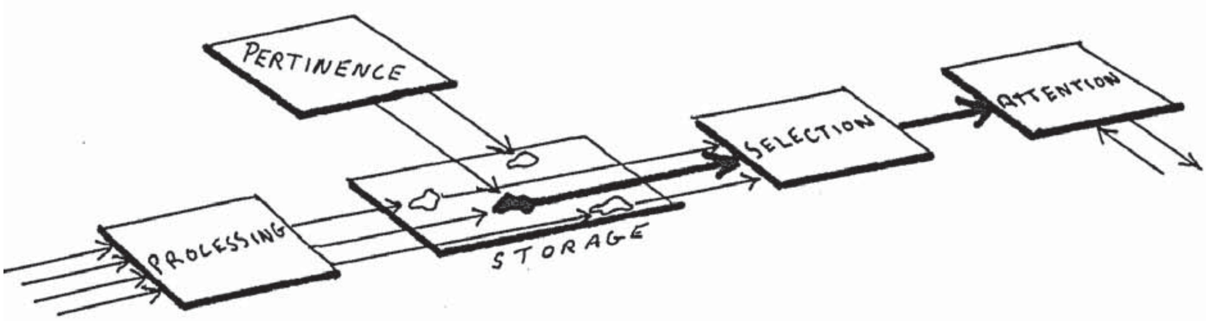
4: 3 Structure and Content

To continue with the analysis of the experiment, it is essential to note the importance of the fact that the predictor variable and criterion variable are defined in terms of the experimental situation. Control by randomisation of confounding variables can only therefore be effected by replicating the observation on a number of different subjects within the same experimental situation. This leads directly to the theoretical distinction between content and structure. Structural elements are those elements which are assumed to be general across all normal members of a species, whereas content elements vary from one individual to another. It is necessary to distinguish between the two because the former are susceptible to experimental investigation and the latter are not: the distinction is necessary for psychology to function as an experimental science.

The point is explicitly made by Schroder et al (1967) "....two distinct classes of information are relevant to the understanding of adaptation. 'Content' variables provide information about the acquisition, the direction, and magnitude of responses..... 'Structural' variables provide a metric for measuring the way in which a person combines information". In more general terms, structural elements include things like eyes, ears,

retinae, attention mechanisms, memory stores, limited capacity processing mechanisms, thought processes, whereas content elements are particular views and sounds, signals, thoughts, concepts, attitudes, beliefs, etc. The structural elements operate on the content elements, and within limit operate irrespective of particular content elements.
Thus theories of memory separate iconic and echoic memories, (Neisser, 1967), short term memory and long term memory (Atkinson and Shiffrin, 1968; Baddeley, 1966), short-term visual and short term auditory memories (Sperling, 1967; Gardiner and Thompson, 1973), in addition to linguistic and motor memories (Morton, 1969). All these distinctions are maintained on the basis of structural information like the type of error made (accoustic vs. semantic confusions) and the amount of information forgotten over different time periods and with different associated memory conditions (primary vs. recency effects, proactive and retroactive interference, free-recall vs. serial recall, interpolated tasks vs. unfilled delays, etc).
These structural elements of memory are assumed to be common in different individuals. Content elements are not entirely independent of structural elements: for instance, the degree of organisation of content is known to affect forgetting (Miller and Selfridge, 1950); Bonsfield, 1953; Deese, 1959), and coding strategies are known to affect short-term memory (Miller, 1956). It nevertheless remains true that the two types of element are defined independently, and that research on the content elements is usually only undertaken by reducing them to structural elements (e.g. dimensions such as "degree of organisation").

The same distinction is applied in theories of attention and perception. The active processing mechanisms are defined in terms of external features. Broadbent's (1958) original filter theory of attention suggested that a channel was selected on the basis of physical features such as a pulse-frequency and intensity. Treisman (1960) produced evidence that the filter could only be partial: some information was passed from the non-attended channels. Deutsch and Deutsch (1963) suggested an alternative attention theory, which assumed that all incoming information was processed to the extent that its "importance" could be assessed, and then the most "important" message would be selected for further processing. Norman's (1970) account of attention followed similar lines. In all of these theories, the nature of the selective filter is defined in terms of externally referenced physical features, precisely what types of information can be selected, how much processing is required before selection, and what selecting mechanisms are required. The theories are abstracted on the basis of normative data, that is to say they are intended to apply to different people. The role of content elements is even more crucial in theories of attention: the main basis for selecting one signal rather than another is in terms of its meaning (at least according to Norman, and Deutsch and Deutsch), but this can only be incorporated in the theory by reducing it to the structural dimensions of "pertinence" (Norman, 1968), or "importance" (Deutsch and Deutsch, 1963), and including a structural pertinence assessor (Norman), + or variable threshold (Deutsch and Deutsch. Thus the content and structural elements are maintained separately. In terms of any real-life criterion, the main burden of explanation rests on the cognitive elements, but they cannot be investigated by experiment. These examples from memory and attention will serve to illustrate the content-structure distinction. It will be apparent that this distinction is general in theories in experimental psychology, without listing those theories here. Indeed, it is so universal and unquestioned that it is initially difficult to conceive of how these theories (and the ones outlined above), might exist without making the distinction. Needless, to say, it is not necessarily present in those areas of psychology which are not based on experiment (e.g. psycho-analytic theories, Piaget's developmental theories) although it may be (the dimensional personality theories, such as that of Cattell, 1965). Up to this point, psychology has made some possibly questionable, but explicit, presuppositions, in order to be able to operate as an experimental science. Now, however, two implicit, and hence unquestioned presuppositions are made:



+ Diagram 4:1 Norman's (1968) Model of Attention

1. The population of observations is assumed to be the same as the population of people. Control by randomisation only gives results which can be generalised if the sample observed is randomly drawn from the population. This is shorthand for the more precise rule that "the sample of observations is randomly drawn from the population of possible observations". As stated above, the experimental method can only involve randomisation across people in the same situation. It omits the possibility of sampling observations across situations but in the same person. This is not possible using the experimental method, because the predictor variable is inevitably defined in terms of the physical experimental situation: it is defined as much by what it is not (the physically controlled confounding variables) as by what it is. The embodiment of experimental effects in theories are thus situation-specific. The theoretical entity of the attention mechanism is thus <u>demonstrated</u> by experiment to be general across people, but specific to the dichotic listening situation in the laboratory. The possibility of theoretical entities which are specific to a single individual but which affect his behaviour in widely differing situations, cannot be investigated by the experiment, because they would be internally referenced and non-manipulable.

2. Structural elements, which are demonstrated to be general across organisms, are assumed to be, <u>without demonstration</u>, general across widely differing situations. They function, within limits, on different content elements in different individuals. The short term memory store is not only present in different individuals, but also operates

on all different types of information, regardless of the significance of that information. The selective attention mechanism operates regardless of content, detecting pure tones in the laboratory, as well as detecting verbal signals at a cocktail party, and allowing a sleeping mother to awake only to the cry of her baby. The attention mechanism and the memory store can be claimed to be demonstrated as organism-general, in that they have been abstracted on the basis of normative data, but they are only <u>assumed</u> to be situation-general.

The experimental method thus involves a biased selection of behaviour from the population of observable behaviour. In the physical sciences, the experimental method prescribes the conditions under which hypotheses are supported or rejected: in psychology, it also prescribes which phenomena are valid subjects of investigation. To the extent that theories of psychology are abstracted from, and tested against, experimental evidence, they inevitably involve a distortion of our knowledge of the functioning of the organism. The question remains as to how serious this distortion is. Does the content/structure distinction seriously affect our understanding of the organism, or is it inaccurate in only minor details? The remainder of the present chapter will argue that the use of the content/structure distinction, and the restriction of research to structural elements, involves a serious distortion of the theoretical concepts of cognitive psychology. To substantiate this argument, it will be necessary to demonstrate the subsidiary propositions:

1. That the determining variables of behaviour are internally-referenced and idiosyncratic, and are therefore not structural.

2. That "structure" and "content" are two aspects of the same thing. These propositions will be examined in the context of a number of areas of psychological investigation.

4 : 4 Perceptual Processes

Proposition: that the determining variables of perception are internally referenced and idiosyncratic

It is necessary at the outset to comment on the term "internally-referenced", which should be taken to mean "internally-defined". Internally referenced variables should be distinguished from variables which are internal in operation, but which are externally referenced. Thus, the hunger drive as a determining variable of behaviour, is an internal variable, but is conceivably externally-referenced exclusively in terms of external variables such as "hours of food deprivation" and "percentage of normal body weight". To substantiate the claim that the hunger drive is internally referenced, it is necessary to show that it operates in a way which is independent of external and objective (i.e. interpersonally agreed) influences. The initial test of whether behaviour is internally or externally-referenced must lie in the evidence on the nature of perception, for if external variables are to determine behaviour they must do so through the perceptual process. It used to be thought, following Locke and the English empiricist philosophers, that perception was based on the purely passive transducing process of sensation, the primary data on which perception operates being "sense-data" which are "given" and invariant across different people and different situations. This tradition is the basis of much of American psychology, to the extent that S-R theorists and neo-behaviourists almost ignore the study of perception in favour of learning. The analogy between the eye and the camera was accepted literally until quite recently. The European tradition, in contrast, developed from different philosophical premises. Musserl, Mach, von Ehrenfels), which emphasised the need to account for the phenomenal world. This position was embodied in the Gestalt school of perception: Gestalt theories implied internal-referencing of perception and behaviour, but in their emphasis on the explanations of good figure in terms of psychological fields of activity in the brain, they went beyond the evidence. It is perhaps unfortunate that the falsity of this reduction led to the rejection of the whole Gestalt position, with the result that many interesting phenomena were ignored for a long time.

It now seems quite clear that neither of these theoretical positions is tenable, and modern views of perception show something of a rapprochment between the two. For instance, the proposition that perception is based on "given" sense-data, which have a point to point correspondence with the physical stimulus elements, is now found to be quite wrong. Forgus (1966) suggests that the perceptual system is hierarchical in nature with more and more information being "extracted" from the input as it proceeds from the lower to the higher levels. He distinguishes five stages:

1. The detection of stimulus energy and a discrimination of change in stimulus energy.
2. The discrimination of a unified brightness or figural unity as separate from the background.
3. The resolution of finer details, which gives rise to a more differentiated figure.
4. The identification or recognition of a form or pattern.

5. The manipulation of the identified form: this happens for example in problem-solving and social perception.

This view of perception as involving successive levels of processing from the lowest to the highest is supported by physiological evidence. Thus at the level of the retina ganglion cells are found which are receptive not just to stimulus, but to a combination of stimuli: these are the so-called "on-off" circular fields, in which the cell responds only to a spot of light in the centre of the field in <u>combination</u> with no light on the surround, or vice versa (Kuffler, 1953; Spinelli, 1966). Progressive specialisation of cells is found throughout the visual system up to the level of the striate cortex and beyond (Polyak, 1927; Talbot and Marshal, 1941; Thompson, Woolsey and Talbot, 1950; Clare and Bishop, 1954; Dorg, 1958; Vaslola, 1961; Otsuka and Hansler, 1962; Polley and Dirkes, 1963). Perhaps the most well-known work of this nature has been done by Hubel and Wiesel (1959, 1961, 1962, 1963, 1965) in which cells at different levels in the occipital cortex have been found to respond only to complex combinations of stimuli, such as a bar in a particular orientation, stimuli with one or more right-angles in them, and like stimuli in two orientations 90° apart. Although much of this work was originally done on cats, rabbits, and other mammals, there is considerable evidence (Morrell, 1967) that the same type of response selectivity is found in the human visual system.

This physiological view of perception is congruent with the theories put forward by psychologists with regard to pattern recognition and like phenomena and emphasises the active, interpretive nature of perception, but other psychological research, particularly into the evolutionary and social significance of perception, into the effects of motivation, expectancy meaning, and culture on perception, serve to emphasise even more the belief that perception is not objective, but a response which is tuned to serve the needs of a particular organism at a particular time. The role of motivation was first emphasised by the New Look theorists, for instance, in an experiment which showed that the perception of the size of coins was influenced by the importance of money to the individual (Bruner and Goodman, 1947). Although there were difficulties with this particular experiment (Carter and Schooler, 1949; Pastore, 1949), the hypothesis that size perception is influenced by perceived value was replicated in further experiments (Bruner and Postman, 1948; Bruner and Rodrigues, 1953; Vroom, 1957).

The effect of motivation was examined more directly in an experiment by McClelland and Atkinson (1948); who asked three groups of sailors with different levels of hunger (defined in terms of hours of food deprivation) to report what they could see on a complex slide presented below the recognition threshold. The results showed that the mean number of responses associated with food increased with food deprivation. That the effect of the significance of the percept can act higher up the perceptual system than in size estimation is shown in the visual illusion connected with the Ames room. This well-known illusion involves a specially-constructed room which is lower on one side than the other; but which projects an image on the retina identical to that from a normal cuboidal room. The perceptual system persists in interpreting the Ames room as a normal room, to the extent that the perception of the size of adults standing inside the room is drastically altered. However, it has been reported (see Gregory, 1966; p.180) that wives viewing their husbands in this situation see them as normal, and the room as distorted. If perception involves extracting information or constructing a model of the physical stimulus, in accordance with the needs of the organism, one would hypothesise that expectancies would play an important role, and this was also demonstrated by the New Look theorists. Bruner and Postman (1949) presented normal playing cards and "monsters" (e.g. a red King of Clubs) tachistoscopically. They found that the monsters had a much higher recognition threshold than the normal cards (114 ms vs. 28 ms). The pre-recognition faults were of three types, all of which supported the view that perception is geared to a "best bet" in normal circumstances.

1. Dominance. The expected theme blocked out the abnormality.
2. Compromise. For example, the black six of hearts would be reported as purple or brown.
3. Disruption. The whole stimulus was faultily perceived.

Perhaps the most controversial phenomena originally demonstrated by the New Look theorists was the subception, or perceptual defence, phenomenon. Bruner and Postman (1947) found that the recognition threshold of words varied with their emotionality. In particular, the recognition threshold correlated with the association reaction time. They obtained an inverted U-curve, and postulated two processes of perceptual defence and perceptual sensitisation with extremely emotional words. A similar experiment by

McGuinnies (1949) found that taboo words had a significantly higher recognition threshold than neutral words. Both of these experiments were widely criticised, in that they did not control for word frequency or individual differences. In addition, response bias effects (virgin college girls being unwilling to repeat obscenities to austere college professors, unless absolutely sure) were not allowed for.

Experiments claiming to control for response bias have been performed in large numbers (see Minard, 1965) and with the result that some found perceptual defence phenomenon whereas others did not. As recently as 1967 Neisser argued that the phenomenon was an artefact, and suggested that experiments which claimed to control for response bias effects and also demonstrate the phenomenon were the result of experimenter effects. However, Neisser appeared to be unaware of at least two experiments which could not be criticised in this way and which clearly demonstrated the perceptual defence phenomenon. Dixon (1958) reasoned that response bias effects could be avoided if the taboo words were presented below the word-awareness threshold, rather than below the recognition threshold. He used a piece of apparatus in which the two spots of light were presented to one eye, and two words (one taboo and one neutral) were presented at the same light intensity, to the other eye. The subject had control of two joy-sticks, the movement of which caused the spot of light together with the word to get brighter or darker.

His instructions were to move the joy-sticks until one spot of light was just visible and the other was just invisible. In this way the awareness threshold of each of the words could be ascertained with absolutely no possibility of word recognition, and hence response bias effects. A pilot using experiment using this apparatus found that the awareness threshold was affected by the presence of taboo words. Heider (1961) used the same apparatus, but used neutral, and emotional but non-taboo words, as stimuli (e.g. Cancer, recent, stance, breast). The words were controlled for length and frequency etc. He also obtained ratings from the subjects as to how unpleasant these words were, and correlated the ratings with the awareness thresholds. He obtained significant results.

Some of the early arguments against the perceptual defence phenomenon suggest that it requires an internal homonculus to decide which words need to be repressed. Furthermore, it is claimed that such a capability on the part of the organism could have no evolutionary significance. These arguments seem to have less validity if one views perception as an adaptive response of the organism, selecting and suppressing stimuli according to perceptual context and organismic needs, The perceptual defence phenomena can then be seen as part of a sensitisation-defence continuum involving a low-level pre-setting of the perceptual system in favour of certain types of useful stimulus information and against other stimuli. Finally, if the argument that perception is partly an organismic response, varying with the demands of the environment and the needs of the organism, is valid, there should be clear differences in the perceptual processing of stimuli in individuals from different environments and different cultures. It is beyond the scope of the present monograph to review cross-cultural studies (particularly in view of the specific methodological difficulties to be considered in that area of research), but a recent review by Lloyd (1972) cites evidence to show that clear cultural differences are found in colour perception, the pictorial representation of space, and response to visual illusions. It is interesting to note that all these examples are at the cortical, rather than the retinal, level of the perceptual system, and are likely to involve social learning.

There is ample evidence that perception is an active, adaptive and constructive process. It is not exclusively definable in terms of the physical stimulus: the needs of the organism, and the functional significance of the stimulus in the context of the environmental situation must also be taken into account. However, a given person's motivations and needs will change from time to time and from situation to situation. Furthermore, different people will be perceiving in different physical situations, and will be under the direction of different motivational states. Perception cannot therefore consist of structural elements which are invariant across different contents, situations and individuals.

The research reviewed above has demonstrated differences in subject's percepts under different motivations and stimulus conditions, and these differences were demonstrated in relatively tightly controlled laboratory situations, where some large sources of variance are controlled. In the more complex forms of perception common in everyday life and in the field situation these variations across individuals and situations, are likely to be immeasurably increased. If it is possible to demonstrate variations in tachistoscopic recognition thresholds to printed words, how much more likely it is that there will be variations in the perception of other people in life situations? And when one goes beyond that to the realms of epistemic perception - the perception of ideas, attitudes, concepts, and arguments - such internally-referenced variations are likely to be the source of by far the larger proportion of cross-situation, cross-individual behaviour variance. The determining variables of perception in the field situation of employment, education and social life are likely to be idiosyncratic in a way which is not seen in the over-simplified environment of the laboratory experiment.

Laboratory experiments have shown that a number of variables effect perception: what they cannot show, or even investigate, is how these individual variables can integrate into patterns over different individuals at different times and in different situations, and yet result in a composite wholistic percept which is appropriate to that individual, time and situation. Theories of perception which are restricted to structural concepts and hypothetical mechanisms, assumed to operate irrespective of different individuals, stimuli and situations, cannot hope to account for complex perceptual processes of the field situation, in the same way that stereochemical theories of hydrogen and oxygen cannot account for the qualities of water. To understand complex perception, it is necessary to investigate it at its full level of complexity, in terms of its significance in the functioning of unique individuals in given situations.

Laboratory experiments, for the reasons mentioned, above, involve a distortion of our conception of the nature of perception, by implying the structural nature of explanatory concepts. This distortion is present even at the relatively low levels of perception capable of laboratory study, but becomes more serious and more visible as one moves from low-level perception to high-level perception, and from a controlled laboratory situation to a complex field situation. This is not to argue that structural explanation mechanisms are incorrect in any absolute sense. All explanatory concepts are relative to the purpose of explanation. Structural theories, concepts and methods of investigation are therefore appropriate to structural questions. They are, for instance, perfectly satisfactory if examining the difference between human and non-human perceptual systems, or the difference between human perception and machine perception. They are not, however, appropriate to questions involving the comparison of one human with another (individual differences) or questions involving the relationship of human perception to other human faculties or activities (e.g. learning, job-performance, social behaviour) and especially not in life situations where several of these 'faculties' and innumerable laboratory variables are simultaneously operative.

Proposition : That "structure" and "content" in perception are two aspects of the same thing.

The above pages have suggested that mechanisms and processes of perception have been assumed to be structural because of the limitations of the experimental method, which is the primary tool of investigation. Evidence has been cited which shows that perceptual processes are not always invariant across different individuals and different situations, and it was further argued that the importance of such inter-individual, inter-situation variations increases as one moves towards higher perceptual and cognitive functioning and towards functioning in the life situation. If the structure/content distinction is an inappropriate one in considering complex functioning, the logical consequence is to abolish the distinction in theory, and devise research methods which do not embody it as an assumption. However, it is perhaps necessary to give concrete examples of a non-structural theory of perception, and to attempt to show that such a theory is in fact a plausible substitute for the structural mechanisms which the writer has tried to criticise.

As mentioned above, an early non-structural theory of perception was propounded by the Gestalt school of psychology, which had its roots in rationalist theories of philosophy and in phenomenological psychology. A major premise of these theorists is that it is not possible to separate the mechanisms of perception from the stimuli which are perceived: perception is "pregnant" with the perceptual world, in the sense that perceptual mechanisms cannot be considered to exist without also considering what is perceived. This point is made by Merleau-Ponty, (1947),

"The unprejudiced study of perception by psychologists has finally revealed that the perceived world is not a sum of objects (in the sense in which the sciences use this word), that our relation to the world is not that of a thinker to an object of thought .... As a result we cannot apply the classical distinction of form and matter to perception, nor can we conceive the perceiving subject as a consciousness which "interprets" "deciphers" or "orders" a sensible matter according to an ideal law which it possesses. Mater is "pregnant" with its form, which is to say that in the final analysis every perception takes place within a certain horizon and ultimately in the world. We experience a perception and its horizon "in Action" rather than by "posing" them or explicitly knowing them (Merleau-Ponty, 1947)".

The present author interprets Merleau-Ponty's terms "form" and "matter" as, respectively, "structural mechanisms of perception" and "the perceptual world" or "the content of perception".

Thus in Gestalt psychology, the laws of "good figure" were an aspect of the stimulus"perceiving-organism interaction, rather than a transformation applied to a passive stimulus by an active perceptual process. The tendency of closure was as much a feature of the stimulus pattern as of the perceptual system, and was specific to certain types of stimuli. Gestalt psychology was largely ignored in Britain and the United States, possibly because of different philosophical premises, and possibly because of linguistic difficulties (Von Fiandt, 1966, points out that English is a matter-of-fact language, highly unsuited to describing the phenomenal world, and that Koffka's "Principles of Gestalt Psychology", was possibly never fully intelligible to Americans), but also because of over-optimistic phsiological underpinnings of the theory. It is unfortunate that, by generalisation, the philosophical point has been ignored as well.

Piaget (see Piaget, Albertini, and Róssi, 1944; Piaget, Vinh-Bang and Matalon, 1958; Piaget 1955) adopts a similar theoretical approach. Rather than viewing perception in isolation, he sees it as an active, adaptive process, analogous in many ways to, and subsidiary to, intelligence. For Piaget, the infant's perceptions are meaningless unless and until they are assimilated into the sensory-motor schemata which forms the basis of intellectual development. For instance, the developing perceptual constancies derive their main significance in the context of the contemporaneous development of intellectual operations, such as the object concept and the spatial groups. Thus the perceptual system is not a mechanism for extracting information from the environment, but a constructuve and operational process definable partly genotypically (in terms of the organism's "biological nature") but mainly phenotypically and in terms of its significance for the adaptive functioning of the individual. Whereas the English Empiricists started off with a consideration of the physical stimulus and the sensation, the Piagetian approach takes the maxim "in the beginning, there was the response". For perception samples only those aspects of the environment which are of relevance to responses (motor reponses and intellectual responses). The statement that "All Chinamen look alike" is an example of this.

Piaget's philosophical position has methodological consequences: it emphasises observation on individual organisms, not of isolated variables, but of variables seen as a function of their organic significance. It is perhaps surprising, therefore, that Piaget has paid so little attention to individual differences. His main concern has been to construct, by abstraction from observations of specific individuals in specific situations, the general schemata of perceptual and cognitive functioning. From the point of view of this thesis, it is important to note, that these schemata are not merely <u>assumed</u> to be general across situations and individuals: they are based largely on non-normative data, and occasionally on very small numbers of subjects (the main observations which form the basis of Piaget's theory of sensory-motor intelligence were made on Piaget's son), are postulated to generalise across individuals and situations, and are found to. Indeed, it is a tribute to Piaget's calibre as a theorist that they do.

The importance of differentiating perceptual contexts is a major feature of Werner and Wapner's Sensory-tonic theory of perception, and of the research programme to which it has led, (Werner and Wapner, 1949; 1952, 1956; Wapner, 1964). Werner and Wapner emphasise the need to consider perception in terms of the total state of the organism in addition to the characteristics of the stimulus. Wapner (1964) amplified the assumption of the research programme in a cube-diagram (see Diagram 4:2) which states the contexts in which perceptual phenomena should be observed. Thus they suggest four levels of cognitive operations, which function on three

Diagram 4 : 2 Schematic Representation of Research Programme, from Wapner (1964).

different types of object or event, under the influence of internal or external states of the organism. The fourth dimension of the diagram represents the developmental change of the organism, defined in terms of their "Orthogenetic" principle of developmentally increasing differentiation and hierarchic integration. The programme aims to place cognitive and perceptual phenomena in an organismic-developmental perspective by pursuing investigations in each cell of the cubes: in this way it hopes to take account of all possible interaction effect.

Wapner (1964) notes three principles which have shaped the programme's approach.

1. Behaviour must be considered in relation to the context of total organismic activity. With respect to perception they use the formula

$$P = sRo$$

i.e. perception is a function of the relationship between the proximal stimulus and the organismic state. Thus the two ends of the polarity of proximal stimulus, and state of the organism, can both be systematically varied to give effects on perception. Both "sensory" and "tonic" effects are equivalent to each other, in that both can produce the same end result in perception. This was demonstrated experimentally in that both tilting of the subject in his chair (sensory effects) and acoustic or electric stimulation on one side (tonic effects) and could change the perception of the vertical in a darkened room (Werner and Wapner, 1952).

2. Subsystems do not operate in isolation, but rather organismic subsystems interact. This leads to

experimentation on intermodal effects of different kinds of stimulation (Goldstein, 1955) and on figural adaptation (Werner and Wapner, 1955).

3. The organism is a system which exhibits directedness towards goals and furthers its goals by a multiplicity of means. Thus different perceptual mechanisms are postulated to be functionally equivalent in that they can lead to the same perceptual ends. For instance, Werner (1940) shows how object constancy may be achieved by perceptual, conceptual, and physiological means. Flavell (1963), points out that this concept is similar to Piaget's "partial isomorphisms".

Although this brief summary cannot do justice to the extensive research findings and theories of Wapner and Werner's programme, it is sufficient to indicate some of the theoretical pre-suppositions. Thus perception is not regarded as a structural mechanism, but is investigated in its different organismic and developmental contexts. The perceptual processing is not assumed to generalise across stimulus situations and organism states without prior investigation, although this has so far necessarily been limited to a few phenomena. However, the programme is still experimentally based, and this has two consequences:

1. It restricts the investigation of the effect of organism-states to what must be a rather minor level. While it is possible to change the organism state by electrical and emotional stimulation, and by the use of drugs (e.g. LSD-25), other more drastic changes of motivation, emotion, and circumstance cannot be experimentally manipulated for ethical or practical reasons. For instance, it is not

possible to manipulate experimentally such emotional effects as bereavement, job redundancy, withdrawal of motherly love, or traumatic experience.

2. It involves summation across individuals, and therefore, although it recognises the influence of a wide variety of organismic and environment variables, it implies that they operate on perception in a mechanistic and summational way. This implication is belied by observations: the percept is certainly geared to a given situation, and a given organism state, but is also a phenomenal whole, and a functional unit. This whole is based on a wide combination of variables, in fact such a wide variation that the particular patterning affecting one subject is unlikely to affect anyone else. A method which involves summation across individuals cannot investigate this unique wholistic percept, nor how it is built up from non-unique elements, nor how it comes to direct behaviour and action.

It now becomes possible to abstract from the theories of perception we have considered, some fairly generally-agreed charactistics - it is in the details that they differ. Thus many theorists agree that perception is active and constructive and hypothetical rather than camera-like (Piaget et al, 1944; Bruner, 1957; Neisser, 1967). These theories suggest that the perception of cues (Bruner) or stimulus elements (Piaget's "encounters" is an analogous concept, which however emphasises the active component of fixation) is used as the basis for a model or construct or perceptual hypothesis. The model is a "Best Bet" combination of the stimulus elements, taking into consideration expectancies, motivations, stimulus significance and other organismic-context variables.

Piaget has emphasised that perception must be largely response-based i.e. derives its significance from incorporation into action schemata. This view will be less familiar in Britain and the United States, where empiricist philosophy is more deeply entrenched in theoretical concepts. However, if one accepts the view of the developing organism as having to construct a model of an initially non-sensical environment, it is difficult to see how he would select information in the way he does unless they were selected for relevance to (or association to) motor responses. There is a virtually infinite number of models which could be fitted to the regularities of the perceptual world, and there must be a way of selecting among them to derive the one which humans use.

Also, the present chapter, following theorists such as Piaget, Bruner, Werner and Wapner, has tried to emphasise the necessity of seeing the perceptual process as inseparable from the actual stimuli perceived and from the organism-contexts of which they form a part.

If this analysis gives the impression that there is a different perceptual process for each individual and each situation, let us rectify that impression, now. The organism enters the world with a genotype which is to a large extent common across the species. All men have eyes and retinae and striate cortices. The genotype interacts with the environment which is also to a large extent common: we all live in the world. Also perception is geared to survival, which involves basically common response patterns (e.g. obtaining food and eating) to basically common needs (hunger, thirst, etc.) and finally, a relatively objective language provides the mechanism for a social regulation of behaviour and hence perception. From a cosmic point of view therefore, in comparing men with Martians or men with plants, structural theoretical entities, and hence experimental methods, provide an eminently satisfactory mode of investigation and explanation. With such a point of view, variations between individuals within the species, and variations between situations within the species' life-contexts, are of little importance. The present thesis has adopted as its criterion the human field situation, however. Against that criterion, individual variations and situational variations are of prime importance: it may turn out that perceptual mechanisms are invariant across different situations and individuals, although evidence has been presented that they are not always, and that evidence having emerged almost despite methodological assumptions to the contrary. With that evidence as a basis, it would seem advisable to pursue research which is actively looking for cross-situation, cross-individual variations: if such research is fruitless, then at least the structural aspects of theoretical explanations will be demonstrated, rather than methodologically pre-supposed.

4 : 5 Memory Processes

As was pointed out in Section 4:3:1, the theoretical concepts of memory are intended to be structural. Distinctions between iconic memory, echoic memory, short-term memory and long-term memory, primary and secondary memory, etc. are assumed to be species-general: they apply to all people in a wide range of situations. Memory research, even more than perceptual research, has been experimentally-tested and laboratory-based: and thus the structural assumption has been tested to a correspondingly lesser extent. However, the importance of the content of memory systems on remembering can be seen in the research into different types of coding. While coding is an active "structural" process applied by the cognitive system, it is inextricably linked to the content of memory: a given coding rule can only be applied to a particular range of cognitive contents. However, accepting that memory systems are partly definable in terms of coding and content, it is more difficult to show that they are also internally-referenced and idiosyncratic. Some types of coding (e.g. phonemic coding) are not at all idiosyncratic. Other types of coding, for instance semantic coding, aremuch more likely to be so. The present section will therefore examine the evidence relating to the importance of semantic coding in memory, and will support a "Process" interpretation of memory phenomena which is a more parsimonious account of the evidence. Finally, critical experiments and research implications are noted. This interpretation of memory is dossonant with the "structure-content" distinction, and is not investigable to any great extent by the experimental method. To the extent that this interpretation is acceptable, therefore, the "two propositions" will be demonstrated.

Proposition: That the Determining Variables of Memory are Internally-Referenced and Idiosyncratic

a) Iconic and Echoic Memory.

Neisser (1967) suggested that the initial phases of perceptions processing involved an extremely short-term memory system which he called an "iconic" memory. Strong evidence for this form of memory was presented in experiments by Sperling (1960), and Averback and Coriell (1961).

Sperling presented tachistoscopic arrays of letters for a duration of 50 ms., certainly too short a duration to allow eye movements. The array consisted of rows of letters, e.g.

TDR

SRN

FZR

Usually, the subject could read four or five of the letters, irrespective of how many were presented. However, when the experimenter arranged to signal which row the subject should read first (e.g. by presentation of high, intermediate or low tone), the subject could still read four or five letters, and the individual row was reported with 100 per cent accuracy. This was true even when the auditory signal was presented <u>after</u> the tachistoscopic stimulus. To produce this performance, the subject would have to focus attention on a particular part of a stored image.

By varying the time between the tachistoscopic presentation and the auditory signal, and noting the presence or absence of this effect, it was possible to measure the duration of the iconic memory. Sperling's data suggested that this duration was about one second.

Averbach and Coriell's experiment was similar, except that the subject only had to report a single letter, which was indicated by a black bar. Their results were similar, as pointed out in a joint paper by Averbach and Sperling (1961). However, their experiment appears to be more difficult to replicate, with a tendency to produce effects of visual masking and erasure (see Neisser, 1967; pp.22-35).

The coding used by iconic memory can be investigated by varying the cues which signal the part of the stimulus to be reported.

Wright (1965) showed that items can be selected on the basis of colour, and cues of location, (Averbach and Coriell, 1961) and shape (Turvey and Kravetz, 1970) can also be effective. Mewhort (1966) has suggested that the memory is responsive to linguistic habits, with better recall of letter sequences that resemble English (THEMLEAKE is more accurately recalled than AEMHTLKEIE), but Baddely and Patterson (1971) suggest that this is a function of read-out rather than storage. Furthermore, Turvey (1967) found no evidence of progressive learning in iconic memory.

Thus there does not appear to be any evidence that iconic memory is responsive to semantic coding, although there does not appear to have been a great effort to find it. The evidence on echoic memory, the auditory counterpart to the visual icon (Neisser, 1967) (Massaro, 1970), leads to the same conclusion. Furthermore, Schiller (1965) has shown that the erasure phenomenon, which is a function of the visual icon, is at least partly located at a retinal level. It is perhaps likely, then, that the icon is justifiably assumed to be a structural mechanism, applying across different people and different situations. An experiment to test this assumption will be suggested at the end of this section.

b) Short Term and Primary Memory.

Distinctions between Short-term memory (STM) and Long-term memory (LTM) have held wide currency in psychology in the last twenty years, and the relevant research is extensively reviewed in Murdock (1967, 1972), Norman (1969), Broadbent (1970), Kintsch (1970), Tulving and Madigan (1970), and Craik (1971). Craik (1971) notes that:

"The principal features of STM are its limited capacity, ease and rapidity with which items are retrieved, rapid forgetting, and apparently no permanent record......... studies by Conrad (1964) and Baddeley (1966) and others have suggested that verbal items are held in STM in terms of their sounds and not their meanings. In contrast, it is believed that LTM has a much larger capacity, shows slower forgetting (or even no forgetting) and stores words principally in terms of their semantic-associative features". However, evidence has mounted to show that STM/LTM are not as distinguished as was originally thought, and Kiss (1972) notes that "a structural schema based purely on the retention interval (short-term or long-term) has now become quite unsatisfactory........." Furthermore, Melton (1963) has long been opposed to the distinction, and presented evidence to suggest that retention of verbal items over a short period was affected by the same sort of variables as affected LTM. Waugh and Norman (1965) attempted to resolve these inconsistencies by proposing a primary memory/secondary memory distinction (PM/SM), PM holds only a limited number of items, items being displaced by later stimuli, and that items are then transferred to SM. Both PM and SM are therefore involved in most laboratory tasks, and the two components needed to be distinguished. The usual way to measure PM capacity was by using the later items from a free-recall memory task, and applying a correction formula, based on the assumption that items recalled from terminal list positions could plausibly come from either PM or SM. The middle list position could only come from SM, and thus gives an estimate of SM accuracy, which could then be applied to the terminal list position recall.

Craik (1971) lists an impressive array of evidence in support of the PM/SM distinction. Furthermore, research into variables which might affect PM capacity have shown that PM is surprisingly "structural" in nature; amongst variables which have no effect are the following:

Presentation rate (Murdock, 1962; Waugh and Norman, 1965; Glanzer and Cunitz (1966)).

Proactive inhibition (Murdock, 1972)

Word frequency (Raymond, 1969)

Concrete-abstract dimension (Paivio and Csapo, 1969)

Language of stimuli, presented to tri-linguals (Tulving and Colotla, 1970)

Repetition of stimuli (Glanzer and M inzer, 1967; Waugh and Norman, 1968)

Semantic variables (Kintsch and Buschke, 1969; Levy and Murdock 1968).

There is some slight evidence that PM may be non-structural. Thus the storage capacity of PM appears to be 3-4 units, but these units may be either letters or words. Murdock (1961) showed that 3 words followed the same forgetting function as 3 letters, and Craik (1968) showed that PM capacity for words was invariant with different numbers of syllables in the words. This use of higher or lower-order units suggests that PM is post-perceptual and can vary with the stimulus context to a certain extent. However, to argue from this to idiosyncratic variables is obviously unreasonable. However, Craik and Levy (1970) found that PM was detrimentally affected by semantic clusters of words. There are two possible interpretations for this:

1) PM is affected by semantic coding
2) Semantic clustering makes it more likely that words are recalled from SM, in which case they

cannot be retrieved from PM, and PM effects are consequently low.

Craik and Levy favour the second interpretation: the present author, with different a priori expectations, would prefer to leave the question open.

Finally, there is some evidence to suggest that PM is an artefact of laboratory tasks. Shallice and Warrington (1970) present clinical evidence which suggests that verbal material need not be held in a short-term store: when the input is highly compatible with the analysing systems (e.g. meaningful stimuli or pictures) analysis will begin at a deep semantic level. This suggestion is consonant with the interpretation to be presented below, but that interpretation also suggests that clinical evidence from brain-damaged subjects may be less relevant to normal functioning than is usually assumed. However, there is also supporting evidence from normal subjects. Sacks (1967) studied memory for sentences and found that subjects tend to remember the meaning of sentences, but not which words are used. Presumably, there is little or no remaining STM trace if the subject is not forced to attend to the early stages of processing.

If this suggestion that PM is laboratory-task specific is true, clearly there is a need for research into the limits of task specificity and what other memory characteristics vary with tasks and situations.

To recapitulate, there is only slender evidence that PM is affected by any semantic coding variables. To be fair to the present argument, however, it is necessary to point out that individual differences have not been investigated, and if semantic effects vary with different individuals, they may be cancelled out when summing scores across individuals. Here again, it may be that PM is justifiably assumed to generalise across individuals and situations. Research designed to elicit such differences needs to be done.

c) Long Term Memory and Secondary Memory

Long-term memory is probably far more important in everyday life than short-term memory, but is less studied in the psychology laboratory, and less well understood. In contrast with iconic and primary memories, there is ample evidence of the importance of coding. In a classic experiment, Carmichael, Hogan and Walter (1932) presented figures which were susceptible to two different sets of labels e.g.

BOTTLE  STIRRUP

When subjects were asked to reproduce the figure, there was clear evidence that their drawings moved towards the stereotype suggested by the label i.e.

BOTTLE  STIRRUP 

The most dominant feature of long term memory is that it is semantic and associative in its classification. This is demonstrated in a wide variety of experiments dating back to Bartlett's (1932) study. More recently, Paivio (1969) showed that the imagery-value of nouns predicted the ease with which they could be learned, and Bower (1971) showed that instructions to subjects to use imagery would improve learning.

Another important variable in LTM is the degree of organisation of the content or stimulus material (Mandler, 1967), although a distinction is made between the effect on storage and the effect on retrieval (Mandler, 1971). Tulving and Pearlstone

(1966) showed that recall could be improved by prompting subjects with the taxonomic categories from which items were taken: this has been interpreted as showing that more is stored than can be retrieved. Also, a list of words drawn from a few taxonomic categories will be better recalled than a list drawn from many. Bower, Clark, Lesgold and Winzenz (1969) demonstrated the effect of organisation even more dramatically: they presented lists of 112 words which were divided into four hierarchies. when the presentation of the lists emphasised the categorical organisation, 73 of the 112 words were recalled, in contrast to only 21 correctly recalled when the words were presented in random order.

Tulving has performed a series of experiments which emphasise the importance of subjective organisation. Thus Tulving (1962) noted the tendency of clusters of words to be recalled together irrespective of their initial list positions. Tulving (1966) required subjects to learn a list of 18 unrelated words for a free recall procedure. Half of the subjects had previously learned 9 of the words, whereas the other half had learned another 9 word list. Tulving found that the subjects who had learned 9 of the words were initially better in the 18 word list, but showed much poorer subsequent learning than the second group. Presumably the subjective organisation that they had imposed on the 9 word list would not extend to the 18 word list, and inhibited further learning.

The obvious importance of stimulus-linked codings implies a considerable variation of memory procedures across different situations and contexts. Also, semantic codings are similar to higher-level perceptions in that they must be based on a large number of more basic elements and yet retain a composite and "appropriate" wholeness. To a large extent, semantic coding will be linguistically based, but this does not mean that it is structural: Whorf's studies on the linguistic differences in diverse cultures and languages, and, even more, the number of arguments in the world which turn out to be "a matter of definitions" are both indications that the meaning of a word varies from one individual to another. An essential aspect of long-term memory will be missed if these cross individual variations are not investigated.

Proposition: That content and structure in memory are two aspects of the same thing.

Kiss (1972) notes that there is a trend away from "structural" memory theories of short-term and long-term memories towards functional memory systems attached to specific modalities, e.g. iconic, echoic, visual, primary, etc. *One is tempted to extend this trend, and predict that when the computer simulation approach comes to investigate the body concept, wine tasting, and other cross-modal phenomena, there will be a rush of olefactory-gustatory-short-term-memories, kinaesthetic-visual-auditory-long-term-memories, and the like, a situation which would be reminiscent ofthe late nineteenth century motivation theorists who proposed an innate instinct to play the piano. However, this is clearly a fallacious argument (perhaps an extensic ad absurdum?), and, despite giving a certain illicit satisfaction, will not be continued here.

Kiss also points out that these theories almost universally assume active processing mechanisms which operate on passive

---

*In the present use of the word "structural", these memories are still structural, but are so to a lesser extent: the range of generalisation of the theoretical criterion has been reduced to stimuli from a specific modality.

information, a necessary pre-supposition when theorisation is based on the conceptual background of the Von Neumann computer organisation. From the basis of automata theory, Kiss suggests an alternative conception:

"The conception of memory suggested here is active, in the sense that the memory system itself has certain processing capabilities. It can undergo autonomous changes without the influence of a central processor and it produces outputs in response to interrogation, rather like an information-retrieval system, instead of merely providing a set of pigeon holes into which information can be placed and later read out with complete certainty. We look upon memory as a sub-system within the over-all information processing system of the organism". (Kiss, 1972, p.336).

The memory system is thus conceived to be an automation and is describable in the language and mathematics of automaton theory. Amongst other things, Kiss suggests that it has the following features:

1) The memory system can be in a number of states. The state can change autonomously when there is no input, or in relation to in input.

2) "The 'contents' of the memory system are the different states". Information is stored in the system if it has a state in which an appropriate output is produced.

3) The retrieval mechanism of the system is the 'transition' function of an automaton i.e. the rules which determine how one state changes to another state. "Given a set of inputs, an item is retrieved from memory if the set of inputs drives the system.......into the state appropriate to that item.

This extremely interesting theory clearly implies the importance of the content of memory, and in fact cuts across the usual content/structure distinction. The usual 'contents' (words, ideas concepts, images, etc.) are a function of the structure of the system, and partly determine the processing of the system. It does, of course, erect new structural concepts - automaton, transition function, state of the system, etc. - but these are on a completely different level of generality compared with the usual theoretical concepts. And, of course, it is not possible to comment on them without further elaboration and empirical research.

However, the important suggestion from the present point of view is that the conventional contents of memory are seen to be a function of the structure and processing of cognition. Contents manifestly differ across situations and individuals and are likely to be idiosyncratic. Clearly the investigation, in any depth, of this type of theory cannot therefore be undertaken using a normative experimental method.

The application of automata theory to memory is of course new, but this is not the first suggestion that memory involves active contents. Bartlett's (1932) theory of memory relied heavily on the concept of schema, which in turn originated in Head's (1920) investigations of the mechanisms of postural control. The schema concept can perhaps be likened to an analogue, or continuously-changing automaton, and is best described in Bartlett's own words:

"'schema' refers to an active organisation of past reactions, or of past experiences, which must always be supposed to be operating in any well-adapted organic response. That is, whenever there is any order or regularity of behaviour, a particular response is possible only because it is related to responses which are similar, and which have been serially organised, yet which operate not simply as individual members coming one after another but as a unitary mass". (Bartlett, 1932, p.201).

Here again the contents of the memory system cannot be separated from the structure: a stored memory derives its effect, and hence its meaning and theoretical definition, from inclusion in the schema, which acts as a whole. As a whole, composed of a vast array of associated elements, it must of course by idiosyncratic and unique: it is not conceivable that any two people would have the same pattern of stored elements, nor the same schema, and consequently it is not possible that any two people would approach a new stimulus, task or situation in exactly the same way. It is interesting to note that while Bartlett used experiments, it was not the experimental method as used to-day. It involved stimuli which were as close to life as possible, and did not sum responses across the different subjects. Each subject was investigated in his own right, and only then were generalisations made.

Piaget is another theorist who makes use of the concept of 'schema' and in fact takes a similar view of memory to that adopted here. Although Piaget's major research is concerned with a different aspect of cognition, one can glean from his writings, comments, relating to his view of memory:

".....various kinds of behaviour - so different from one another that the only common feature is the conservation of the past...... - are generally designated 'memory'. In most cases memory is thus fused with habit or with that special aspect of habit which involves the recognition of signs. As for evocative memory, ...it probably only represents the figurative aspect (recollection images) of the conservation of intelligence schemata". (Piaget,1971,p.3). "I do not consider 'memory' and 'logic' to be two distinct departments: the essential data are the perceptual sensorimotor or operational schemata: the conservation of these action schemata is what constitutes memory, and their organisation is what constitutes logic...." (Piaget,1971,p.13). Here again, there is no active processing mechanism manipulating a passive information store: there is no need to postulate one. Any system, biological or mechanical, which changes in a non-random way, exhibits memory, or representation of past events. If a system is changed by an input, then the new state naturally involves a memory of the input. In terms of the cognitive systems, the input will only produce a partial change, indeed a miniscule change in relation to the entire system. The new state constitutes a memory of the input, and re-activation of that state constitutes recall. This cognitive system is pregnant with information, and is a process view of memory rather than a "structural" view, and suggests that there are as many "memory stores" as the are different processes in the individual. Memory is an aspect of the process as it changes over time, but no more than an aspect: it cannot be separated from the process and cannot be investigated except in relation to the process. Section 4 : 4 suggested that perception and cognition are motor-oriented i.e. they can be defined only in terms of the motor responses and response schemata which the organism has developed in order to satisfy its basic needs. In the sam way, memory is an aspect of these response schemata. In concrete terms, this implies

that memory is context specific: to the extent that individuals operate in different life contexts and use different response schemata, there are different memory stores.
In this conception of memory, primary memory is probably an artefact of the psychology laboratory, where subjects are forced to attend to the lower levels of processing. In complex life situations, stimuli are almost always meaningful - indeed the inexplicable often provokes fear or astonishment - and processing is on the level of semantics. Hence the results of Sachs (1967), that subjects remember sentence meanings but not the actual words. Hence also the results of Bransford and Franks (1971), who presented simple sentences which combined to form complex ideas: subjects "recognised" complex sentences embodying the ideas, even though they had never been shown the complex sentence before, and indeed preferred them to the shorter sentences which had been presented. This study emphasises not only the importance of semantic coding, but indicates how ideas can come together to form a temporary unit directing recall.
Baddeley (1970) showed that SM relies heavily on semantic coding, and Bower (1970) showed that it can use phonemic information where that will aid performance, for instance with rhyming words. Phonemic coding may also be used where semantic coding is difficult (as with nonsense syllables - Gruneberg, Colwill, Winfrow and Woods, 1970) or too demanding on processing capacity (as with a demanding supplementary task - Eagle and Ortof, 1967). All these results suggest compatibility with a process view of memory; which puts forward the notion that different tasks will involve different schemata, with their corresponding manifestations of 'memory'. The theories of memory reviewed above maintain the distinction

between storage and retrieval of information. Indeed, Baddeley and Patterson (1971) state:

> "Because of its enormous storage capacity, one of the major problems of LTM lies in retrieving what has been stored". (p.240)

Process memory, in contrast, does not distinguish between these. Storage and retrieval are both a function of the change in response schemata. For instance, in the laboratory experiments which show the 'effect' of organisation of stimulus content on retrieval, process memory would suggest an entirely different interpretation. In these laboratory tasks, the subject is being asked to recall a number of elements which are not usually associated in the same response schemata. The subject's difficulty is the retrieval of the individual stimulus items. Thus words from few taxonomic categories are words from few normally-unassociated cognitive schemata, and the task schema is more easily constructed. Bower et al's (1969) 112-word hierarchy is the provision of ready-made task schema. Perhaps the classic example of the construction of a task schema is the use of mnemonics, which can be translated into process-memory jargon as "the translation of a schema-unrelated collection of stimuli into a familiar task schema". Thus the subject who mentally moves around the house and takes out the stimuli from different places is connecting the usually unconnected stimuli into familiar schema.

Perhaps one of the best demonstrations of this interpretation was given unwittingly by Mandler (1967). He required a group of subjects to sort unrelated words into subjectively determined categories. After reaching a criterion of sorting the words consistently, he asked them to recall them. He found that subjects who were only instructed to sort the words, recalled them as accurately as subjects who were instructed to memorise them.

The writer earlier made some facetious remarks about cross-modality memory systems. These remarks were only facetious as far as the structural memory stores are concerned, however. The process memory interpretation suggests that these memories are likely to be the rule in complex functioning of everyday life. It would suggest that driving a car, for instance, has its own visual-auditory-motor "memory" at specific levels of processing - corresponding to the sensori-motor and conceptual schemata which control driving.

Furthermore, this "memory system" is different from the one used by the motorcyclist, where the critical stimuli are different. For instance, if asked to recall the features of their morning drive to work, not many car drivers, but most motorcyclists will remember particular places where oil collects on the road (e.g. where cars park). Process memory suggests that this is not merely a question of differential importance, but a question of differential cognitive schemata which imply a differential perception and recall of the environment.

Finally, evidence based on brain-damaged patients was quoted above to support the present argument. This may be ever more suspect in its relevance than is usually assumed, however. Luria's (1966) impressive investigations of higher cortical functioning emphasise the importance of interactions between different parts of the cortex in producing any cognitive phenomenon. If this is true, then a lesion in one part of the cortex will involve a temporary disturbance of functioning, but will allow functioning to recover, to a certain extent, as new patterns of interaction and new pathways are set up. It was this phenomenon which underlay Lashley's early theory of "mass action". However, if different cortical areas and different interactions are being used, it is plausible to assume that the patient is fulfilling the same cognitive goal by means of different cognitive processes. Memory, from the present viewpoint, is merely one aspect of those processes: it is possible therefore that brain-damaged patients used different memory systems from normal patients, and from each other.

It should be pointed out that the view of memory, put forward by Bartlett and Piaget, and elaborated here, is not "a theory" of memory in the same sense that Waugh and Norman's (1965) PM/SM distinction is a theory of memory. It is not at the same level of specificity, and cannot predict many of the phenomena predicted by the other theories. It is instead a "meta-theory", suggesting the types of phenomena which are relevant, the methods which might be used to investigate them, and the form which theories of memory should take. It is not a theory in that it is difficult to think of an observation which could disprove it: in Popper's (1959) terms, it lacks any immediate "falsifiability". This does not mean that it is beyond empirical data. In fact, it is on the same level as the "implicit" meta-theory presupposed by the experimental method: while it cannot be falsified by any single observation, it can be found to be unsuccessful in producing predictive theories, and it can give a more or less

parsimonious interpretation of research findings.
This being so, it is not possible to suggest any critical experiment or observation which would disprove the process view of memory. However, the process memory interpretation suggests that memory cannot be separated from process, context, and individual variations. This is an absolute proposition, but the extent of its practical application can be investigated. The most simple way would be to use two groups of subjects who had widely different life contexts and hence different cognitive schemata. It would then be necessary to devise stimuli which were associated in the life contexts of one group, but not in the other. Recall of these stimuli could then be tested at the different 'levels' of memory processes - iconic, ST, LT, etc. For instance, the following experiment would investigate the possibility of individual differences in iconic memory.

Subjects Stimuli

	Subjects	Stimuli	Prediction
Group 1	Engineers	Engineering jargon	Long icon
Group 2	Artists	Engineering jargon	Short icon

Alternatively, the same point might be demonstrated by:

		Condition I	Condition II
Subjects:	Engineers	Engineering jargon	Non-engineering jargon

Prediction: Standard deviation of the length of icon is greater in condition II

The author thinks it unlikely that individual differences in iconic memory would be found, at least with respect to semantic coding. It is not difficult to devise similar experiments to investigate the proposition at different levels of memory functioning.

The major research implication, if this view of memory is accepted, refers to the mode of investigation. Semantic codes manifestly differ across different individuals, and individuals vary their cognitive processes and memory processes across different life contexts. The experimental method may be used to test this supposition, as suggested above, but cannot be used to investigate the phenomenon in detail. A research method capable of doing so is therefore of prime importance, and it is to be hoped that such a method would speedily lead to an account of the development, organisation, and function of life contexts together with their cognitive schemata.

4: 6 Higher Cognitive Functioning

Without making a formal definition, higher cognitive functioning is taken to refer to the functioning associated with attitudes, belief, logic, thinking, decision making, problem-solving, planning, complex skills (conceptual and motor) and knowledge. More than anything, it is what distinguishes Homo Sapiens from other animals. It is also the type of functioning most distinctive of everyday life and least studied in the laboratory.

Proposition: That the determining variables of higher cognitive functioning are internally-referenced and idiosyncratic

It is almost superfluous to 'demonstrate' this proposition with regard to higher cognitive functioning: the evidence is all around one in art, literature and common sense. It is common observation that a person's knowledge, beliefs, attitudes and interests influence his perception and categorisation of the environment, and his responses to it. The University student's subject of study comes to affect his perceptions of the world: the engineer will see structures which are hidden to the artist, the artist will see light effects which are hidden to the musician, the musician will hear modulations and phrases in a bird's song, and the psychologist will see "defence mechanisms" where another will see "fact". Laurence Sterne (1760) gives an account of a related phenomenon, when he portrays the effects of Uncle Toby's Hobby-horse. Uncle Toby's Hobby-horse was his love of fortifications: not only did he follow all the accounts of campaigns, battles, and sieges, and re-enact them in his back-garden, he came to interpret everything in life in terms of fortifications.

It is common observation that the course of life and even sometimes its continuance is affected by knowledge. The author would not have spent the last three years researching for and writing the present thesis, had he not had, and been able to, demonstrate a certain knowledge. Qualifications, occupations, and hence even place of residence, and friendship patterns are influenced by knowledge. In extreme situations, knowledge can become the difference between life and death, as with the would-be sailor, who goes to sea in a canal barge. There can be no doubt then that the knowledge, beliefs, attitudes and skills possessed by an individual have a profound and diverse effect on all aspects of his life. The question is, to what extent can these cognitive elements be described in structural and externally-referenced terms? This question leads us back to the nature of knowledge itself. The nature of knowledge has been a central topic of philosophical enquiry since man began to reflect on himself. It is a topic on which volumes have been written, and in which the central problems are still far from resolution. It is clearly beyond the scope of the present thesis to follow, or even summarise the debate. However, neither is it possible for the writer to refrain from getting his feet wet to a certain extent: if, as is being argued, the knowledge of individuals is a major determining variable of their behaviour, it is necessary to make a psychological investigation of knowledge, and hence to adopt a position with regard to the philosophical pre-suppositions. The position adopted here, and set out below, is based on the ideas of Piaget (1970), (1971). Piaget intends his position as a philosophical argument, and it has been criticised as such by, amongst others, Hamlyn. It is adopted here, however, primarily because of its congruence with other analytic concepts of psychology, particularly the "process" interpretation of memory and perception described above. One of Piaget's arguments, with which the writer would concur, is that psychological investigations are beginning to cast light on epistemological questions which have previously been discussed only by the philosopher. Piaget suggests that knowledge is ultimately derived from either or both of two sources:

1) Innate knowledge - biologically programmed. Reflexes and sensations come into this category.

2) Environmental knowledge.

Innate knowledge is likely to be general across a species to a certain extent, but will also differ with individuals. All normal members of the human species perceive light and sound and can manipulate the environment. Thus all knowledge is ultimately based on light, sound and physical displacement.

Also all normal members of the human species have the capacity to make sounds and develop a language (in contrast to a dog, which does not have the biological capacity to develop a language): this potential is a necessity for "secondary knowledge".

Much knowledge is obviously a knowledge of the environment. However, this knowledge can only be developed as a function of certain ways of perceiving the environment. To the extent that these ways are common across the species, this knowledge is objective and externally-referenced. But Section 4:4 argued that perception is not externally-referenced. It is subject to individual differences in motivation and expectancy. It is also a constructive process that incorporates a mass of stimulus variables and differences into a composite and appropriate whole. In the final analysis it is internally-referenced and idiosyncratic. Knowledge, also, in the final analysis, must be internally-referenced and idiosyncratic.

There are of course important control and regulation systems which minimise the idiosyncracies of knowledge. All knowledge is ultimately aimed at adaptation to the environment and at satisfaction of the organisms' needs. To the extent that all humans are in the same environment, and receive the same feedbacks from the responses open to them, knowledge is common and objective. There can be few people who do not "know" that brick walls are not put there for the banging of the head.

A second important regulatory system is, of course, the social system. Parents, peers, and teachers, exert the most immediate feedback: education can be interpreted as an organised attempt to communicate the objective knowledge of a culture. Tendencies towards idiosyncracy of concepts and beliefs can therefore be corrected to the extent that they can be verbalised and discussed.

It is not therefore unreasonable to attempt to conceptualise perception and knowledge in a structural way. But such a conceptualisation will miss out two important aspects of knowledge:

1) The non-socially regulated knowledge. Some forms of knowledge are taboo, and not talked about. Sex in the abstract used to be such an area, but is perhaps now talked about a good deal more than most subjects. Sexual knowledge by an individual of particular people is still taboo, and one would speculate that such knowledge would manifest highly idiosyncratic categories. Other forms of social behaviour are also usually left unmentioned, and "social knowledge" of people is not likely to be under social control. There are also areas of knowledge which social forces cannot "get at": the phenomenal percept is one of these and conscious thought is another. The author has found many laymen are surprised when he says: he is studying "styles of thinking". Most people do not seem to have considered the possibility that other people think in a different way from themselves. From the point of view of the layman, therefore, the author's categorisations and knowledge of conscious thinking are idiosyncratic.

2) High level knowledge, which is a construct based on lower level "objective" elements. Although it is difficult to demonstrate this point, it seems likely that lower order beliefs will be subsumed within higher order categories which act as a whole or unit in cognitive functioning. One might

hypothesise that attitudes are examples of such units. While they cannot be verbalised in their entirety, they affect perception and thinking in a characteristic way, produce a 'bias' or 'emotional flavour' to certain percepts, and prejudice thinking towards certain conclusions. Such "wholes", because they are the resultant of an array of parts, cannot be structural. They cannot be defined exclusively in terms of the parts which go into their composition, but must also refer to the cognitive and perceptual phenomena which they control in the individual.

A review of the laboratory evidence to demonstrate the idiosyncracy of higher cognitive functioning would fill a book, and is surely unnecessary in the present context. It will suffice to mention a number of phenomena each of which emphasises the directive and processing effect of cognitive content:

Problem solving:
- -functional fixedness (Maier, 1930)
- -insight (Maier, 1931)
- -perceptual and emotional blocks (Osborn, 1953)

Logic:
- -effect of emotional content on syllogistic reasoning (Lefford, 1946)
- -implicit presuppositions in syllogistic reasoning (Henle, 1962)
- -inductive evidence (Gilson and Abelson, 1965)

Reasoning:
- -importance of direction (Maier, 1930: Bartlett, 1957)
- -content in reasoning (Wason and Johnson Laird, 1972)
- -plans and strategies (Miller, Gallanter and Pribram, 1960)

Vocational Choice:
- -as a function of self concept (Super 1957; Hunt, 1967)
- -as a function of interest (Strong, 1943; 1955)

Persuasion:
- -resistance to (Hovland, Janis and Kelley, 1953)

Cognitive Dissonance
- -individual differences (Miller and Rokeach, 1968).

Proposition: That structure and content in high cognitive functioning are two aspects of the same thing.

Piaget (1971) argues that autoregulation is an essential characteristic of all forms of life. All biological organisms are involved in an endless fight to build up and maintain an organisation, which is in some respects invariant and stable, in the face of external and internal perturbations. This is as true in talking about plants, as in talking about the higher mammals. The perturbations are environmental changes and breakdowns or random disrupting influences within the organism, and the mechanisms of autoregulation are the responses open to the organisms. Autoregulatory systems are visible in the plant suffering from desiccation (perturbation), which alters its water intake (response) right up to the spaceman who is faced with an inoperative heat-shield (perturbation), and who improvises (response).

The important aspect of this argument is that these systems form a continum throughout nature. Piaget argues that cognitive processes, and particularly intelligence, are not different in principle from the homeostatic regulations of primitive creatures. They are, in fact, a highly differentiated form of autoregulating system. From the primitive responses of plants, he moves on to the more complex response systems of mammals, and to the sensori-motor responses of the human infant. His account of the development of sensori-motor intelligence, and its consequent elaboration through pre-operational schemata, concrete operational schemata to formal operational schemata, has of course been the major part of his life's work, and is well known. All these examples are exemplars of autoregulatory structures. *

*"structures", henceforward in this chapter, are used in their Piagetian sense, and not in the sense of "invariant across individuals and situations", as the word has been used up to the present point.

Piaget's earlier work (Piaget and Lambercier, 1942-58), emphasises that perception also is an aspect of these schemata. Piaget sees perception not as a passive, receptive process, but as an active, constructive process which is just as much a response to the environment as is a motor act. Theories of perception in this country and the United States are at present moving towards that sort of interpretation (e.g. Neisser, 1967). Thus the perceptual construct is not "abstracted" from the proximal stimulus, but imposed upon it, and is a very real tool in categorisation of the environment and the delineation of response systems. From this point of view, knowledge schemata are another form of cognitive schemata slightly less general than the formal operations of intelligence. The knowledge schemata are active categorisations of the perceptual environment, which posit functional invariances i.e. invariances in terms of response-payoffs. They can thus be interpreted as the limits of application of sensory motor and intelligence schemata. To take the concept of "food" as a particular knowledge schema, it is an active categorisation of numerous environmental stimulus-elements (plums, apples, eggs, cheese, etc.) which posits a response-pay-off invariance (they are all good to eat), and can therefore be thought of as the limits of application of action schemata associated with eating (chewing, conveying to the mouth, salivating, etc.). But they are not peripheral to action schemata: they are an integral part of the total response to the environment. "The primary meaning of the hypothesis is that knowledge is not a copy of the environment but a system of real inter-actions, reflecting the autoregulatory organisation of life just as much as the things themselves do" (Piaget, 1971,p.27). Therefore there is no need to postulate active processing mechanisms which operate on passive information: the information itself is an active response to the environment. Something is only "known" to the extent that it is incorporated, or assimilated (to use Piaget's terminology) into action schemata.

In concrete terms, this suggests that the concept of "table" is not defined in terms of some template of "tableness", or even purely in terms of a feature analysis of the perceptual image of tables (legs, flat surface, hard, etc.). This suggests that an integral part of the concept of table, or at least of how it develops, is the assimilation of the visual stimulus into motor schemata i.e. that you can put things on a table, eat from it, sit up to it, etc. Admittedly, when the label of "table" is added, the concept gains a cognitive independence which allows it to be separated from the immediate sensori-motor context: a table can be recognised when floating in a harbour, so that one cannot "sit up to it", and the verbal concept can be used when the physical stimulus is not present. But that should not be allowed to obscure the origins of the concept, or its nature as an active operational entity, as an active structuring of the environment which is instrumental in the direction of response systems.

Cognitive schemata are thus structures, in the sense of the word as used by the structuralist philosophers. They have the following characteristics:

1) They are transformational. They are systems of operations, in which "processing" is implicit.

2) They are autoregulatory, i.e. they incorporate feedbacks

which are aimed at self-maintenance of the structure.

3) They exhibit "wholeness". Although structures are composed of elements, the elements are associated by rules. It is this system of rules that is the structure. Because the system is self-maintaining, it can act as a "whole" or a "functional unit". Thus the concept of "table" is initially a structure of sensori-motor operations, but is also a functional unit in cognition.

That both knowledge and processing are two aspects of cognitive structures or schemata is the central point of this chapter, and this has direct consequences for the types of research method used by psychology, and on the suitability of the experimental method, in particular. The nature of structures as "wholes", which contain but relate vast numbers of subordinate elements, reinforces the previous argument as to the importance of idiosyncracy. Structures, as autoregulatory systems, develop <u>within</u> the individual, and reflect both the individual organism and the individual environment. To the extent that individual environments vary, and to the extent that the developing structures are not normalised by social and linguistic control structures, to this extent cognition is idiosyncratic.

4 : 7 Cognitive Style and Cognitive functioning as Empirical Epistemology

The argument seems to have moved a considerable distance away from "The nature of cognitive style", but the distance is perhaps more apparent than real. It may be useful to have an overview of the hypotheses that have been presented. Firstly, cognitive style was originally taken to mean "individual differences in cognitive functioning", a definition which suggests that conceptualisations of individual differences must be solidly based on theories of cognitive functioning. Some of that analytic concepts common in such theories were critically examined, particularly in terms of their relationship to the prime research method of psychology, the normative experiment.

It was suggested that the experimental method involved a number of assumptions, which might be incompatible with the nature of cognitive functioning. For instance, it assumes that the determining variables of behaviour are in fact internally-referenced. Thus a given experimental situation might mean two things to two people; also, while it may be possible to manipulate these variables to a certain extent, it is both ethically and practically impossible to simulate variations which occur in everyday life. Finally, "systematic" manipulation of variables is in any case impossible until one can give an account of the internal system of which they form a part.

The external referencing of determining variables also means that experimental results are specific to the experimental situation, whereas psychological theories are often assumed to generalise. Evidence was presented to suggest that these are often false generalisations. However, this is not to say that determining variable s do not generalise, (indeed structures of knowledge and structures of intelligence show very wide generalisation). It <u>is</u> to say that generalised theoretical entities should be <u>observed</u> to do so, and not <u>assumed</u> to do so. The use of control by randomisation makes the method normative. Summing across subjects implies that determining variables in any situation are the same for all subjects. Evidence was presented to show that individuals differ in response to determining variables, and, indeed, in the nature of the determining variables which influence them. It was argued that the role of idiosyncracy in cognition is often underrated.

Finally, a point which was not made before, but can now be appreciated, the experimental method involves the isolation of a particular variable, together with the control or randomisation of other variables. It is a systematic control within a "frozen" situation. However, if the suggestion of the schema, or structure, as the basic unit of cognition, is accepted, there is a contradiction. The structure, although a "whole" in one sense, in another sense is a system of interacting parts. The interaction is primary to its nature, not a secondary consideration. A research method which involves the isolation of one part from its functional relationship is not a good way of investigating structures. These presuppositions of the experimental method are embodied in the structure/content distinction, which is implicit in many theories of cognitive functioning, it was suggested. Arguments and evidence were presented to suggest that this distinction is theoretically false, and in practice gives a distorted view of cognition, the more distorted as one moves from the sensory and motor aspects towards higher cognitive functioning, and from a laboratory criterion to predictions in the field situation of everyday life. In contrast to the structure/context distinction, a view of cognitive functioning was presented which leans very heavily on the ideas of Piaget, and which is also implicit, to a greater or lesser extent in the writings of other theorists (e.g. Bartlett, 1932;

Hebb, 1949; Miller, Gallanter and Pribram, 1960). This view, which emphasises the active, interactive, auto-regulatory nature of cognition, is dissonant with the assumptions of the experimental method. This, in turn, leads back to cognitive style, suggesting that a normative, dimensional analysis, which embodies the same pre-suppositions, is likely to be inadequate and possibly misleading.

This chapter has also suggested that the importance of what is normally called the content of cognition, has been under-estimated. Considering its intuitive prominence, there has been a surprisingly small research effort. Even Piaget, who has come to the author's intellectual rescue many times, is of little help here. Piaget's researches have been into the structures of intelligence, and havé largely ignored cognitive content. It is true that he shows how intelligent operations originate in the sensori-motor schemata of the neonate, and these sensori-motor schemata are strongly tied to their behaviour contexts. Even at the level of the concrete operational structures, there is still a link with context: the operations can only be performed on concrete subjects, although they do generalise across different types of situations. However, Piaget's main effort is put into the development of the structures with the end-point of formal operational structures in mind. The formal operation structures are characterised by being general to a maximum degree, attaining the fully cognitive status where they can operate in the abstract.

The structures of intelligence are not the only cognitive structures which need investigating. The structures of knowledge are also important, and possibly more important with respect to normal everyday functioning. The structures of knowledge can be thought of as along a different dimension from the structures of intelligence. The structures of intelligence are internalised motor operations, and are to a large extent general across different types of knowledge, (although not completely so; see, for example, Wason,1969). The structures of knowledge are internalised perceptual operations, involving an adaptive categorisation of the stimulus world. An investigation of the structures of knowledge may very well be the key to questions of how people interpret the complex situations of everyday life, and how individuals respond to these situations, whether the response involves the mending of a car or the choice of a career. It seems probable that a psychology of the human being in his natural environment will require an account of the organisation of knowledge schemata, in the individual, and across individuals, and how these schemata develop from the behaviour contexts of early life. It may be objected that this is relegating psychology to a branch of philosophy. The author would prefer to turn this round and say that what was a branch of philosophy should now be open to psychological investigation. There does not appear to be any insurmountable barrier to the application of the empirical and scientific methods which can give an account of the nature and organisation of knowledge, in the individual and in the social system. There is no barrier, provided the assumption is <u>not</u> made that the normative experiment is the only tool available to scientific psychology. Chapter 6 will suggest an observation method which involves presuppositions which are compatible with this view of cognitive functioning, and chapter 8 will present the results of a pilot study using it.

The cognitive structures of knowledge, attitudes and social functioning have so far scarcely been investigated. Needless to say, the present study can scarcely hope to scratch the surface of this area. Nevertheless, one can hope that a serious research effort would shortly bring the possibility of understanding behaviour in the real-life contexts of education, the family, friendships, and work.

CHAPTER 5

PRESUPPOSITIONS OF THE NATURE OF EDUCATIONAL SITUATION

5 : 1 Preliminary Considerations

As with research in cognition, any research in education must be based on a pre-supposed model, in this case of the nature of the educational situation. In this model, the presuppositions of educational research will be, to a large extent, congruent with the pre-suppositions of cognition, and one can therefore expect many of the points made in the last chapter to apply here. It will also involve pre-suppositions which are specific to the research methods, and environment, of education. Parts of the basic model will be explicit and parts will be implicit. Implicit pre-suppositions are particularly difficult to explicate. Many of the explicit presuppositions which are listed below have been proposed by other researchers as criticism of research methodology. The justification for dealing with them here is that, even though most researchers are aware of them, they remain implications of the research: a research method which does not involve these presuppositions has not yet been developed. Furthermore, it is very rare to find a general survey of the basic model of the educational situation: individual criticisms and points are usually presented in the context of individual papers. This makes it difficult to obtain a coherent view of our research methodology, and its goodness-of-fit to the modern concept of the educational situation.

The present chapter will attempt to set forth the explicit presuppositions, and some of the implicit presuppositions, of that sub-category of educational research which attempts to predict academic performance in tertiary education, and the research methods used. It is thus necessary to deal with criterion variables, predictor variables, and the logical and statistical procedures which allow comparison of the two.

This sub-category of research is still vast in extent, and the author has been forced to select from it. Differences between countries in the administration of education results in educational research projects often being less comparable than other types of research. The present chapter is therefore based primarily on research in British universities, and includes studies originating in other countries only where (a) there is insufficient British research on the point of discussion, (b) a theoretical point or argument is raised which is not apparent in the comparable British research.

When one wishes to examine the presuppositions behind research, and the methods used in research, it is necessary to bear in mind the purpose of the research. The use to which research information will be put inevitably determines how research questions are formulated, and which observations, of the infinite possible range of observations, are actually recorded. Although the purposesof a particular piece of research are rarely mentioned, at least in the research reports, it is possible to infer that the ability to predict academic performance might prove valuable for the following reasons:

1) It would enable the maximisation of the effectiveness of financial resources. Where financial resources are scarce relative to the demand for University places, there is a duty to maximise the effectiveness of those resources by selecting

students who are able and willing to take advantage of them.

2) It would help to minimise human wastage and suffering. It is often argued that a failing student has not only resulted in financial waste, but has also suffered a blow to his self-esteem and confidence. This blow may be both undesirable and unjust.

3) It would elucidate the psychological components of the academic skill, thus leading the way to:

(i) the tailoring of the educational situation to the needs of individual students

(ii) the counselling of students who are not fulfilling their potential

(iii) the design of assessment procedures which are more valid measures of the attainment of educational objectives.

It should be noted that the criterion for 1) and 2) is actuarial: if research findings allow the reduction of wastage rates from 30% to 20% of student entry, then the aims of the research have been partly fulfilled. The criterion for 3) is more strict and more important, although probably more remote: it requires the application of knowledge to the individual student.

5 : 2 Criterion Variables

The general criterion of "academic performance" has been measured in a number of different ways. The most obvious measure is examination results, based on 1st or 2nd year examinations, or Finals. The over-all grades or degree divisions can provide an adequate differentiation between students, and examination marks have been used where greater precision is sought. Other researches have simply used failure rates, or similar classifications (normal academic progress vs. failed or delayed academic progress). It is necessary to bear in mind that the particular nature of the criterion variable will determine the personality and individual difference correlates that are observed. Examination results, in particular, have certain limitations which are often ignored. Although giving a semblance of objectivity and reliability, the present examination procedures lack both of these features. Examinations are marked subjectively by markers who may or may not have in mind the precise features they are looking for in an examination paper. Even if these features are known, they are unlikely to be related to the educational objectives of the degree course, not logically; let alone empirically.* Furthermore, there is no specified relationship between examination results and any external (e.g. vocational) criterion. Examinations cannot, therefore be demonstrated to be valid measures of anything other than the skill of sitting examinations.

Many papers have also pointed out the unreliability of examination marks (e.g. Hartog and Rhodes, 1935; Dale, 1958; Furneaux, 1961; Cox, 1967).

It has been estimated (Jones and McPherson, 1972) that between 20% and 70% of the variance in examination marks is attributable to marking unreliability alone: in addition to differences in marking criteria between different individuals, departments and Universities, many examiners will admit that they cannot

---

*One might surmise that responses to examination stress, and the luck factor in question-spotting would be two unwanted sources of variance.

mark consistently over a single day. There is also the possibility of response unreliability, the possibility that day by day variations of mood, health, and weather will significantly affect examination responses, making them an unrepresentative sample of even examination ability.
Despite these very real drawbacks of examination marks as criterion scores, the over-riding consideration which determines their importance must be the fact that they exist and are used. Although imperfect, virtually all students sit examinations and are awarded a degree on the basis of them. The degree grade itself will often be critical in a student's later life: without a grade of 2:1 it is difficult for a student to enter a University as a lecturer or research worker. The attainment of a 2nd Class Honours or above earns a teacher £312 more per year compared with a Dip. Ed. teacher. The Civil Service is known to place high emphasis on degree results when considering prospective candidates. Thus the very real importance of a degree, and hence examination results, to a student, makes it of prime importance as a criterion variable.
The use of failure rates as a criterion avoids some of the errors involved in examination marks. Although based on examination results, the differentiation between students is much less, and errors of unreliability are consequently less*. Furthermore, students are rarely failed on the basis of a

---

* Error variance is still present, however. In an individual follow-up of failing students, Wilson(1971) found frequent effects of chance factors affecting results. One student turned up a day late for his re-sits, and hence failed. Another spent the Summer in the USA on a trip arranged early in the year, before the June exam.results were known - she was thus unable to attend the resits. However, the mere possibility of re-sits must reduce the unreliability of a single examination mark.

single examination: in many Universities, the staff bend over backwards to allow for emotional and health problems which cause a temporary period of unsatisfactory work. The failure rate is therefore more of a long term measure than examination marks. It is also a fairly valid measure of one criterion, financial and human wastage which results when a student enters for a course of study but does not complete it.

Unfortunately, failure rates and wastage rates do not measure only academic criteria. When a student fails, it may be as much a reaction to the situation and environment of his course, as a reflection on his academic abilities. Apart from more subtle reactions, mental disturbance or pregnancy may cause a student to fail.

Failure rates also have an inbuilt statistical characteristic which many researchers appear to ignore, as Wilson (1971) pointed out. Failures are an 'extreme' group taken from the population of University students, and the characteristics of these students are not likely to be normally distributed. This will make the data unsuitable for input to correlation and regression analyses, theoretically at least. In practice, researchers tend to disregard the pre-requisite conditions of parametric statistics, although it is not known how great an effect this will have on the resulting statistical conclusions. (Siegel, 1955).

<u>The Pre-suppositions of Criterion Variables</u>

i) Reliability. Criterion variables should be reliable if there is to be a chance of successful prediction, and therefore examinations and failure rates are pre-supposed to be reliable. In fact, there is a general awareness (see above) that examination results are highly unreliable (it is difficult to see how one could, in practice, get evidence on the reliability of failure rates). Jones and McPherson (1972) have pointed out the implication that one cannot expect to attain correlations of unity between any predictor and examination results: they suggest correlations of 0.7 to 0.8 are the maximum possible. Presumably, if examination marking were to become more objective, higher levels of prediction would be possible. Furneaux (1961) notes that this is indeed the case in the United States, where the prevalence of multiple-choice examinations has led to higher predictor-criterion correlations than are found in this country. (Kelsall, 1963, has suggested that, because of the reliability factor, the most one can hope for is prediction, not of rank order of accepted candidates, but of the degree of risk attached to different groups).

ii) Fundamental Unity. This is the pre-supposition that all examination performance is a psychologicaly unitary thing towards which a variety of characteristics might contribute. This assumption is implicit in the normative measurement method combined with the correlation/regression type of statistics/analysis. The converse pre-supposition is that examinations require a variety of skills, abilities, and responses, and that a given score can be attained in a variety of ways. Hypothetically, foru students might have reached the same low score by four different paths: one had poor handwriting, the second had a poor memory for details, the third had little analytic ability, and the fourth was unoriginal. There is, in fact, some recent evidence which bears on this point. Thompson (1973) reported case study material which indicated that students reach the same academic criterion (high or low on 'A' levels and University first year examinations) by diverse motivational routes. The high academic achievement of different students was often founded on different basic motives. To the extent that examinations are intended to be a measure of motivation to work and achieve, these observations would indicate that they are not a pure or unidimensional measure, or indeed any measure at all, but rather a conglomeration of variables. If this alternative pre-supposition, that the criterion is a conglomerate or system of sub-variables, is nearer reality, then a normative approach which involves the summation across students obviously has little change of successful prediction of the performance of individual students: each "path" to a given criterion score is likely to have different predictor characteristics associated with it. Whether or not the pre-supposition of fundamental unity has been believed by the researchers to be veridical, there has until recently been little choice about making it: the alternative in normative measurement method, cluster analysis, has only recently become viable as a result of the growth in the size of computers (Entwhistle and Brennan, 1971), and the idiographic measurement method remains faintly unrespectable, and lacking in methodological rigour. The pre-suppositions of the measurement methods are dealt with in greater detail in Section 5:4.

5:3 Predictor Variables

The cognitive style measures, their predictive success in the educational situation, and their presuppositions, have been reviewed in the preceding chapters. Cognitive style, however, is only the tip of the predictive iceberg, as far as academic performance is concerned.

Table 5:1 lists the major studies which have been made using different predictors, together with the typical results of these studies.

Table 5 : 1 Predictors of Academic Performance

1. Entry Knowledge

Higher School Certificate

Williams (1950) Warburton (1952) Dale (1954) Petch (1959) Forster (1959)	General agreement that there is a significant association, with the exception of Petch. Correlations between 0.25 and 0.8, but generally 0.5. Wide variations in correlation between faculties (Forster) and department (Warburton, Williams). A tendency towards better prediction in Science as compared with Arts.

Scottish Leaving Certificate Examinations

Scottish Council (1936) Gould and McComisky (1958) Pilliner (1960) Nisbet and Welsh (1966)	General agreement of a "significant" but moderate degree of correlation (Gould and McComisky), with wide variations between subject area (zero to 0.7, in Pilliner's study). Also variations from year to year: Pilliner's over-all correlation was 0.06 in 1953 and 0.33 in 1954. Science generally more predictable than Arts, although there are exceptions to this (Pilliner: French and Latin results had high correlations, whereas Maths. had a very low correlation).

University First Year Examinations

Pilkington and Harrison (1967)	r = 0.524

General Certificate of Education

Austwick (1960)
Nicholson and Galambos (1960)
Petch (1961)
Richards and Wilson (1961)
Himmelweit (1963)
Bagg (1968)
Bagg (1970)
Pilkington and Harrison (1967)
Hamilton (1968)
Maclay (1968)
Abercrombie, Hunt and Stringer (1968)
Wankowski (1968)
Cohen and Child (1969)
McCulloch, Foren and Hitch (1969)
Smithers and Batcock (1970)
Sherwin and Child (1970)

General agreement on a significant and moderate degree of relationship, with correlation ranging between 0.2 and 0.6, with exceptions (Himmelweit, Hamilton and Ba gg). Wide variations between subject areas (Austwick; 0.126 - 0.635) and faculty (McCulloch, Foren and Hitch; Social Science - rho = 0.377. Joint Honours - rho = -0.037).

Many of the studies used categorisation of groups and differences of means testing rather than correlations, which makes quantification of relationships between predictors and criterion difficult.

2. Cognitive Tests

These include AH5, Bedford College Test, Newnham College Test, NIIP Group Test No. 33, Thurston's Primary Mental Abilities,

intelligence tests by Cattell, Spearman and Valentine, in addition to a variety of lesser known tests.

White (1931)
Dale (1935)
Petrie (1948)
Himmelweit and Summerfield (1951)
Nisbet and Buchan (1959)
Valentine (1967)
Himmelweit (1963)
Kelvin, Lucas and Ojha (1965)
Pilkington & Harrison (1967)
Hamilton (1968)
Ryle & Lunghi (1968)
Banks, Kardak, Jones & Lucas (1970)
Smithers & Batcock (1970)

Many investigations have found moderate correlations with performance, but others have found none (e.g.Hamilton). The predictive value is less than that of 'A' levels, and when combined with 'A' levels, cognitive tests contribute only a negligible gain (Pilkington & Harrison). The level of correlation is not sufficient to allow prediction of rank order, generally being within the range of 0.2 - 0.4.

3. Personality Tests

Bernreuter Personality Inventory

Brotemarkle (1933) Nemzek (1938) Munroe (1942) Super (1942)	Brotemarkle found that "neurotic tendency" predicted academic aptitude at a low but significant level. The other studies found no relationship.

Guilford - Zimmerman Temperament Survey

Goedinghaus (1954) Bendig & Sprague (1954) Witherspoon & Melberg (1959)	Witherspoon and Melberg found significant correlations around 0.2. The other two studies found correlations varying between 0.2 and 0.4, depending on the particular GZTS factor, the highest being Restraint (similar to extraversion/introversion).

Maudsley Personality Inventory and Eysenck Personality Inventory

Furneaux (1956) Broadbent (1958) Lynn (1959) Bendig (1960) Lynn & Gordon (1961) Savage (1962) Kelvin, Lucas & Ojha (1965) Lucas, Kelvin & Ojha (1966) Ryle and Lunghi (1968) Entwhistle, Percy & Nesbet (1971) Kline and Gale (1971) Wilson (1971)	Strong evidence that introverted students perform better (Furneaux, Broadbent, Lynn, Bendig, Kelvin et al, Entwhistle). Contradictory evidence on neuroticism: Furneaux, Lynn, and Wilson found that neurotics performed better, Lynn and Gordon, and Savage (in Australia) found the opposite, and all other studies found no relationship. Smithers & Batcock, and Entwhistle, Percy and Nisbet, found variations in different subjects and Wilson found that the prediction of success and failure required different regression equations. Where relationships found, correlations were of the order of 0.2.

Minnesota Multiphasic Personality Inventory

Hahn and Singer (1944) Owens and Johnson (1949) McQuary (1953) Schofield (1953) Frick and Keener (1956) Burgess (1956) Stone & Ganling (1956) Grace (1957) Weiss, Segal and Sokol (1965)	Relationships have been found in the following dimensions: extroversion, immaturity, psychopathology, hypomania, and schizophrenia (all -0.2 —0.3- negatively correlated at a low but significant level).

Rorschach Inkblot Test

(often modified for group administration)

Munroe (1942) Munroe (1945) Montalto (1946) McCandless (1949) Osborne & Sanders (1949) Gaier (1952) Rust and Ryan (1953)	Contradictory results. Munroe found that neuroticism scores predicted at levels superior to the Bernreuter and to intelligence. Osborne and Sanders found significant differences in the scores of achievers and non-achievers. Other investigators found no relationships. Lavin (1965) concluded that the Rorschach "is a poor instrument for the prediction of academic performance".

16 PF

Cattell (1957) Cattell, Beloff & Coan (1958) Locke (1958) Holland (1959) Holland (1960) Holmes (1960) Warburton (1963) Locke (1963)	Cattell and Beloff found that the correlation of performance with intelligence was raised from 0.5 to 0.7 by the addition of 16 PF scores. Other investigators have found significant although less spectacular correlations with the following factors: emotional stability, high self-sentiment, confident adequacy, composure and persistence.

The compilation of this table is largely based on previous reviews by Eysenck (1947), Entwhistle (1971), Kelsall (1963), Lavin (1965), and Hack (1971).

Table 5:1 shows that the best predictors of educational performance are the various measures of entry knowledge-attainment at an earlier level of education. This would tend to argue that the determinants of performance are fairly stable, operating in a similar manner at both earlier and later levels of the educational process. However, it should be remembered that the levels of correlation are generally of the magnitude of 0.5, accounting for only 25% of the total variance in performance. Furthermore, there are wide variations in correlation in different subjects, departments and faculties. There are valid reasons for accepting this as an adequate level of prediction: the G.C.E. and other school-leaving examinations are based on school performance, and schools show

many differences in teaching method, environment, and even subject of study, from Universities. A further point is that the prediction of University performance is not a design feature of school-leaving exams. This reasoning is supported by Pilkington and Harrison (1967), one of the few studies to explicitly compare the predictive validity of G.C.E's and University first year examination. The latter was the superior measure, although the correlation was still only 0.524.
The group of cognitive tests are second to entry knowledge in level of prediction, although there are wide differences between individual tests, and again across different subjects. Many of these tests are expressly validated against University performance (e.g. the AH5, the Valentine), and hence the relatively low correlations of 0.2-0.4 are disappointing. They are not, however, surprising. As can be observed at an intuitive level, intelligence is only one factor: dimensions of personality, study habits, and factors of environment confound the validity of any single predictor.
The various personality measures, although not expressly designed for the educational situation, do have some slight predictive value, although results here are even more contradictory and confusing. Perhaps the strongest trend in the evidence is that introverted students perform better, although variations between subjects and departments are apparent here also.
The author's view of these research reports mirrors exactly the following comment in a paper by Lin and McKeachie (1973):

> "..... the list of personality variables that have been found to relate to academic achievement is more impressive for its length than for the magnitude or replicability of the relationships found. Since academic achievement clearly involves many variables and their interaction......

this would in itself be not disheartening were there some theoretical framework within which the various findings make sense, and upon which further research could be planned. But this too is lacking".

The judged adequacy of these measures as predictors of educational performance will of course depend on the criterion of adequacy. In terms of the research aims mentioned above, measures of entry knowledge, and cognitive ability, and possibly also the personality measures, go some way towards fulfilling the first two aims i.e. of maximising financial resources, and of minimising human wastage or suffering. None of these measures, however, makes a significant contribution to prediction at the individual level, to identifying particular students at work, or of outstanding performance, and diagnosing the particular combination of factors operating in those cases.

It is also necessary to consider the adequacy of these studies which have used a large number of measures together with statistical techniques such as multiple regression and multiple correlation. None of the above measures would be expected to predict individually: one might expect that the objective of the present research strategy is to obtain a high level of prediction by using a weighted combination of tests from a large battery. One such study was that of Himmelweit and Summerfield (1951), who gave a battery of cognitive and conative tests to 232 students at the L.S.E. in 1947-48. The battery included measures of (1) general knowledge, (2) ability to read tables, charts and graphs, (3) seven tests from Thurstone's Primary Mental Abilities, (4) two tests of completion and direction, (5) a short, non-verbal, and relatively pure, intelligence test (NIIP Test 70/1), and

(6) an accuracy of performance measure based on the NIIP test. The resulting multiple correlation was 0.601. Dale (1954) noted that the student sample included a wide range of ex-Servicemen, and was therefore likely to be heterogeneous with regard to intelligence, thus inflating a level of prediction which might be found in more settled times. Furthermore, as Hope (1968) points out, multiple regression techniques capitalize on chance variations, and the value of the multiple correlation obtained is likely to drop when applied to fresh data. This was confirmed by Dale (1954), who obtained a correlation of 0.41 when the same equations were applied to the 1949-50 intake of the L.S.E.

The heterogeneity of intelligence argument was also supported in a later study by Himmelweit (1963), still at the L.S.E. The previous successful tests, together with Furneaux's speed and level tests of intelligence, were given to the same groupings of students as previously. The results were:

For

Law students	r = 0.51
Sociology students	r = 0.48
Economics students	r =
Statistics sub-group	r = 0.43
Economics sub-group	r = 0.34
Government sub-group	r = 0.29

One of Himmelweit's conclusions was:

"it looks as if in each case the test which best predicts performance is the one which measures not the quality most related to the degree course, but the quality without which the student might come a cropper in that particular course. It is the limiting ability that predicts (i.e. capacity to memorise parrot-like in Law; ability to deal with figures in Economics, and the ability in the Statistics course to deal with the Economic History paper)......"

This conclusion is interesting in that it points to

1) wide variations between subjects

2) a non-linear, "necessary but not sufficient", relationship between individual variables and educational performance.

Thus it appears that studies which use a battery of predictive measures are also lacking when it comes to predicting for an individual. To paraphrase Himmelweit and Summerfield (1951), it is possible to predict the extremes of performance, but not the borderline cases, and that is where prediction is most needed. Even the moderately successful levels of prediction that have been obtained can to an extent be called illusory: in addition to differences between subjects, and date of testing (Himmelweit, 1963), differences between Universities are likely to affect the regression equations.

The Presuppositions of the Predictor Variables

The predictor variables used in the educational situation are in fact a subset of analytic concepts of cognition discussed in the previous chapter. Consequently, the arguments and criticisms developed there apply also to this subset, and need not be repeated. It is, however, worth noting the presuppositions, and the conclusions which can be drawn:

1) The dimensions are assumed to be reliable and general i.e. an individual's score on that dimension is held not to vary across different times, and across different situations where that dimension might be relevant. While there is generally supporting evidence with regard to reliability, the generality of theoretical dimensions is seldom tested. Although the relative contribution of specific and general factors to an over-all intelligence score, and in various

applied situations has been throughly explored, the same cannot be said for "over-all" personality dimensions and cognitive style dimensions. Indeed, what little research there has been into the cognitive complexity dimension has suggested high domain specificity.
The presupposition of reliability and generality of dimensions is a part of the "structural" nature of the concept of dimension: the dimension is structural, and must therefore be distinguishable from aspects of "content", i.e. variation from situation to situation and person to person. The arguments, quoted in the previous chapter, against applying this distinction to higher cognitive functioning, apply with even greater force in the context of the field situation of Tertiary Education.

2) The test construction procedure often involves the attempt to isolate a "fundamental dimension". Tests are constructed so that items inter-correlate with each other to a maximum extent, and inter-correlate with other tests to a minimum extent. Thus many tests are intended to be measures of basic and independent psychological entities.* This characteristic of the dimensional approach minimises the appearance of interactions between variables, and renders the theoretical explanation of any interactions which do appear, almost impossible. Chapter 4 argued the importance of interactions in cognition, and noted the possibility that the interaction is a primary characteristic of cognitive functioning. Furthermore, Chapter 4 favoured the "scheme", or the "Structure", as a conceptual unit which reflects some of the important

---

* There are notable exceptions to this, Cattell's 16PF being one of them.

characteristics of cognition, and it is difficult to see how these concepts can be reconciled with the concept of "dimension" as a basic cognitive unit.

Finally, it is sufficient to note the sociological detail that the theory of testing appears to have lagged behind the technology and methodology (Shouksmith, 1971). It is necessary, therefore, to examine the research methods used in predicting the educational situation in order to discover the origins of our model of that situation.

5: 4 The Research Method

Before discussing the presuppositions of research method, the more mundane difficulties, specific to research in education, should first be mentioned. In any applied science, the theories and methods developed in the pure part of the subject often need modification before they can be put into practice. Similarly with Education, the "field situation" of education produces difficulties which are often not found in the more abstract psychology laboratory situation. Thus where the psychologist recruits subjects who are asked to perform in their spare time, the educational researcher's subjects are often being investigated in University time. This places an extra constraint on the researcher to fit in with the University's schedules and timetables. This factor, together with the duration of many educational investigations (3 years, when predicting degree performance), makes for poor, or non-existent, experimental design, and hence conclusions with low internal validity. The non-scientific way in which data leads to conclusions in education is perhaps masked by the high "common-sense" validity of the conclusions, in contrast to pure psychology where data often inexorably lead to counter-intuitive conclusions. In compensation for this lack of internal validity, the external validity of educational research is usually high, again in contrast to pure psychology where it is equally usually low.

As a result of the field situation character of education research, sampling is also difficult. Whereas the population of tests and theories in Psychology is the "man-in-the-street", the population of educational research is the student in education, a person who has already gone through a rigorous selection process. This means that the educational population is likely to be more homogeneous, and distributed in a non-normal way. In practice these considerations need not weigh too heavily: except in certain applied disciplines, most psychological research in fact uses University students as subjects,* and thus does not differ from education.

The educational researcher must be especially vigilant for experimentor effects, in particular the Hawthorne effect. By selecting out a sub-sample of students to test and experiment on, he may very well be enlisting their interest and raising their motivation to do well.

In contrast to the laboratory experiment, the field investigation often lacks control of possible intervening variables. While this is generally true of educational research, it is particularly true of the prediction of academic achievement. In this case, the testing of students in University time,

---

* It is interesting to speculate whether cognitive psychology would be what it is to-day if 15-year old school leavers replaced University students as the main source of subjects.

together with the immense number of possible predictive variables, and the timespan over which the process occurs, makes it impossible to make a study the slightest bit experimental. The typical study in this area is what Campbell and Stanley (1963) call "quasi-experimental", involving an experimental effect and an observation:

X ________________________ O

Experimental effect -	Observation(s) -
Educational processes	Performance, predictive variables*

As Campbell and Stanley point out, this type of study has negligible internal validity, being open to all eight classes of extraneous variable which they mention. Furthermore, the adage that "correlation" does not indicate direction of causality" applies with particular force: the stimulation of change in cognitive and personality variables and in attitudes would be part of the objectives of the educational process, and would thus render the use of these variables as predictors of the process an absurdity. Here again, this is yet more likely with cognitive style dimensions.

Much research in this area attempts to compensate for lack of experimental control by the use of multi-variable statistical analyses. However, these methods, of which multiple regression analysis may be perhaps taken as a model, themselves make presuppositions about the nature of cognition and the nature of the educational situation.

---

* Predictive, as used here is of course logical prediction, although it <u>may</u> also be chronological prediction. Even if chronological, however, the design is still as above, and not O——X——O: this latter would require the observation of the same variables in the pre-test and the post-test.

These presuppositions are fairly widely known amongst researchers, and are also fairly widely ignored. It is worth re-stating them here:

1) The requirement of interval data. The vast majority of multi-variate analyses used are parametric and require interval data. Most psychologists would agree that psychological variables are seldom on an interval scale, and sometimes not even on an ordinal scale. However, the interval data assumptions of parametric statistics are widely ignored, often with the optimistic comment that parametric tests are "remarkably robust". However, the truth is that the amount of distortion engendered by breaking this assumption is a complete unknown, when the distribution of scores and the nature of the scale is not known (Siegal, 1956). The recent development of non-parametric methods of multi-variable analysis (Torgerson, 1958) will perhaps lead to a return to "virtue".

2) The requirement of a normal distribution. Here again, the use of variables which are known not to conform to the normal distribution (e.g. field dependence), together with the highly selected nature of student samples, militates for the development and usage of statistical methods which do not make this assumption.

3) Adequate numbers. As more and more predictive variables are found, and need to be taken into account, the increasing classification of samples increases the number of "cells" and decreases the numbers of students per cell. This means that extremely large samples are required for a broad multi-variate analysis. The problem is, of course, exacerbated when any of the variables are not normally-distributed.

4) Normative dimensions. The usual multivariate analysis itself requires the use of dimensions of individual difference on which every individual in the sample can be placed, and is probably the main motivation for the extensive use of the dimension concept despite its unvalidated nature. Other taxonomic units, for instance the type, and the hierarchical structure, have until recently been suitable only for simple statistical analysis, rather than for multi-variate analysis. However, recent developments in numerical taxonomy (Sokal and Sneath, 1963) and cluster analysis (Wishart, 1969) open up new possibilities in this area. Cluster analysis is similar to factor analysis in that it attempts the statistical representation of points in an n-dimensional hyperspace, where n is the number of variables used to classify each point (a person). A factor analysis attempts to reduce the information in the space by positing vectors from the centre of the space so as to maximise the variance accounted for. Cluster analysis attempts to reduce the information by positing clusters of points, which are analogous, as Entwhistle and Brennan (1972) point out, to galaxies of stars. Thus whereas a factor is itself a dimension, which is correlated with the predictive dimensions, and on which an individual can be given a score, a cluster is a group of people with a similar pattern on all the variables of the analysis. Thus clusters may be, in psychological terms, approximations to "types" of people. Furthermore, some clustering procedures begin with n clusters (where n is equal to the number of cases, in this case people), and progressively collapse the most similar clusters until there remains one large cluster. The way in which the clusters collapse, and the "fusion

coefficients" can be plotted graphically in a "dendrogram", which may give some indication of hierarchical structuring. Cluster analysis has as yet been little applied in psychology, and less in education. However, there has been some premising research into the structure of semantic domains (Miller, 1964; Osgood, 1968; Fillenbaum and Rappoport, 1971).

Furthermore, Brennan (1971), and Entwhistle and Brennan (1972) report the application of cluster analysis to a battery of 23 variables associated with academic performance. They report a twelve cluster solution, with 3 high performance clusters, 3 low performance clusters, and 6 intermediate performance clusters. The clusters have a certain intuitive validity, even though there is as yet no way of eliciting a predictive equation or applying tests of statistical significance.

Cluster analysis thus appears to be a promising technique, but it is not without its difficulties. The whole clustering procedure is based on the similarity matrix, which states the similarity or dissimilarity of each point with each other point. Entwhistle and Brennan use a simple Euclidean distance function: the Clustan User's manual, however, allows for any of forty different metrics to be used. Little is known of the psychological assumptions and practical differences involved in these metrics, and the assumption of a Euclidean space.*

Secondly, there are a variety of different clustering criteria which can be used. Different procedures maximise the variance included ina cluster regardless of shape, or maintain a spheroid cluster shape, collapse clusters according to the

---

* in itself challengeable (Scott, 1963).

nearest neighbouring cluster, or maximise the distance of the furthest neighbour. Where the distribution of points in the hyperspace is relatively homogeneous, these and other different clustering criteria lead to radically different results. In the absence of ways of deriving predictive equations and examining the predictive validity of the different procedures, it is difficult to know which procedure is in some way consonant with the structures of cognition.

Thirdly, although cluster analysis is less restrictive in its assumptions than normative multi-variate analysis, and can pick out, types and hierarchies, it cannot handle other types of structure, in particular mixtures of dimensions and types. Deese (1969) and Osgood (1968) note some other structured forms which may be represented in cognitive functioning and which conflict with cluster analytic assumptions.

Finally, as with factor analysis, the results from a cluster analysis depend on the input to the analysis. Thus the identification of cognitive and behavioural domains which are in some sense real, and the correct sampling from those domains, would appear to be a logical pre-requisite for meaningful results. In the case of the variables usually associated with academic performance, it is not possible to isolate particular domains, nor to sample them in any systematic fashion, and this must cast doubt on the reality of any clusters found.

One might also point out that the chief advantage of cluster analysis, i.e. its ability to pick out patterns of scores, and hence interactions between variables, is underused if the input to the analysis consists mainly of the conventional dimensional measures, as is the case with Entwhistle and Brennan's (1972) study. As was pointed out above, the theory and construction of such dimensions often minimises possible interactions.

5) Linear casual relationships. Multiple regression analysis as used in the prediction of educational performance also assumed that the relationships between variables are linear, and by implication, that there is one route to academic success and one route to academic failure. Although there is little clear evidence of curvilinear relationships (and, one is tempted to add, bearing in mind that there has been little vigilance for them), it seems intuitively likely that they proliferate in the educational situation. The 'limiting abilities' mentioned by Himmelweit (1963) and reported above, without which students cannot do well, are examples of curvilinear relationships which are only minimally picked up by correlation and regression. Similarly, it would seem likely that students attain a given level of academic performance by various cognitive routes: that academic performance is not an index with any normative psychological reality.

Conclusion

An examination of previous attempts to predict academic performance using a variety of predictors has shown that these have met with only slight success. Although the use of objective measures in selecting is justified in an actuarial sense, our knowledge of the individual differences in response to education is insufficient to allow identification of borderline cases and students at risk, and to give diagnostic or corrective help to individual students. This failure of research is to a large extent due to the unreliable and invalid nature of degree results as a criterion. However, a further important factor must be the discrepancy between the reality of the educational situation, and the model of that situation presupposed by the conventional research methods.

The conventional research method involves the use of normative individual difference dimensions, and the combination of the dimensions into a normative predictive equation. This implies a model of the educational situation involving students who are responding to a common environment, and who can be "quantified" on a unitary scale of academic performance, which is predictable from a linear combination of dimensional differences.

The common environment does not exist, however; neither "objectively" nor psychologically. The conventional research methods can, to a certain extent, take into account "objective" environmental variations, such as different departments, and subjects of study, tutors and lecturers, teaching methods, lodgings, etc., although the classification matrix, and hence the sample size required to fill it, quickly becomes very large. However, it is unable to take into account the perceptual and cognitive factors which make common objective environments psychologically diverse. A common Department, tutor, teaching method and subject of study are likely to be perceived in a variety of ways by different students: it was the argument of Chapter 4 that these

perceptual and cognitive differences are not likely to be reducible to dimensions.
Furthermore, the variety of psychological paths which can lead to a given examination score is dissonant with the normative combination of scores in a regression equation. There is, therefore, a strong requirement for the development of research methods which do not involve normative and dimensional assumptions, and which allow for the construction of internally-referenced, or intra-individual, structures but which remain rigorous and ultimately objective.

CHAPTER 6

A METHOD FOR THE INVESTIGATION OF COGNITIVE FUNCTIONING IN THE FIELD SITUATION

6 : 1 Psychology as a Science

The previous two chapters have examined some of the difficulties associated with explanatory concepts in cognitive psychology and education, and have traced them back to assumptions made by normative experimental research methods. But the basic problem, the root of the conflicting evidence and low predictive validity in education and cognition, is simpler than this. It is simply that psychology has not yet identified the determining variables of behaviour; we do not yet know what conceptual units to use when talking about functions of the organism. It is largely for this reason that the experimental method is not likely to prove efficient in the study of cognition. The externally-referenced critical variables are not the determining variables of behaviour: the externally-referenced 'control' variables are partly irrelevant to the organism, and other variables which may be important are not controlled because they are not externally referenced, and are therefore difficult to identify.

Does this mean that cognitive psychology, by the nature of its subject matter is not scientific? The answer to this question will naturally depend on which definition of science is adopted: Kuhn might say that cognitive psychology is still at the pre-scientific stage of "puzzle-solving", whereas Popper might disagree. It is both beyond the scope of this thesis, and beyond the expertise of its author, to adjudicate in a question which has received so much attention from philosophers (e.g. see Lakatos and Musgrave, (1971): however, the writer senses that Popper's viewpoint is the more widely-accepted, and will therefore accept it for the purposes of this analysis. Although it is impossible to summarise adequately a 450-page work (Popper's - The logic of Scientific Discovery, 1959), one of the main arguments was that the scientific method is not inductive, but hypothetico-deductive. Science is not a painstaking collection of observations, from which a general law is inferred, but a process of hypothesis, deduction of consequences and testing of deductions. An associated point is that scientific laws can never be proved, but are liable to disproof by dissonant observations. This is the principle that theories are scientific because they are inherently falsifiable. The origin of the hypothesis, according to this view, is independent of the testing of the hypothesis: the hypothesis may indeed come about through inductive processes, but it may equally be an intuitively-based 'hunch'. It is the possibility of falsification, and the search for dissonant evidence, which is the characteristic of the scientific-method.

The experimental method is not at all necessary to this view of science. The experiment is an efficient inductive method: it enables the scientist to build up a knowledge of the relationships between a number of independent variables and dependdant variables, and to state with confidence that the observed relationship is not "polluted" by confounding variables (although in the relatively uncontrolled experiments of psychology, this confidence is often unfounded). However, Popper has shown that this type of exercise can never lead logically to laws of science.

It is true that experiments can also be used in the search for dissonant evidence, although this is often not the case in psychology, when experiments are normally used to provide confirming evidence (much of Broadbent's work is a notable exception). The point is that while experiments can be used in this way, they are not a necessary part of the search for dissonant evidence. A simple observation, provided that it is objective (i.e. independent of a particular observer), is sufficient to disprove a hypothesis. As was pointed out in Chapter 4, the main function of experiments is to control possible confounding variables (and such variables can be confounding even in the search for dissonant evidence, leading to a Type II error), but this function, in the context of the internally-referenced variables of cognition, they cannot perform.

The normative nature of the research reviewed is not even a necessary part of experiments: a demonstration of Boyle's Law in physics does not need to be repeated on 30 flasks of "gas". It is a device to allow the control of non-physical confounding variables by randomisation, a device which, as Chapter 4 argued, is defeated by the possibility of idiosyncracy.

Thus normative experiments are not at all a necessary part of psychology's self-concept as a science. Even without experiments, psychologists could still make hypotheses, based on intuition or inference, hypotheses which can be tested against objective observation. Admittedly this procedure is less efficient and more time-consuming than an experimental approach on appropriate subject matter, but it must surely be more efficient than an attempt to change the subject matter to fit the experimental method. It was suggested that a necessary precondition of a science of psychology is that the observations made should be objective. Objectivity is often taken to imply the principle of replicability, the principle that any doubtful observer should be able to repeat the observation and come to the same conclusion. On the basis of the need for replicability, self-reports and introspective data are often held to be an unacceptable input in a scientific theory. However, this is to misunderstand the nature of objectivity, even as it is applied in that paragon of scientific virtue - physics. For instance, Einstein's theory of relativity predicted that light would be subject to the same gravitational forces as particles of matter, a prediction which would be tested by observing the apparent position of a given distant object when the sun passed close to the path of its light. An apparent change in position was observed, but this is not an observation that could be replicated in any practical sense. That distant object might never again be in the same relationship to the earth and the sun. The observation was still objective, however, in that any scientist in the same position could have made the same observation. The same argument applies to self-reports and introspections: although there are numerous difficulties in assuming that these types of data indicate anything literal about the subject, the observation of self-reports and introspections is still objective. They are still valid objects of study, even though their significance to organismic functioning is not clear.* This point of view is confirmed by the physicist P. W. Bridgman (1959, p.241):

> "Introspectional report is, by the very fact of being a report, public, and as such, proper subject for psychological enquiry.......it is not meet that an outsider like me should question the judgement of the psychologist that he can at present spend his time more profitably on more public matters than introspectional report, but as a physicist I may be permitted to express a mild surprise that this should apparently be made so much a matter of principle by the behaviourist, or that any method should be discarded which might conceivably help in unravelling the incredibly complicated maze of present psychological phenomena".

A more detailed discussion of this issue can be found in Boring (1961) and Kessel (1972).

6 : 2 Cognitive Psychology and the Problem of Taxonomy

Any research method presupposes a model of the nature of the phenomena which are to be investigated. In cognitive psychology, the basic units of functioning and the determining variables of cognitive phenomena are still unknown: the problem of a suitable research method therefore reverts to the original problem of taxonomy. Until the basic units

---

* There is of course the unlikely possibility that they are random, in which case they are still valid, but unworthy, objects of study.

of functioning can be identified, it is not possible to evaluate the potential of a research method.

Chapter 4 discussed a variety of assumptions about cognitive functioning, and came to the conclusion that the most veridical conceptual unit of analysis yet proposed was the schema concept, or the structure, as used by Piaget and Bartlett. This concept is similar in many ways to the TOTES, plans, and strategies of Miller, Gallanter, and Pribram's (1960) analysis, and also to the reverberatory circuits and phase sequenciesof Hebb's (1949) earlier work. All of these concepts imply that processing is implicit in the 'content' of cognition: that cognition is 'pregnant' with information. All of these theories have been recognised as important developments, and have stimulated a change of course in research.. It is also true that, with the exception of Piaget's work the other theories have <u>not</u> led to direct research implications: they have not been susceptible to verification or falsification. The author believes that this is because they are metatheories as well as theories, and the appropriate methodological conclusions have not been drawn. A meta-theory cannot be tested directly, but only embodied in a methodology, which itself is subjected to extended testing. The theories of Bartlett, Hebb and Miller et al are dissonant with the assumptions of the experimental method, and can only be tested by embodying their own assumptions in a compatible research method. Piaget's position has led to more direct research, but then Piaget uses a unique research method. It is experimental, but not normative, and relies heavily on logical analysis and abstraction. Thus the characteristics of experimenting on individual subjects, and abstracting conslusions across different subjects and different experimental situations are both compatible with his assumptions about the basic structural nature of cognition. However, this is only true in so far as he is investigating the structures which are likely to be species-general and situation-general. Piaget has not attempted to investigate the cognitive structures of social, complex perceptual, and epistemic functioning; structures which are likely to be idiosyncratic, and would therefore require a different research method. In attempting to design a research method capable of handling these phenomena, it is therefore necessary to begin from the meta-theory of the nature of cognitive functioning. From the present theoretical position, therefore, such a method must reflect the following key characteristics:

1) Cognitive elements are internally-referenced, and combine into structures which are idiosyncratic. Observation must therefore be made on the individual and 'defined' or 'referred' in a circular way to other aspects of the individual, rather than to objective features of the environment (as with S-R theory, and experimental methodology).

2) Cognition is systemic, and probably structural. The determining variables of functioning are not individual objective variables, but systems of variables, internally-linked. Thus the sampling of observations must be widespread over a single individual, rather than narrow and over many individuals. The experimental method, which involves a single, or a few variables, repeated over many individuals cannot

investigate the importance of those variables in the economy of the organism, especially if that importance differs from one organism to another. A widespread sampling of observations in a single organism must therefore be made, in order to allow the detection of patterns, systems and structures.

3) Importance of context. Allied to the previous point, observations must reflect the different cognitive contexts in which an organism functions, rather than being restricted to a single context. It is only by comparing functioning across different situations that the generality and functional importance of cognitive structures can be assessed.

4) Importance of cognitive content. While not prescribing which phenomena are valid objects of study, the schema or 'structure' concept suggests that cognitive content has previously been under-studied. Thus observations of perceptions, attitudes, beliefs, opinions, classifications and knowledge are of prime importance in the understanding of functioning in real-life situations.

6 : 3 The Method of Structured Observation

6 : 3 : 1 Logical basis:

One research method which is compatible with the assumption (it is not claimed to be the only method) is the method of Structured Observation, and is shown schematically in Diagram 6:1. Its logical basis is the same as that of the scientific method generally, and involves the measurement of determining variables and the explication of their relationship to a criterion variable. It is in the nature of the observations and, in particular, in the hypothesis of the relationships, that this differs from the normative experimental method. The primary data have the same logical status as normal experimental independent variables, with the exception that they are not assumed to be externally-referenced or 'physical', and the criterion variable is exactly the same as the experimental dependent variable. The predictive indices, criterion prediction, and the lack of control groups, are where this method differs from experiment.

Diagram 6:1. Method of Structured Observation

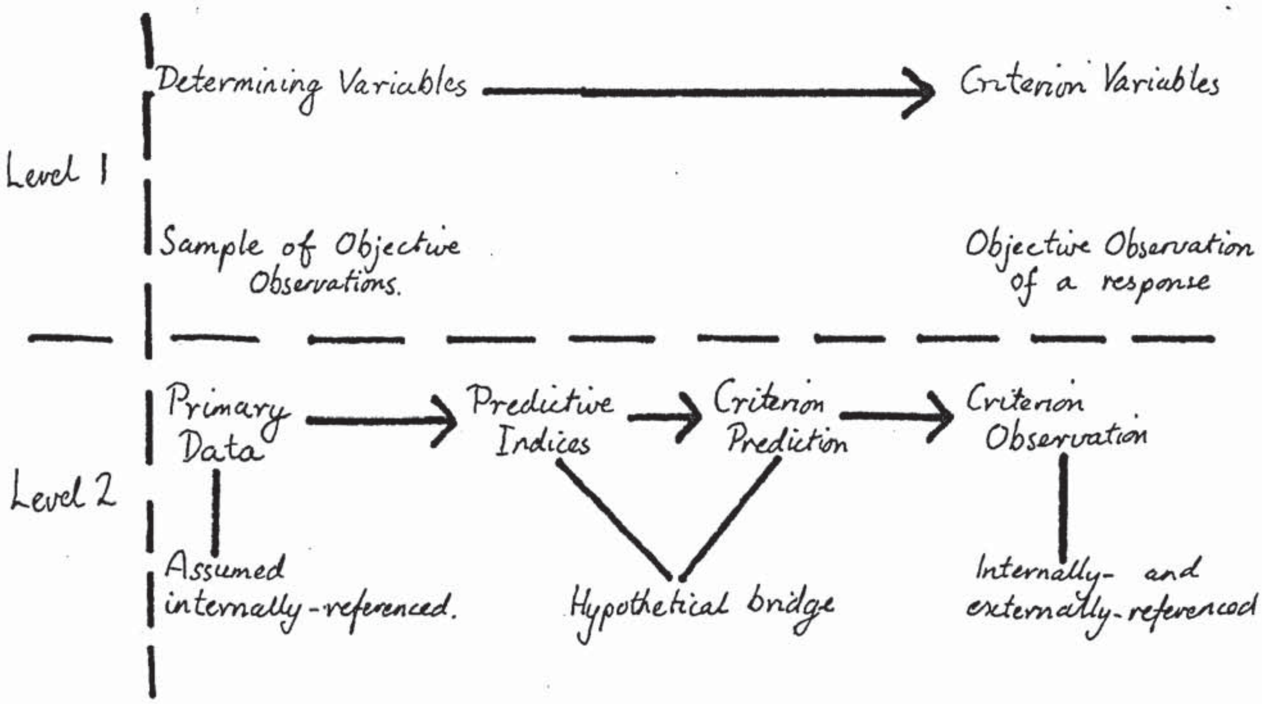


Diagram 6:1 Method of Structured Observation

The method is shown at two levels. At Level 1, the logic of the method is similar to the logic of the scientific method, involving a prediction from a determining variable (usually called the independent variable of an experiment) and the criterion variable (the dependent variable). Observations of both variables must be objective, in the sense that they are not affected by a particular observer. It is at Level 2, which shows the details of the predictive process, that the structured observation method differs from experimental methods and other observation methods in psychology. The primary data (observations of the determining variable) are sampled from a wide area of potentially relevant variables, and are internally-referenced. The criterion may be either idiographic or normative. The primary data are tested against the criterion by a two-phase process:

a) construction of a hypothetical bridge, which consists of formulae which specify how a criterion prediction can be derived from the primary data. Calculation of the criterion prediction is intra-individual.

b) the testing of the criterion prediction against the criteria.

6 :3:2 Primary Data

The only restriction here is that the observations must be objective, in the sense that they must be independent of a particular investigator. This, of course, is easier to preach than to practice, as can be seen from reports of "experimenter effects" from the most unlikely situation (Rosenthal, 1966).

The primary data will comprise a wide sampling of observations reflecting diverse aspects of an individual's functioning in different situations, but all of which are conceivably relevant to the criterion. Also, at least some of the observations should allow expression of idiosyncracy - i.e. would be material 'produced' by the subject. Bearing in mind the sole restriction of objectivity, numerous techniques can therefore be used to collect the primary data. Amongst these might be the following:

1) Semantic differential techniques.
2) Kelley 'rep' test techniques - i.e. involving ratings on subjectively based categories.
3) Rating scalesof other types.
4) Check list.
5) Normal questionnaire questions, perhaps combined with rating scales.
6) Sociograms.
7) Interview data and introspections - opinions, arguments and stated beliefs.
8) Projective material, sentence and paragraph completions, essays, etc.
9) Probes. The use of "experimental" phenomena which have been found to be content-linke; for instance cognitive dissonance, retroactive and pro-active interference in memory, memorability and subjective structure, recognition thresholds, G.S.R.'s etc.

With the use of tape-recorders, there is no difficulty in making all of these observations objective. Even projective material and production material (essays, etc.) are objective in the simple sense. The limits of usage of these types of material are therefore defined by the difficulties in making their interpretation objective, i.e. the difficulty of converting idiosyncratic primary observations into objective predictive indices.

This problem, in the form of translating idiosyncratic symptoms and reports of symptoms into a normative theory of function and malfunction, has long been the main difficulty of psychoanalysts and clinical psychologists, and is exacerbated in that context by having to look at, quite literally, the whole man, and predict the whole range of responses rather than just a single response. The problem may therefore be less acute in cognitive psychology; the author can propose no simple solution beyond the rather weak suggestion that an account of the structure of knowledge, perceptions, beliefs and other cognitive content might help.

6 : 3: 3 Predictive Indices

The predictive indices are a bridge between the primary data and the criterion. Their distinguishing characteristic is that they are calculated separately for each individual, rather than being based on a normative analysis. Where a normative criterion is used, the rules used for calculating the predictive indices must also be normative, and they can then be thought of as hypothesising general cognitive structures. In this case they differ from the normative experimental method in that:

1) they can embody relationships which are not capable of investigation by normative statistics e.g. the "necessary but not sufficient" relationship.

2) they are not restricted to a particular experimental situation, but can relate primary data which originate in different behaviour contexts are hypothesised to be functionally similar.

In discussing the construction of predictive indices, it is necessary to distinguish between those based entirely in the restricted-alternative data, in which all observations are classifiable (e.g. conventional questionnaires), and the "production" data, in which there may be idiosyncracy. These different types of data give rise to different problems in the construction of predictive indices.

a) Restricted-alternative data. There is no great logical problem here. The investigator simply sets out rules which specify how different primary observations may be combined to form a predictive index. Overlapping indices (which include a common primary observation) may be hypothesised, as may "higher order" predictive indices, based upon a combination of lower order indices. The investigator can make a number of alternative hypotheses, which involve combinationsof different predictive indices being used in the calculation of the final criterion prediction. This criterion prediction is tested in the normal way against the observed criterion, and if the two corresponde, the hypothetical structure is supported.

"The investigator <u>simply</u> set out the hypothetical structure": "simply" refers to the logical aspects of the procedure, but not, unfortunately, to the practical aspects. The procedure is logically simple in that it is a classical hypothesis testing situation. The hypothesis, and in this case the structure of predictive indices, are independent of the hypothesis-test. The matiematical correctness of any procedures used in deriving the predictive indices are irrelevant to whether those indices are a good predictor of the criterion: indeed one could elicit the structure of predictive indices from a soothsayer on Brighton Pier, and if found to be a good predictor, they would be perfectly valid. Of course, the likelihood of their being a good predictor is another matter.

It is in the hypothesising of predictive indices which are <u>likely</u> to give a good criterion prediction that the main difficulty lies. A large part of the procedure must inevitably be intuitive: what the investigator thinks likely, by intuition or prior observation, to be related to the criterion, and the relationships between these primary data, which seems a likely reflection of cognitive functioning. This heavy intuitive contribution must be inevitable, because of the huge number of possible combinations from the wide sampling of primary data, and because of the large number of possible relationships between different indices.

Mathematical and statistical techniques may also be helpful in investigating which predictive indices have a predictive potential. If sufficient numbers of subjects have been investigated, conventional multivariate analysis (multiple regressions, factor analysis, etc.) will give indications of which primary data are related to the criterion, and how they might be combined into higher order predictive indices. However, such indications are only a first step towards the hypothesising of particular relationships and particular predictive indices: normative multivariate analysis, as has been pointed out, embodies assumptions which are contrary to those espoused here, and are not suited to eliciting non-linear relationships. Furthermore, the side sampling of primary data not only increases the time which must be spent on each subject, but increases the number of subjects who must be investigated if factor-analytic and regression methods are to be useful.

The recently-developed technique of cluster analysis is likely to be of considerable importance in the construction of predictive indices. Cluster analysis is a procedure for identifying clusters of people with a similar pattern of scores on a number of observations, and thus reflects an assumption of the present work, namely that different people can arrive at the same cognitive goal (the criterion score) by a variety of cognitive routes (the 'systems' of predictive indices). It can thus give an idea of which primary observations contribute towards a predictive index, and which primary observations distinguish between predictive indices. It is, however, still only a first step at the present state of play: the type of relationship which might hold between two or more primary observations and the criterion cannot yet be deduced from a cluster analysis, although it may be possible to develop an appropriate method in the future.

Graph plotting of combinations of primary observations against each other and against the criterion may also be helpful in eliciting likely predictive indices, as may other methods of pattern analysis which the author has yet to hear of.

b) Production data. The building of a bridge between primary observation and criterion is far more difficult when the primary observation includes idiosyncracy. Logically, it would seem to require a hypothesis of the rule or code by which idiosyncracy can be generated in the first place. This itself could not be a simple classification code, but would have to embody the type of logic involved in a generative grammar à la Chomsky, perhaps a generative grammar of semantics. This, of course, is far beyond present psychological knowledge.

However, while such a theory is necessary for a truly objective interpretation of idiosyncratic or production data, it may be possible to take some account of idiosyncratic or production data, it may be possible to take some account of idiosyncracy by using the investigator's intuition, or another judge's intuition, as a "natural" theory of semantics. The human judge can "interpret" production data, and judge its relevance to, e.g. compatibility or incompatibility with, performance on a criterion. The judgements required can be more or less complex, and more or less explicitly specified, and hence would involve a greater or lesser sacrifice of objectivity. Research has tended to show that fairly high inter-judge agreement can be obtained where the judgements are tightly specified and not too complex.

It may be argued that the attempt to interpret idiosyncracy in the context of real-life functioning is exactly the problem faced by clinicians and psycho-analysts, and many experimentally-trained psychologists would claim that no satisfactory solution has been found. In that case, why should an unsuccessful method be adopted in the exploration of cognitive phenomena? In the first place, the criterion in cognitive psychology may be simpler and more isolated: interpreting idiosyncracy in its relevance even to phenomena of the complexity of job selection or academic performance, is still easier than interpreting symptoms in their relevance to a shadowy and undefined malfunction of behaviour or mind. In the second place, Chapter 4 has argued that idiosyncracy is present in any case in complex perception and higher cognitive functioning: in which case the attempt must be made to study it whether methods are sufficiently refined or not.

6 : 3 : 4 The Criterion

The criterion, against which the hypothetical structure is tested, can initially be either normative or idiographic. If idiographic, and research is carried out on a single subject, it must of course involve sufficient alternatives such that the prediction of a correct response is not attributable to chance. Taking the conventional 0.05 level of significance, this therefore requires the correct prediction of one response from a set of twenty possible responses.

While there are situations where investigation of a single subject is worthwhile, and possibly preferable (e.g.clinical situations), a science must always take such investigations as a first step towards generalising the conclusions to other subjects and situations. A psychology of John Smith is not of much benefit to anybody else. In most cases, therefore, the criterion must be normative and objective.

A further necessity is that the criterion must be of functional importance to the organism: it is little use investigating the eye-blink rate unless eye blinks can be shown to be non-random. Similarly, academic performance might be a poor criterion for the girl who only goes to college to find a husband (although, converseley, it might also be a good criterion, if the finding of a husband requires her to do less work).

The testing of the criterion prediction against the criterion can be effected by the normal statistical techniques, except that non-parametric methods might be more appropriate to the nature of the data.

6:4 The Problem of Control

The control of possible confounding variables is the major tool of science in the establishing of valid conclusions. It is the efficiency of control that allows a scientist to state that under certain conditions, $x = f(y)$. The method of structured observations does not involve any physical control, beyond the selection of specific observations, nor any control by randomisation, beyond the final hypothesis test. This would seem to lead to the conclusion that the method, although logically sound, is fore-doomed to failure: it simply will not lead to any conclusions at all. However, Chapter 4 argued that physical control, at least as applied to higher cognitive functioning and in real-life situations, is more illusory than real. Firstly, physical variables which may not be determining variables are under experimental control, and secondly, internally-referenced determining variables are not susceptible to physical control. Since also such internally-referenced variables are unknown and may be cross-situational as well as cross-individual, neither can they be controlled by randomisation. The method of structured observation is therefore, in practice, in the same boat as the experimental method, at least as far as cognitive phenomena, and real-life criteria are concerned. Furthermore, if one accepts the common complaint that experimental results do not generalise to real life situations, the method of structured observation could not be more unsuccessful than normative experiments in giving an explanation of cognitive functioning in the field situation. However this is the negative side of the coin: on the positive side, the method of structured observation gives a hope of efficient investigation by substituting for physical control, measurement control. Instead of physically controlling confounding variables, or rendering their effect neutral by randomisation, the present method attempts to measure them and allow for them. The measurement of confounding variables is made more likely by the wide sampling of the primary data: the discounting of confounding variables is made possible in the structure of the predictive indices. The success of measurement control thus depends upon the development of knowledge and techniques for constructing predictive indices in particular cognitive domains and behaviour contexts.

Is the method of structured observation likely to succeed in producing valid explanations of cognitive functioning in life situations? It is certainly a "high-risk" procedure: if the criterion prediction does not predict, there is no way of locating the error within the structure of predictive indices. It may perhaps be likened to the "focus-gambling" strategy in concept acquisition (Bruner, Goodnow, and Austin, 1956). However, the method does involve presuppositions which are similar to the model of cognition adopted here:

1) Determining variables are internally-referenced, and to some extent idiosyncratic. Even where the primary data are not idiosyncratic, the complexity of the predictive indices allows some reflection of idiosyncracy. Furthermore, the construction of predictive indices is intra-individual.

2) Cognition is hierarchical and systemic: the inter-

action between variables is primary. The predictive indices can be constructed to reflect these assumptions.

3) It is necessary to examine the "functional relevance" of cognitive elements. This is facilitated by a wide sampling of primary data observations, which are not necessarily situation-linked.

4) Cognition is oriented to action-schemata. The primary data and predictive indices (even when these are intended to be measures of perception) are defined in terms of the response.

The method therefore seems more promising, to the author, than the experimental method, which makes contrary assumptions. Perhaps it can be said that it is at least worth trying. The results of such a trial are reported in Chapter 8.

SECTION II

EMPIRICAL

CHAPTER 7

AN EXPERIMENT:

COGNITIVE STYLE, TEACHING METHODS, AND ACADEMIC PERFORMANCE

7 : 1 Introduction and Aims of the Experiment

The review of cognitive style variables (above) suggested that cognitive styles are potentially useful predictors of educational criteria, despite the previous paucity of research in this area. It was also noted that the field situation of education presents some methodological problems, the most prominent of which is the practical difficulty of carrying out a rigorous and controlled experiment, which will give conclusions with a high internal validity. The control problem arises both from the wide number of possible confounding variables, and from the purely practical difficulty of manipulating a life process which is not, after all, in existence for the convenience of investigating psychologists. The lack of methodological rigour was observed in connection with the review of individual difference dimensions used as predictors (Chapter 5), and can also be seen in the review of teaching methods, below.

An initial aim of the present investigation was therefore to counteract this tendency, and to produce an experimental study which was as rigorous and controlled as possible.

General academic performance, as measured by final examination results, is an unsuitable criteria for an experimental study. Evidence as to the unreliability of final examinations has already been presented. But, in addition to this, examinations are a compound criterion of many types of performance in many different sub-environments of the University. An experiment using such a criteria would have to control, at the very least, subject of study, teaching methods, tutors, and working environments. To circumvent these control problems, the present experiment investigated performance in response to a single course of study, delivered by a single teacher in a single environment. This has the effect of increasing the internal validity of any results, but of decreasing the external validity. However, the detriment to external validity was minimised by using a Complementary Studies Course*, delivered to students from a variety of Departments of the University.

Instead of using a single teaching method as the medium of delivery of the course, the two teaching methods of lectures and project work were used, contrasting one performance against the other. Although there has been a considerable amount of research into teaching methods, there has been very little research into individual difference measures as predictors of teaching methods, and the writer has found no study which used cognitive style dimensions in this connection. Furthermore, the methodological difficulties of the educational situation have meant that these studies tend to be methodologically poor.

In assessing the results of studies on teaching methods, it is necessary to examine closely the criteria by which performance is measured. Apparently disparate criteria like preference and performance might be more closely related in terms of cognitive variables than different types of performance. Memory for the factual content presented in the teaching methods is probably quite different from ability to use that content. If performance is measured in terms of examination or essay results, it is

---

*i.e. a General Studies programme which 1st and 2nd year students attend.

often not apparent what criteria are being used in marking the examinations and essays. Without precise specification of the criteria by which performance is measured, one experiment cannot be compared with another. Measurement criteria are also important in the specification of the aims of the teaching method. Varying teaching methods are employed in tertiary education for a number of reasons, one of which is to train different skills. If lectures are intended to promote factual knowledge, and seminars are supposed to promote critical faculties, it is obviously inadequate to compare lectures and seminars on a memory criterion alone. Unfortunately, education is not quite so orderly as this account would suggest, and the objectives of different teaching methods are implicit rather than premeditated. Thus the educational researcher can only use a spread of measurement criteria which he hopes will reflect the various objectives. Most of the studies reported below do not take into account the aims of the different teaching methods. Nor do they take account of possible interactions between aims, and individual difference variables.

There are a number of studies who have compared lectures and small-group discussions. Thus, Ward (1956) compared performance under lectures and discussion groups, and found that 'academically superior' students showed greater retention of materials presented, greater comprehension of the work, and a greater expression of individual differences, in group discussion rather than lectures. The reverse results were found for 'academically inferior' students. Unfortunately it is difficult to infer exactly what psychological functions underlie 'academic superiority'.

Koenig and McKeachie (1959) used a number of personality measures to predict criteria of preference on three teaching methods - lectures, small-group discussions and independent study. They hypothesised that:

1. Students with a high need for affiliation would prefer small-group discussions.

2. Students with high self-reliance would prefer independent study.

Neither of these hypotheses was supported.

3. Students with a high need for achievement would prefer small-group discussions and independent study to lectures. This hypothesis was supported for women, but not for men.
4. Middle need for achievement students would prefer lectures. The rationale is that these students have previously been shown to be high in fear of failure, and will consequently prefer the more directive situation of the lecture. The hypothesis was supported.
5. A further non-hypothesised finding was that women of greater flexibility participated more in small-group discussions. However, this relationship was reversed for men. Flexibility is negatively correlated with authoritarianism, and this result has possible implications for cognitive predictions.

Beach (1960) compared performance, as defined by knowledge increment, under four teaching methods:

1. Independent study
2. Lectures
3. Discussions
4. Autonomous small groups (five people in each)

The sociability of subjects was measured using the appropriate sub-scale of Guilford's STDCR, and it was hypothesised that subjects would do better when their sociability matched the sociability of the teaching method. The results generally confirmed this hypothesis. High sociability subjects did better in discussions and worse in lectures. There were no significant differences where independent study was concerned, but with autonomous small groups the more sociable subjects did worse than the less sociable. This last result was contrary to expectations, and no explanation could be offered.

In a questionnaire study, Rogers (1960) identified two independent factors relevant to lectures, a general factor of listening ability, and a smaller factor of self-reliant or independent thinking. There was only a slight relationship between listening ability and academic achievement.

McCollough and Van Atta (1960) found that flexible students benefited more from independent study than do those with more rigid personalities.

These results are summarised in Table 7:1. There is nothing very striking in them, other than the reversal of relationship between flexibility and participation in small group discussion from women to men.

Teaching Method	Better performance from:	on criterion:
Lectures	Less sociable (Beach, 1960)	Knowledge
	Middle need for achievement (high fear of failure) (Koenig and McKeachie, 1959)	Preference
	Academically inferior (Ward 1956)	Performance
	High listening ability (Rogers 1960)	Examinations
Discussions	More sociable (Beach 1960)	Knowledge
	Academically superior (Ward 1956)	Performance
Small-group discussions	Less sociable (Beach 1960)	Preference
	High need for achievement women only	Preference
	High flexibility women	Participation
	Low flexibility men (Koenig and McKeachie 1959)	Participation
	No differences in need for affiliation.	
Independent study	High need for achievement women only (Koenig and McKeachie 1959)	Preference
	High flexibility (McCollough & Van Atta 1960)	Performance
	No differences in self reliance of sociability.	

Table 7:1 Personality correlates of performance in different teaching methods

It is difficult to find any explanation for this, and it may be a result of differential sampling. Generally, the results show that flexible students perform better on small-group discussions and independent study than do rigid students. One might hypothesise that such students would perform better on all teaching methods, so the significance of these results cannot be gauged without further research. Beyond this, the results show that sociable people prefer sociable teaching methods, that better listeners perform better than worse listeners in lecture situations, and that academically superior students perform better under discussions than they do under lectures. Beyond the flexibility results, there is little evidence here to help with cognitive style predictions.

Of partial relevance to the present study is a survey by McLeish (1968) into student attitudes towards lecture methods (oracular methods) and seminar/tutorial methods (participatory methods). It is of partial relevance only in that ratings of favour/disfavour are not the same thing as performance. However, it is a reasonable assumption that a person likes what he is good at, so there should be at least a weak relationship between these two measures.

The attitude questionnaire was initially administered to a sample of teacher-training students and staff at Cambridge Institute of Education. The results showed that all groups disfavoured the lecture and favoured seminars and tutorials. There were no differences between the latter two. When students were classified into high status and low status, according to their examination mark or tutor's estimate of that mark, high status students showed a much greater tendency to favour participatory methods as contrasted with lecture methods. This result agrees with Ward's (1956) finding that academically superior students perform better in discussion groups.

The students had also been administered with the Survey of Educational Opinions, and a number of personality and attitude scales. McLeish therefore obtained three groups from students with high preferences for oracular methods, high preferences for participatory methods, and intermediate preferences for both. These groups were compared on 34 variables, 20 of which differentiated significantly between the groups. McLeish summarises the results as follows:

The lecture-centred group favour formal methods of teaching in primary and secondary schools more than do other groups. They are relatively tough-minded and strongly favour the values of security, submission and workmanship. The intermediate group are more interested in power; they are in favour of utilitarian values and corporal punishment; and they express high regard for aesthetic values. Their uncertainty score was highest. The discussion-centred group, strongly disfavouring the lecture, were more prone to naturalism in their thinking - that is they believed that spontaneous growth was better than a forced development. They value new experience and freedom for themselves more strongly than do other groups. Their certainty score is highest and they are also more anxious. They were more radical in their educational views, were more extroverted, and were more satisfied with their jobs as teachers than the other two groups.

McLeish (1968), p 23

Relationships between these groups and religious and political affiliations were also established. No Conservative disfavoured lectures, and no radical favoured them; the highest number of Church of England members appeared in the intermediate group.

As the sample in the college consisted entirely of "mature and experienced teachers", the survey's political and religious findings are unlikely to generalise to a university sample. However, the general characteristic of formalism, tough-mindedness, workmanship, conservatism and lower academic status for those who prefer lectures, and naturalism, freedom, certainty, extraversion, radicalism and higher academic status for those who prefer participatory methods, are likely to generalise. Again they implicate flexibility and authoritarianism, or the more general quality of cognitive complexity, into the preference variance for teaching methods. If the trend follows through, individuals of greater complexity should do relatively better on seminar methods (and probably project work), and individuals of lesser complexity should do relatively better on oracular methods.

The present study involves measuring the levels of student performance under two different teaching methods - project work and lecturing. It is desirable to specify precisely the teaching methods used in the experiment, so that it is possible to infer how the experimental results will generalise to the field situation (i.e. education in Great Britain). It is also desirable that the teaching methods used should simulate as closely as possible their counterparts in the field situation in order to obtain maximum external validity.

The problem of simulating the field situation is realised upon consideration of the questions "what is a lecture?" and "what is project work?". A few minutes thought indicates a long list of dimensions on which different lecture courses and different project work courses can vary. These lists are not exhaustive. Lectures can vary on:

- different aspects of the teacher variable

a) status

b) knowledge

c) personality

d) enthusiasm

e) social interaction style

f) conception of unstated objectives

g) speed of delivery

h) clarity of exposition

- degree of structure of the subject matter
- use of questioning or discussion
- use of visual aids, printed notes etc.,
- transmission of facts vs transmission of ideas i.e. information or stimulation
- motivation of the students
  a) necessary for career
  b) necessary for exams

c) necessary for both

d) necessary for neither

- subject matter.

Similarly, project work can vary with respect to:

- tutor variable

- allocation of time and physical resources

- individual vs small group vs large group

- group characteristics

a) personality interactions

b) interaction styles

c) different leadership styles vs leaderless.

- tutor/student ratio

- type of problem situation

a) structured vs unstructured

b) open vs restricted solution

c) assigned vs selected

d) informational vs creative vs production

- motivation of students.

'Project work' and 'lectures' are not unitary in nature, but rather a collection of processes and practices with common characteristics. This being so, it is not possible to construct a template, measure the template and then infer to the field situation. If the template matching model is not applicable to teaching methods, to use an analogy from research on perception, then the feature analysis model is applicable. The lecture methods and the project work methods are not unitary processes which can be simulated, but they can be defined in terms of their common and individual attributes. Their common attributes would be represented by a number of statements about their position in the educational process, and the aims which they share in common with the rest of education.

Their individual, or differential, attributes are those attributes which serve to distinguish between the two teaching methods. There are certain superficial distinguishing attributes which can be listed immediately, and refer to the physical form of the teaching method. Thus lectures are primarily didactic, with a relatively active teacher and relatively passive students, they have a lower teacher-student ratio, and take place in a single auditorium. Project work is a participatory method in which the students are more active. This consequently requires a higher teacher/ student ratio. Project work may be confined to one room or the educational institution, or may include the life situation for which education is preparing the students. This latter is particularly prominent in a "Technological University" with a fair proportion of Sandwich Course students.

Beyond these physical defining attributes, there are more tangible attributes which are difficult to identify. A promising method of identifying these attributes might be to compare the objectives of the different teaching methods. However, objectives are usually explicated with respect to a course of study, and involve detailed statements of the knowledge objectives. It is rare for researchers to consider the relative values of different teaching methods in promoting their objectives, and difficult from the literature, to abstract a set of objectives applicable to the teaching methods irrespective of their course content.

In the absence of stated objectives, the functions and primary characteristics of lectures and project work, as they are rated by students and staff in questionnaire studies, can be examined. There have been a number of studies of how students and staff perceive the lecture method (Hale Report, 1964; Marris, 1965; Saunders et al, 1969; Students' Society Committee of the Royal Dental Hospital School of Dental Surgery, 1966). Generally speaking, the results showed that staff were more favourable to lectures than were students. There was much inconsistency in the different views expressed, and some evidence that individuals differ systematically in their opinions of lectures. Thus applied science students felt that lectures should give information, whereas arts students thought they should be stimulating (Smithers, 1970a). Smithers (1970b) found that extroverts viewed lectures as a performance in which the lecturer should be entertaining, confident, and at ease. He also found that unstable students and authoritarian students wanted the lecturer to be definite and certain: he should give full notes, set clear goals, and convey information lucidly.

Bearing in mind these variations in opinion, a large number of statements referring to the functions of the lecture, or the ideal lecturer can be culled from the four reports mentioned above. These 'objectives' can be classified as:

1. Logistic. e.g. "most economical way of presenting information" and "the only way to deal with the shortage of books."
2. Procedural. e.g. comments about ideal pace, delivery and exposition.
3. Affective. e.g. "sharing of enthusiasm" and "stimulation of individual work."
4. Cognitive.

The present discussion is not concerned with logistic and procedural considerations. Affective objectives probably play an extremely important part in defining the difference between lectures and project work: one of the most obvious features of project work is that it simulates the life situation and is 'relevant', and should thus raise the student's motivation to participate. Unfortunately, while intuition indicates the importance of affective factors in different teaching methods, it is particularly difficult to measure such factors: furthermore, the present study is mainly concerned with cognitive styles.

The statements referring to cognitive factors are presented in Diagram 7:1. classified according to Bloom's taxonomy of educational objectives (Bloom, 1956). Also shown in this diagram are the cognitive objectives of project work. In fact there has been very little research on the objectives of project work: one study (Eggleston and Kelly, 1970) has succeeded in designing an assessment scheme for project work without even mentioning educational objectives! The present objectives are taken from an associated study in the Department of Education (Small, 1973) in which project tutors rated a variety of objectives as they applied to the project process used in Technician Engineer training. It should be noted that these projects were 'production-type' rather than 'discovery-type'.

The diagram shows clearly that lecture objectives are associated with the lower part of Bloom's taxonomy, whereas project work objectives are associated with the higher part. Although not shown in the diagram, it seems likely that project work will be less efficient for the lower level objectives. An experiment by Bligh, reported in Beard and Bligh (1971), has shown that the performance of students in lectures tends to follow the bias of objectives noted here. Bligh used objective tests at eight cognitive levels to evaluate the effect of lectures in psychology. He found that where terminology, facts, general principles and simple comprehension were concerned, the test results showed comparable gains; where skill in applying knowledge was concerned, there were variable gains, but where analysis, synthesis, and evaluation of results were concerned, the test results showed negligible and non-significant gains. The objectives inferred from statements by teachers and students are predominantly associated with knowledge and comprehension and to some extent application. This is the pattern that student performance follows. Project work objectives are primarily concerned with application, analysis, synthesis and evaluation skills, although whether these objectives are fulfilled is not known.

BLOOM's TAXONOMY OF COGNITIVE EDUCATIONAL OBJECTIVES

KNOWLEDGE | COMPREHENSION | APPLICATION | ANALYSIS | SYNTHESIS | EVALUATION

LECTURES

- Includes in lectures materials not readily accessible
- Brings text-books up-to-date
- Enables most recent research to be reported
- Indicates sources of reference
- Enables demonstration of how to organise a topic, how to build up a complex diagram or argument
- Enables introduction of topics too difficult for student to tackle alone.
- Indication of topics for further enquiry
- Presents material clearly and logically
- Should involve clear, orderly synopses, logically planned
- Should emphasise basic principles
- Introduce subject, and set it in its context
- Provide a framework for the course
- Indicate methods of approaching the subject
- Enables student to understand the basic principle of a subject
- Provide discussion of problems and possible solutions

PROJECT WORK

- Consolidation of course work
- Development of planning ability
- Awareness of interdependence of course subjects
- Development of the ability to communicate
- Development of the ability to acquire and use information
- Development of the ability to evaluate the usefulness of techniques, instruments, tools, or materials for specific purposes
- Development of problem solving ability

Diagram 7:1 Objectives of Lectures and Project Work

An interesting point here is that many of the statements about lectures emphasise that the lecturer should supply a logical structure or framework on which to 'hang' the information: this logical structure appears to be what the project work students are trained to acquire for themselves. The possession of an appropriate super-ordinate structure is the goal of analysis and synthesis, and the key to effective evaluation. It appears that lecture students are not able to acquire this structure for themselves, because of the nature of lecturing, and must thus be presented with it.

It was originally intended that the information provided under the lecture and project work sections of the experiment would be identical: in view of the above inference about the objectives of the two methods, it was decided that it would be preferable if the information were varied such that the lecture students were provided with a super-ordinate structure, and the project work students were not. It would also be desirable to measure the achievement of the different levels of objectives.

Furthermore, if the structuring of the information is as important as is suggested by the statements above, it is reasonable to suppose that individuals may differ with respect to their ability to construct or evolve super-ordinate structures. If there is such an individual difference dimension, individuals high on it (i.e. more able to synthesise structures) would perform better on project work, both relative to low individuals and relative to their own performance under lectures (in which they would be bored). Individuals low on the dimension would function more efficiently where their abilities of comprehension and application were not impeded by their inability to create structures: they would prefer the more organised situation of the lecture.

The ability to create structures would seem to be related to certain cognitive style dimensions, notably cognitive complexity. One might infer that the person most able to create structures would also be most able to tolerate ambiguity and novelty. This in turn would imply that cognitive complexity or authoritarianism were involved, in view of the work connecting these traits with dissonance resolution (reported in Chapters 2 and 3). However, an experiment by Wyer (1964) indicates that things are not so simple. He found that 'integrative ability', a similar concept to ability to create structures, was correlated in the 30's with verbal ability, negatively correlated with cognitive integration. The latter two are the hypothesised sub-components of cognitive complexity. Thus the investigation of the point is not an empty question. Thus the present experiment has the following primary aim:

To investigate the value of selected cognitive style dimensions as predictors of the attainment of specified educational objectives in response to a specified course of study which is presented by means of lectures or project-work.

7:2 Design

The course of study was delivered to two independent groups. The course itself was divided into two sections, each consisting of four two-hour sessions. Criterion tests were administered in three testing sessions, at the beginning of the course and at the end of each session, and the predictor tests were presented at the end of teaching sessions throughout the course. One section of the course was taught by the lecture method, and the other section was taught by the project method. The order of these two teaching methods was counterbalanced on the two groups, although the order of the sections of the course was maintained constant. Orthogonal to these variables was the battery of cognitive style tests: the total sample was classifiable into two groups according to whether they were above or below the median on each cognitive style variable. Thus for some hypotheses each subject was his own control, while for others an independent groups design was applied. The design of the experiment is summarised in Diagram 7:2. It was not practical to set up a further two groups with the sections of the course presented in reverse order, and neither was it possible. Although for purposes of criterion measurement, the two sections were independent, the information presented in the first section was a pre-requisite to a full comprehension of the second section. However, it is not felt that this weakness in experimental design is important; even if order of presentation of sections had an effect, it is difficult to see how such results would lead to theoretical conclusions.

7:3 Subjects

The subjects numbered 16 students, all second year students drawn from the Complementary Studies programme of the University of Aston in Birmingham. The Complementary Studies Programme is a course of general studies which is compulsory for first and second year students at the University.

The breakdown of the total sample was as follows :

	Group 1	Group 2
Male	10	5
Female	0	1
Total	10	6
Origin	Students who had opted for a course on Information Science	Those students who had opted for a course on Biological Sciences
	Diverse departmental origins.	Diverse departmental origins.

The experimental sample above was a sub-section of the students who attended the courses. Students with incomplete data had to be eliminated from the course. The testing sessions T1, T2 and T3, were particularly

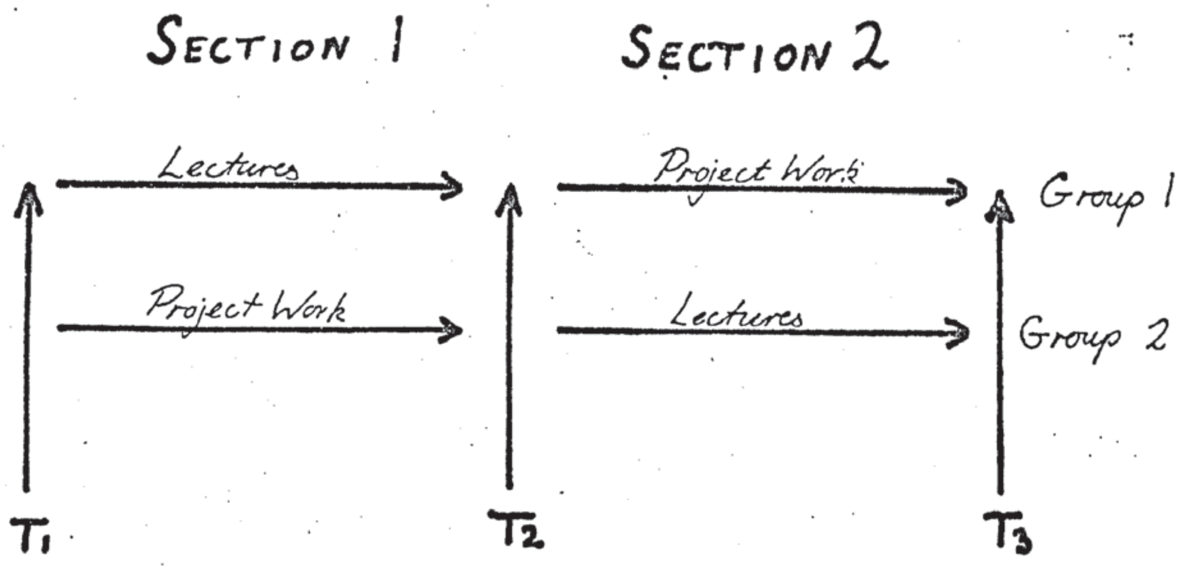


Diagram 7;2. Design of the Experiment.

<u>Independent Variables</u>:

1. Teaching method.
2. Order of presentation of teaching method.

Individual difference dimensions:

3. Cognitive Complexity (ModRCRT).
4. Cognitive Complexity (PCT).
5. Cognitive Differentiation (CFT).
6. Authoritarianism (Bal. F).
7. Automatisation (SCDT).
8. Intelligence (AH6) - 3 scores.
9. Creativity (UoO)*.
10. Convergent vs. Divergent Thinking (AH6 and UoO).
11. Direction of Interest (DIQ).
12. Reflection vs. Impulsivity (CFT-R/I).

<u>Dependent Variables</u>:

1. Attainment of lower cognitive objectives.
2. Attainment of higher cognitive objectives.
3. Creativity (UoO)*.

* Tautological prediction omitted.

important in this respect, since absence from these could not be rectified by testing later in the course. The initial size of the groups in the experiment was respectively 11 and 26, and although a low experimental mortality rate was expected, the unfortunate reduction in experimental sample size from 37 to 16 was not anticipated. It was partly due to a tradition of missing lectures on the part of Group 2 students (whose Complementary Studies course tutor disagreed with the principle of compulsory lectures): this probably caused a reduction in the size of Group to from 26 to 20. However, the major factor was the 1972 Coal Strike, which resulted in a disruption of University timetables, and a re-scheduling of Section 2 of the course from the Easter term to the Summer term: this in turn resulted in the absence of a number of 'Sandwich students', who had gone out into industry, and in a clash between the Complementary Studies course and revision for examinations. It was, naturally, the course which suffered. Thus an initially respectable sample size became only just acceptable.

These factors also affect the randomness of the sample: the initial self-selection of students in opting for the two courses was enhanced by the self-selection of students missing lectures. Thus the present sample cannot be considered representative of any particular population.

7:4:1 Procedure

The course, on human thinking and problem solving, was delivered in two parts:

Section 1	Human Thinking
Section 2	Techniques of Solving Problems

The two sections were delivered sequentially, but with different teaching methods for each section to each group. It is necessary to specify a fairly precise operational definition of "Lecture" and "Project Work" in the context of this experiment.

Teaching Methods: 1. Lecture. The conventional didactic situation in which a relatively active lecturer, at the focal point of the auditorium, delivers an address to relatively passive students. The lectures were two hours in length, broken by a quarter-hour coffee break after the first 50 minutes. Overhead projector slides were used to illustrate points, and to summarise the over-all structure of the particular lecture. Questions and discussion of points were encouraged, but discouraged when a discussion moved too far from the point of the lecture, or continued for too long. The lecturer used a printed "handout" for his notes, and then gave it to the students at the end of the session. The lectures differed from normal departmental lecturing in two ways, both embedded in the "General Studies" nature of the course:

- Students knew that they would not be examined on the course content, and therefore took no notes. The use of duplicated handouts to some extent obviated the need for note-taking, but in this case were probably ignored by the students.
- Because of the lack of extrinsic motivation towards examinations, the lecturer felt a special need to stimulate the intrinsic interest of the students in the course. Thus, the present lectures by containing a lower proportion of "hard facts", names, and dates, and a higher proportion of general concepts, counter-intuitive observations, and weak jokes, differed from the usual departmental lectures.

In pursuit of the earlier discussion of differences between lectures and project work (above), the lecturer made a special effort to give the total course, and each lecture within it, a clear theme or logical structure, and this structure was emphasized in the lectures.

2. Project work. Each group of students tackled one project, which extended over 4 two-hour sessions during one section of the course. The Section 1 project was to produce a "model" of the different factors and variables associated with human thinking. The Section 2 project was to produce a codified procedure for tackling a novel problem. It can be seen that these projects are largely concerned with information acquisition and sorting, and possibly with creative thinking to a certain degree. They are also relatively unstructured, and do not lead to a "correct" solution. They can also be contrasted with projects involving production of hardware, and projects where the student selects his own task.

The project was carried out entirely within the four two-hour sessions, and used as reference material the "handouts" which had served as the basis of the lectures to the other groups. The information available to the lecture and project groups was thus identical, except that the attention of the lecture group was directed to the course and lecture structure, whereas the project group needed to infer this, themselves, from a consideration of the material.

The projects were pursued in group-form rather than individual form. The students were asked to form themselves into groups of three to five people. The supervisor arbitrarily selected a leader for each group, and it was his task to direct the discussion and maintain its relevance. The leadership was changed each session so that each student in a group took a turn as leader.

The supervisor informed the students of the project task, of the availability of information, and of their freedom to work as they wished during the session. He also suggested that one possible method of working was for each student in a group to take and read a part of the session's handout (generally the handout consisted of four or five parts, each part consisting of two foolscap pages double spaced). The group could then come together and attempt to discuss and collate the material with respect to the project task. However, it was emphasized that this was only a suggestion.

The project supervisor moved from group to group with the aim of stimulating discussion where this appeared to be flagging. However, in view of the need to maintain constant the information to the two groups of students, the project supervisor and lecturer both took care to limit their intervention in discussion to the posing of questions, and avoided expressing points of view or delivering further information.

The project supervisor and the lecturer were, of course, the same person, namely the experimenter.

7:4:2 Timetable

The course was preceded by a two hour "testing session" which included an explanation of the nature of the course, and an appeal to the students to cooperate in the research. The students were also told that no specific information about the research would be revealed until the final session at the end of the course. Thus, although the students had no expectations or knowledge of experimental hypotheses, they did know that they were "guinea pigs", and a possible "Hawthorne" effect must be borne in mind.

After the introductory talk, the students were administered the Uses of Objects Test, the AH6 and the Closure Flexibility Test.

Also in this session, a procedure was instituted which attempted to exclude the possibility of experimenter effects, and simultaneously guaranteed confidentiality of results to the students. Each student was handed a folder and asked to label it with a pseudonym. All tests which were taken during the course were placed in the folder and handed in. The experimenter then marked tests, selected the data which were of value, and communicated results to the students (see below). Thus there was no possibility of the experimenter knowing the test scores of any individual student and consequently no possibility of him treating that student differentially in a way which was relevant to experimental hypothesis.

The first testing session was followed at weekly intervals by the four "course" sessions, and by the second testing session. The course sessions were lectures for Group 1, and project work for Group2. Cognitive style tests were administered in the last ten-twenty minutes of most sessions.

At the second testing session, the second Uses of Objects Test was administered, together with a measure of the attainment of educational objectives, and a measure of problem solving ability (see Criterion Variables, below).

There was an unfortunate delay of approximately 9 weeks between the second testing session and the beginning of the second section of the course. As mentioned previously, this was a result of the 1972 Coal Strike, and the consequent disruption of University timetables. The second section, consisting of four two-hour sessions at weekly intervals, proceeded in the same way as Section 1, except with Group 1 doing project work, and Group 2 attending lectures.

The third testing session in the following week included a third Uses of Objects test, and a second measure of the attainment of educational objectives. Finally, the students were given an explanation as to the nature of the experiment in which they had taken part.

It was thisthird testing session which unfortunately proceeded an examination by only one day, and hence caused the heavy experimental mortality described above.

7:4:3 Course Content

The course content was selected with the following aims in mind:

1. To motivate interest in psychology, and its practical relevance to the individual.
2. To give the student a knowledge of the presumed nature of the mind - in particular of perception, cognition and action.
3. To give the student a knowledge of how individuals differ in these.
4. To give the student a knowledge of his differences from others.
5. To increase the students' insight into the workings of his own mind.
6. To give the student a knowledge of the procedures and techniques available for solving problems.
7. To foster the ability to apply this knowledge in:
   a) maximising the speed, quality, and efficiency of problem solving;
   b) maximising the productivity of learning.

The total content was divided into two sections, the first on human thinking, and the second on problem solving, with emphasis on improving one's own ability to solve problems. Each section was divided into four sessions, and each session into two-to-five parts. The content was fully set out on "information sheets", which project work students could use, which formed the lecturer's notes, and which were given to lecture students after the lecture. Each session's "information sheets" were about 8 pages in typed foolscap, usually double spaced. Thus the total quantity of handouts for the course numbered some 66 sheets of foolscap. For this reason, the course content is summarised by title of information sheets, together with a brief explanatory sentence, in Table 7:2. A full set of sheets is available for inspection upon application to the writer.

As can be judged from the content summary, the course aimed at breadth rather than depth of knowledge. This was in keeping with the aim of the course, which was to present information relevant to the individual's ability to solve problems. It was also more suitable in the context of teaching general studies students, where the burden of making the course interesting falls entirely on the lecturer, rather than on the intrinsic motivation of the students.

Table 7:2 Course : Human Thinking & Problem Solving

Summary of content:

Section 1 - Human Thinking

1st Session

Part 1. The Eye and the Camera. Making point that the eye and the camera are not comparable.

Part 2. A Constant Percept of a Changing World. Some information on visual "constancies."

Part 3. A Changing Percept of a Constant World. Effect of motivation on perception, and perceptual defence.

Part 4. The Importance of Expectations. Perception and expectancy.

Part 5. Perception is a Construct. Evidence pointing to the constructive and hypothetical nature of perception.

2nd Session

Part 1. Introduction. The importance of imagery and perception to thinking.

Part 2. The Visual Icon.

Part 3. Pattern Recognition. Template matching and feature analysis.

Part 4. Attention and Synthesis. Neisser's analysis-by-synthesis.

Part 5. Imagery and Dreams. Synthetic processes of perception without the proximal stimulus.

3rd Session

Part 1. Memory: Introduction. Importance of memory to thinking and problem solving.

Part 2. The Nature of Memory. Memory as an aspect of a changing system. Demonstration of Carmichael, Hogan and Walters' experiment.

Part 3. The Principles of Memory. Characteristics of highly memorable information.

Part 4. Cognitive Consistency. The importance of this in thinking. Account of best-known experiments.

4th Session

Part 1. Plans and Strategics. Account of Miller, Gallanter and Pribram's concepts.

Part 2. The Types of Thinking. Serial and parallel processing.

Part 3. Concepts and attitudes. Cognitive, affective and behavioural components.

Part 4. Cognitive biases. i.e. concepts with an implicit emotional 'flavour'.

Part 5. Creativity. Characteristics of creative thinking.

Section 2 - Problem Solving

1st Session

Part 1. Heuristics. Explanation of heuristics (particularly Polya), algorithms, and stages in solving problems.

Part 2. A Preliminary Heuristic model.

Part 3. 1st Stage. Awareness of the existence of a problem. Some relevant considerations.

2nd Session

Part 1. The Creative Solution - introduction.

Part 2. The generation of solutions. Techniques of stimulating creative thinking. Suspended evaluation, forced compliance, synectics, etc.

3rd Session

Part 1. Information Search. Collection of information, importance of scheduling, and information structuring; study habits and speed reading.

Part 2. Analysis of information. Some characteristic mistakes of thinking (de Bono) and logical fallacies.

4th Session

Part 1. Information Search - continued. Confirming and disproving hypotheses. Importance of negative instances of concepts.

Part 2. Conceptual "good figures" i.e. imposition of serial order, and polarisation of information.

One of the aims of the course was to give the student a knowledge of how he differed from others in terms of his thinking. In this context the cognitive style tests which were used as predictors, were offered as a service to the students. Having completed the test, the students' score was fed back to him the following week on an information sheet which gave a brief account of the theory and correlations of the dimension. The aim was also embodied in the constant emphasis which the course placed on different perceptions of the same situation, the ability, as it were, to look through another person's eyes, and consequently change one's own perceptions.

7:4:4 Independent Variables : Test Administration

1. Closure Flexibility (Concealed Figures) Test. This is essentially a group administered version of Witkin's Embedded Figures Test. Information relevant to reliability and Construct validity was presented in Chapter 2. The test was quick, easy to administer and score, and presented no problems. It was intended as a measure of field dependence vs field independence, sometimes called "cognitive differentiation" or articulated vs global functioning, and high scores indicate

field-independence. Although it was not possible to include a standard test of reflection vs impulsivity in this battery, it was possible to derive a score which might approximate the dimension from the CFT. It will be remembered that error scores or RT scores on certain perceptual tasks are normally used as an indicator of reflection vs impulsivity: although the CFT is a rather different perceptual task, the calculation of the usual CFT score involves the subtraction of errors from correct answers. To obtain a rather crude measure of reflection vs impulsivity, the experiment has therefore, additionally, expressed the number of errors as a percentage of correct responses. This second index was then included in the analysis of results.

2. Balanced F Scale. A group administered questionnaire in which the subjects check their agreement or disagreement with a statement on a 6-point scale. The test was reviewed in Chapter 3. There were no problems in administration or scoring. The test was intended to measure authoritarianism, with high scores being more authoritarian. The version of the test used here is presented in Appendix 1.

3. Paragraph Completion Test. A test in which subjects are asked to write a paragraph on a stated subject. The test was reviewed in Chapter 2. The scoring is subjective, based on criteria, and specimen answers, laid down in a scorer's manual (Schroder, Driver and Streufert, 1967). There is said to be a high inter-scorer reliability between trained scorers: the present scorer, in being untrained, or at best self trained, would be unlikely to reach this level. The final score is the mean of the scores assigned to the two best completions. The test was intended as a measure of abstractness vs concreteness, with high scores being more abstract. The test is presented in Appendix 1.

4. Modified Role Construct Repertory Test. A test in which subjects rate ten people who are assumed to be important to them, on ten provided constructs. By means of a matching procedure described in Bien, et al (1966), a measure of the extent to which the subject uses the constructs differentially can be derived. If all the people are given identical ratings on a particular construct, then the subject is not using that construct to differentiate between people. The final score is a summation of the number of times the constructs are used identically on different people. The maximum possible score of 450 would indicate that all constructs were being used identically, and hence that the subject had a simple cognitive system. A score as low as 100 could indicate a relatively high level of cognitive complexity.

The copy of the test used in this study is presented in Appendix 1. The instructions on the first page produced an unforseen difficulty, in that the subject was asked to fill in the name of an actual person who fulfilled the roles in his life (i.e. person you dislike, person with whom you feel most uncomfortable, and immediate superior), but it was not made plain that the ten roles should involve ten different people. In fact, two subjects duplicated a singlename, and three subjects entered the name in three roles, or duplicated names on two roles. The following simple arithmetical correction was applied to the scores of the former two subjects, and they were included in the analysis:

1) When one row comparison is repeated, it will contribute 10 points to the new total. Therefore subtract 10 from the total score.
2) Determine the average number of points per comparison for the other rows. This is done by dividing the remaining score by the number of remaining row comparisons (45 - 1 = 44).
3) Add the row average to the remainder score.

This correction quickly becomes inaccurate with increasing repetition of names, and the other three subjects were eliminated from the data analysis for the purpose of the prediction from this test.

The Mod RCRT was reviewed in Chapter 2: it was intended to be a measure of cognitive complexity in person perception.

5. AH6. This is a relatively short (30 min.) measure of high level intelligence. It yields scores of verbal ability (V), numerical and diagrammatic ability (N & D), and total ability (V + N and D). The subjects are described either by a raw score, or by a grade (A, B, C, D, E) based on norms drawn from the performance of University students. The raw scores were used in the data analysis of this study. The test was simple to administer and score, and presented no difficulties.

6. Uses of Objects 1. The Uses of Objects test which was presented at the first testing session was used as a predictor variable. The test consisted of the instruction "Below are three everyday objects. Think of as many uses as you can for each," followed by " A barrel", "A paper clip", and "a brick". The test was untimed.

The test was intended as a measure of "creativity", although there is little evidence that it relates to any practical creative ability. Perhaps it is better described as a measure of ideational fluency.

The marking of the test presented some problems. It is usually thought necessary to categorise the lists of uses before counting them, in order to prevent some subjects from building up a high score through endless minor variations which are not really different uses. Thus, it is argued, "storage of liquids" and "storing wine" are not different uses of a barrel. However, in marginal cases, the marker's task is difficult: are "storing liquids" and "storing flour", for instance, different uses? They are, of course, different in that they involve different substances, but they are both instances of storage. The assignment of the level of abstractness of the different concepts is left to the scorer. In the present study, a second scorer categorised all the uses suggested for one of the "objects", giving an inter-score reliability of 0.76: an acceptable but unimpressive level of correlation.

The categorisation of uses also allows bonus points to be given for uses which were suggested by less than 25% of the sample: in the present study, an additional point was given for each "rare" use.

7. Convergent vs Divergent Thinking. The score on this dimension is obtained from the scores on the AH6 and the UoO1, in the manner suggested by Hudson (1966). Raw scores on both the Uses of Objects, and the AH6 Total, were rank-ordered, and placed in 5 categories of A-E, based on the proportions 1:2:4:2:1. The bias in grades was then calculated, giving scores from +5 (Intelligence -A; Creativity -E; Converger) to -5 (Intelligence -E; Creativity -A; Diverger).

8. Speed of Colour Discrimination Test (SCDT). The intention of the experimenter was to present the Stroop Colour Word Interference Test, and to use Broverman's regression equations to obtain a score of strong vs weak automatisations. However, enquiries through the National Foundation of Educational Research, and to the USA failed to obtain any copies. As this is not the sort of test which can be easily made up (it is printed in colour), the SCDT, sent as a substitute by the NFER, was used. (See Appendix 1).

The SCDT has two parts. In the first part, subjects have to put the first letter of the colour in which a line of dots is printed. The dots are of four colours; blue, green, orange and red. He is asked to work as quickly as he can consistent with accuracy. In the second part, the subject has to write the first letter of the colours in which words are printed, but the words themselves are the names of the four colours. Again, speed and accuracy are exphasised. If Part 1 gives a control score (C), and Part 2 gives an interference score (I), the score used in the present study was the interference score as a percentage of the control ($\frac{I}{C} \times 100$).

Broverman (1960) reasoned that the performance of strong automatiser would be less affected by the inference of the printed names: according to this reasoning, the interference score of strong automatisers would be a higher proportion of the control score. Unfortunately,the data from this test was incomplete, with two subjects missing.

9. Direction of Interest Questionnaire. This is a very brief (15 item) questionnaire which measures the bias towards an interest in the external world rather than in the internal world of thoughts and feelings. This measure has been found to be related to nurses' attitudes to patient care, and patients' symptomatology and treatment expectancies. (Caine, 1970; Smail, 1970; Caine and Leigh, 1972). It has also been found to correlate with Tomkins' Conservatism scale (-0.65) and Uses of Objects (0.50) (Caine, Wijesinghe, and Wood, 1973). High scorers have internally-directed interests.

This test was included in the battery for personal reasons rather than for theoretical interest.

7:4:5 Dependent Variables : Test Administration

1. Uses of Objects 2 and Uses of Objects 3. The Uses of Objects tests presented at the second and third testing sessions served as criteria to all the predictor variables, with the exception of creativity (uses

of Objects 1) and convergent vs divergent thinking (also derived from UoO1 scores).

The second and third tests differed from the first in the "object" which served as stimuli. UoO2 had "A tin of shoe polish", "A blanket", and "A sheet of currugated iron", and UoO3 had "A key ring", "A pillow case", and "A ten-foot length of rope". The test administration procedures and scoring procedures are the same as for UoO1.

2. Truth-Functional Test (T-F test). This was intended as an objective test of the attainment of educational objectives, and is based on the T-F test devised by Bligh (1971). The T-F test consists of statements with which the subject is asked to agree, disagree, or indicate ignorance. The statements are constructed so that the subject who answers correctly can be said to have gone through certain types of logical operations, and these, in turn, are related to levels of educational objectives on Bloom's hierarchy. Thus, Bligh (1971) gives the following examples:

Q1. Repetition consolidates learning. Agreement with this is said to involve the recognition of a single key phrase.

Q2. The capacities of long and short memories provide evidence that they are of a different kind. Agreement with this is said to imply the ability to make a simple inference and remember it. It should be noted that this is only the case so long as the inference presented by the statement has not been presented by the course of study as a "fait accompli". This reservation applies to all of the "higher order objectives": while it was simple, in the present study, to know precisely what had, or had not, been presented as part of the course, the same cannot be said for the usual University course, which inevitably requires the student to select from and read items from, a reading list.

Q3. If you practise playing table-tennis with your right hand, your left-hand game will remain unaffected. This is said to involve the ability to apply information that has been acquired, requiring the subjects to relate examples that were not presented to principles that were.

Two T-F tests were constructed for use in the present study, one for use in the present study, one for each section of the course. Both tests had 33 questions, and aimed to sample at the six highest levels of educational objective on Bloom's taxonomy:

Type		Number of questions
3	Recognition of key point/principle	8
4	Memory of simply stated relationship	8
5	Making of simple inferences	8
6	Application	5
7	Analytic thought	2
8	Synthesis	2

The questions referred to information presented in the four sessions of each part of the course with approximately equal frequency for each type of objective. Numbers of correct "agree" and "disagree" answers were also approximately equal. The two T-F tests are presented in Appendix 1, with correct answers filled in. (With respect to "correct" answers, many of the statements are so general that no psychologist would unreservedly agree with them: they were nevertheless "correct" in the context of the material presented in the course). For each correct answer, the subject would score two points: "don't know" and incorrect answers scored, respectively, one point and zero.

The construction of questionnaires such as these is difficult, according to Bligh (1971). In the writer's opinion, it is particularly difficult in the content of a general studies course, which deals in generalities, and even platitudes, and which attempts to change the student's viewpoint rather than give him a body of hard facts. The difficulty of creating statements which are based on "fact" or "themes" presented in the course also leads to a doubt about how the subjects judge the statements: it seems unlikely that the subjects' memory of information presented is sufficient to allow them to go through the logical processes 'implied' by the statements. For instance, Bligh (1971), gives the following statement as an example of a test of analytic thought:

> "All the evidence for there being two types of memory has been obtained by psychologists observing other people and animals."

Bligh suggests that this is an instance of checking that "All A is B", and hence requires the recall of instances of A, and their categorisation as B or not B. However, it seems unlikely that students will be able to recall this material, and particularly unlikely in a general studies course. It seems more probable that the answer will be based more on the general impression, which is perhaps nearer to "recognition of a key phrase" than to "analytic thought."

In the present context, the statements requiring "ability to make a simple inference" have the prior requirement that the subjects recall the "fact" on which the inference is based, a "fact" which may be buried in a mound of principles and generalisations. The recall of such a fact may well be a more difficult task than the making of an inference. In these cases again, the writer suspects that answers were based on the general impression of the statement.

For the purposes of data analysis, three scores were derived from the T-F tests: lower order objectives, (Types 3, 4 and 5), Higher-order objectives (Types 6, 7 and 8), and total.

3. Problem Solving Ability. It was originally intended to develop a test of problem solving ability, to check if the course was fulfilling

its primary aim. Furthermore, as the course was primarily aimed at creative problem solving, it was felt that this should involve insightful solutions. However, pilot studies showed that it was not possible to design a short (45 minutes) test of insightful problem solving which would also yield a usable score. It is particularly difficult to design problems whose solution require a flash of insight, and such problems usually take some time to solve. In addition, the requirement of objective marking means that they cannot be open-minded, i.e. they must have correct solutions. Thus in order to obtain a sufficiently differentiated score (e.g. 0 - 5), five problems are required in the test, and each one will take 15 minutes at the minimum. Such a test would not be sufficiently important to merit 75 minutes testing time. Furthermore, an attempt to use a single open-ended problem produced insuperable difficulties in the subjective scoring. It was felt that the design of a scoring manual, which would specify sufficiently precise criteria, would be a major study in its own right. Consequently, no measure of problem-solving ability was used in the present study.

7:4:6 Other Measures

1. Lecture Evaluation. At the end of the lecture sessions (T2 for Group 1 and T3 for Group 2), the students were asked to fill in a lecture evaluation sheet (presented in Appendix 1). This asked them to rate, on a 6-point scale, the following qualities of that half of the course:

- vocal relevance
- interest
- clarity of expression
- organisation of information
- visual presentation
- delivery .. audibility
  .. speed
- rapport with class
- amount learned
- problem solving (improved vs no change)
- subject (easy vs difficult)

Although this was not repeated on any control group, it was felt that it would provide a check against the possibility that the present lecturer was, in any way, strongly divergent from the norm.

The students' suggestions for further courses were also invited. This was, to some extent, a 'safety- valve' question, as no further courses were intended.

The lecture evaluation also provided valuable feedback on the lecturer's mistakes or misjudgements. However, it was desirable, in the present case, that the change in lecturing performance between the two sections of the course should be minimal: the lecturer therefore exerted will power, and refrained from examining the lecture evaluation sheets until the course was completed.

2. Project Evaluation. The project, as used here, is somewhat less usual in the University context, and there was not the same need to check it against the norm. However, a rating schedule comparing it with the more conventional course, and assessing its value, was belatedly designed, and submitted to Group 1, at T3 only. The project evaluation sheet (see Appendix 1) asked students to rate whether the following qualities were increased or decreased by <u>that</u> form of project work, <u>as compared with</u> the typical lecture course:

- interest	- motivation to work
- memory	- ability to use the information outside the course
- comprehension	- inter-relatedness of different parts of the course

The experimenter had no need to exert willpower with this measure, as it was presented on the last session.

7:5 Results

For convenience in examining the results, the main variables used in the design of the experiment are shown in Diagram 7:3

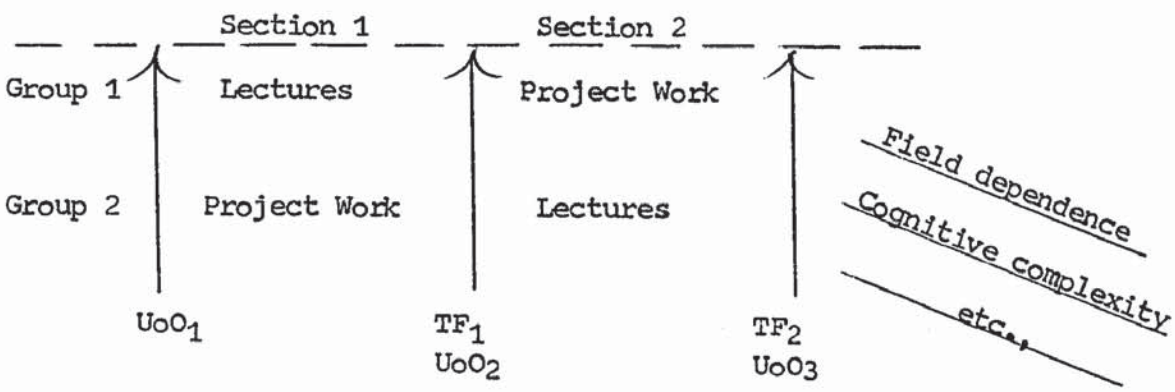


Diagram 7:3. Summary of design: main variable of the experiment. This design is, in effect, repeated for each cognitive style variable.

a) Control variables. It is first necessary to examine the differences between Sections 1 and 2, and Groups 1 and 2. The means and standard deviations of subjects' score on the Truth-Functional Tests are shown in Table 7:3. A visual inspection of this table leads to the following conclusions:

1. Performance, on all indices was better in the second section of the course than in the first.

2. Groups 1 and 2 were approximately equal in performance on TF2, but Group 1 were better on TF1.

TF1			TF2				
LO	HO	T	LO	HO	T		
24	21	45	24	21	45		Random Answering
29.90	23.20	53.10	31.30	29.40	60.60	m	Group 1
5.15	3.66	6.32	3.61	2.62	4.94	s	
25.00	19.00	44.00	31.67	27.83	59.17	m	Group 2
2.77	3.00	4.80	4.23	3.08	5.54	s	
28.06	21.63	49.69	31.44	28.81	60.06	m	Groups 1 and 2
5.01	3.96	7.28	3.86	2.89	5.23	s	

Table 7:3 Means and standard deviations of scores on the Truth-Functional Test for Section 1 (TF1) and the Truth-Functional Test for Section 2 (TF2), for Group 1, Group 2, and the total sample. The scores obtained from the Truth-Functional Tests are:

a) Attainment of lower-order objectives (LO), i.e. Types 3, 4 and 5

b) Attainment of higher-order objectives (HO), i.e. Types 6, 7 and 8

c) Attainment of total objectives (T) where T = LO + HO.

The scores which would be obtained by answering the tests at random are indicated in the first row of the table.

To take the hypotheses in order, it would seem quite reasonable that performance should increase between the first and second sections of the course, as the information contained in the two sections, as mentioned before, was not independent. A knowledge of the nature of perception and thinking is likely to give increased knowledge of different techniques of solving problems. It is therefore reasonable to use a one-tailed test of significance. For this hypothesis, each subject is his own control, and it would be wise not to make parametric assumptions about the data. The Sign test fits these features. Table 7:4 shows the results when the sign test was applied to the direction of change of scores between TF1 and TF2. It will be seen from this table that Group 2 subjects increased their scores on lower order objectives, higher order objectives and total objectives, all increases being significant at a level of p = 0.016. Group 2 subjects increased their scores on higher order objectives and total objectives (p = 0.011), but decreased their scores very slightly on lower order objectives. This would indicate any or all of three possibilities:

1. The second section was easier than the first. This possibility is confirmed by comments to the lecturer, to the effect that the subject matter of Section 1 was difficult and confusing.

2. The TF2 was easier to answer than TF1. Although it is not possible to state whether or not this was so, it would be surprising, indeed, if the two tests were identical in difficulty.

3. The attendance at Section 1 contributed to ability to answer TF2. As commented above, this seems likely.

With respect to the second hypothesis, that there is a difference between groups 1 and 2, a two-tailed test of two independent groups was used, as there are no a priori expectations.

LO	HO	T	
n = 9* 0.746 x = 5	n = 10 0.011 x = 1	n = 10 0.011 x = 1	Group 1
n = 6 0.016 x = 0	n = 6 0.016 x = 0	n = 6 0.016 x = 0	Group 2

Table 7:4 Comparison of TF1 scores with TF2 scores by means of the Sign Test. The table shows the probabilities associated with the observed direction of change in scores between TF1, and TF2 (One-tailed test).

n = size of sub-sample
x = number of subjects whose score (TF2-TF1) was negative
* = in this case, one subject showed no change, and was eliminated from the sample.

With respect to the second hypothesis, that there is a difference between groups 1 and 2, the comparison is between two independent groups, and the Mann-Whitney U test was applied (Siegel, 1956). Although there is no a priori reason to believe that one group would perform better than the other, and consequently a two-tailed test is applicable, this is a control variable, and the negative pay off from a Type II error is greater. It may be convenient, therefore, to use a one-tailed test, rather than to alter the conventional levels of significance. Table 7:5 shows the results of this analysis. It can be seen that Group 1 differed significantly from Group 2 on all classification of TF scores, but not in the second section of the course. Furthermore, there was a significant difference in the higher order objectives, when the two TF tests were summed to give a total course performance. These results indicate that the two groups, although 'randomly' selected (where 'randomly' is used in its conventional euphemistic way to indicate that the experimenter had no control over the selection), in fact differed systematically, and should not be combined. This consideration should be borne in mind when assessing the experimental results, as the disastrous experimental mortality mentioned above means that there is no choice but to combine the groups for the rest of the analysis: they are simply too small to consider separately.

	LO	HO	T	
U =	12^{*}	11^{**}	7^{***}	TF1
	n.t.	n.t.	n.t.	TF2
	n.t.	14^{*}	16	TF1 + TF2

For all comparisons, n1 = 10 and n2 = 6. For these sample sizes, the maximum values of U associated with the different levels of significance are:

$^{*}p$ 0.05 14
$^{**}p$ 0.025 11
$^{***}p$ 0.01 8 n.t. = no observable trend

Table 7:5 Comparison of Groups 1 and 2 on TF1 and TF2 and on total course performance. Mann-Whitney U Test. Table shows observed values of U and critical values of U for different significance levels. For all significant differences, Group 1 scores were higher than Group 2.

Before dealing with the experimental hypothesis, it is worth noting the top row of Table 7:3; this indicates the scores which would be obtained if the TF tests were answered at random (the maximum possible scores are double the value). It will be seen that the mean scores are very little above these random scores, and, indeed, in the case of Group 2 scores on HO and T of TF1, are below the random score. There are two possible explanations for this: firstly, it is possible that the subjects learnt very little from this course, and secondly, it may be that the test was insensitive to what the subjects did learn. It will be seen from the Lecture and Project work evaluations (below) that the students thought that they learned something, in which case the second explanation would be the more likely. In this context, the comments made in the Dependent Variables section (above), with respect to the difficulty of constructing the TF tests, and the logic required to answer them, should be borne in mind. In constructing the tests, the experimenter attempted to make the negative statements (to which the correct answer was 'disagree') sound inherently plausible: it is possible that the scores which were below the random answering level are a result of him doing the job too well.

b) Main experimental hypotheses.

The main aim of the experiment was to investigate the predictive potential of the battery of cognitive style tests, initially with respect to performance under different teaching methods. The paradigm hypothesis is:

"that individuals who are high on CS1 will perform better on project work relative to lectures". The direction of the prediction, of course, varies with the different cognitive styles.

The hypothesis involves each subject as his own control. Many of the cognitive style tests are known to differ from the normal distribution (e.g. the Mod RCRT), so again a non-parametric test is preferable. For the purpose of the analysis, Groups 1 and 2 were combined. The hypothesis was tested by means of the Wilcoxon Matched-Pairs Signed-Ranks test, using the subjects rank on the dependent variables as the input. Thus the procedure was as follows:

1. Subjects were classified into "above median" and "below median" on the cognitive style dimension (CSi).

2. Scores at T2 and T3 were expressed in rank order.

3. From the experimental hypotheses, and for each CSi, it was noted which score (T2 or T3) should be higher. Thus if the hypothesis were "that subjects high on CSi will perform better on project work", the scores were recorded as follows:

CSi	Column 1		Column 2
Above median, Group 1	T3	$<$	T2*
Below median, Group 1	T2	$<$	T3
Above median, Group 2	T2	$<$	T3
Below median, Group 2	T3	$<$	T2

With this tabulation, the numbers in Column 1 should be lower than the numbers in Column 2, if the experimental hypothesis is to be supported. The hypothesis was tested by means of the Wilcoxon Matched-Pairs Signed-Ranks test (Siegel, 1956). This test was repeated on each combination of predictor variable and criterion variable, with the exception that UoO, was not used to predict performance on UoO2 and UoO3. The results from these tests are shown in Table 7:6. As can be seen, only three hypotheses were supported at statistically significant levels: individuals who were above the median on field independence attained lower order objectives

---

* T3 <u>rank</u> is lower than T2 <u>rank</u> i.e. subjects have a higher position in the class at T3.

($p \leqslant 0.05$) and total objectives ($p \leqslant 0.01$) better and showed a greater improvement on the UoO, under project work than under lectures. One other result is worth mentioning, however; there was a trend for authoritarian individuals to perform better on total objectives under project work. This trend was contrary to prediction and was just below the 0.05 level of statistical significance: the maximum value of T for the 0.05 level on a two-tailed test is 17, whereas the observed level was 17.5.

Finally it is necessary to bear in mind that Table 7:6 contains 47 tests, one to two of which should be significant at the 0.05 level, and less than one at the 0.01 level. The corresponding observed figures are 3 and 1, little better than chance.

CSi	LO	TF HO	T	UoO
CFT	p.0.05*	s.t.	p.0.01	p.0.025*
CFT - R/I	n.t.	n.t.	n.t.	n.t.
PCT	n.t.	n.t.	n.t.	n.t.
Mod RCRT	n.t.	n.t.	n.t.	n.t.
Bal. F	s.t.TT	s.t.TT	s.t.TT	n.t.
AH6 V	n.t.	n.t.	n.t.	s.t.
AH6 N+D	n.t.	n.t.	s.t.	s.t.
AH6 T	n.t.	n.t.	s.t.	s.t.TT
UoO1	n.t.	n.t.	n.t.	
Con/Di	n.t.	n.t.	n.t.	n.t.
SCDT	s.t.	s.t.TT	s.t.TT	n.t.
DIQ	s.t.	n.t.	n.t.	s.t.

Table 7:6 Results of the Wilcoxon Matched-Pairs Signed-Ranks tests applied to the change in rank position between project work and lectures for individuals above or below the median on CSi

n.t. = no trend * = $p \leqslant 0.05$

s.t. = small, but non-significant trend ** = $p \leqslant 0.01$

Although the main aim of the experiment was to investigate the prediction of the response to teaching methods, it is also possible to investigate the cognitive style predictions of the response to the whole course. This has now been done, using, as the predictive criteria, the total performance on the TF tests (i.e. TF1 and TF2), and the total change in performance on the Uses on Objects tests (UoO3 - UoO1). These criteria were rank ordered, and tested by means of the Mann Whitney U Test, with subjects above the median and below the median as the two independent groups. The results of these tests are shown in Table 7:7. As before, a one tailed test was used when the trend was as predicted, and a two-tailed test when the trend was contrary to prediction.

	TF1 + TF2			
CS1	LO	HO	T	UoO3-UoO1
CFT	p=0.08 TT	p= 0.14	n.t.	n.t.
CFT - R/I	p=0.08	p= 0.04*	p = 0.04*	n.t.
PCT	n.t.	n.t.	n.t.	n.t.
n1=6, n2=7 Mod RCRT	n.t.	p= 0.051*	p = 0.14	n.t.
Bal.F	P=0.005**	p= 0.07	p = 0.01**	n.t.
AH6 V	n.t.	n.t.	n.t.	n.t.
AH6 T	n.t.	n.t.	n.t.	n.t.
UoO1	n.t.	n.t.	n.t.	
Con/Di	n.t.	n.t.	n.t.	n.t.
n1=7, n2=7 SCDT	n.t.	p= 0.16	p = 0.10	n.t.
DIQ	n.t.	n.t.	n.t.	n.t.

Table 7:7 The results of Mann-Whitney U tests applied to the performance of independent groups (above, and below, the median on CSi) on the criteria of TF performance, and change in UoO performance, over the whole course. Except where otherwise stated, n1 = 8, n2 = 8.
TT = Two-tailed test. Where not stated, a one-tailed test was used.
** = significant at p 0.01 * = significant at p 0.05
n.t. = no trend

It can be seen from Table 7:7 that the prediction from the Bal.F is supported at a particularly impressive level (p=0.005 on lower order objectives; p = 0.01 on total objective). Other tests which gave significant predictions were the CFT - R/I score, and the Mod RCRT. All of these results were as predicted. The scores on field dependence showed a trend towards prediction on LO and HO, but in the case of LO, the trend was contrary to theoretical expectation, and a two-tailed test was used. Neither of these trends was significant. Another non-significant trend worth noting was for individuals high on automatisation (SCDT) to perform better on HO and T. All other results on the TF scores, and all results on the UoO scores were in line with the null hypothesis.

Finally, it should be noted that Table 7:7 involved 47 predictions: by chance, one would expect - 1 - 2 of these to be significant at the 0.05 level, and 0-1 (nearer zero) to be significant at the 0.01 level. The observed figures are respectively 5 and 2.

c) Subsidiary Hypotheses

It is also possible to examine the difference between performance under lectures and performance under project work, irrespective of cognitive styles. From the analysis of the objectives of these two teaching methods (above) it would be hypothesised that lectures are primarily aimed at the lower order objectives involved in communicating a large amount of structured information in a short time. Conversely, it is claimed that the main advantage of project work is that it increases the student's motivation, and involvement with the information, and is likely to improve his comprehension, and ability to use subject matter, albeit at the expense of a lesser breadth of knowledge. These hypotheses can be examined by comparing the levels of performance after lectures and after project work. Thus TF2 (LO) scores should be higher than TF1 (LO) scores for Group 2, but lower for Group 1. The converse relationship should hold for HO scores, and a two-tailed test is required for TF - T scores. Table 7:8 shows the results when these predictions are compared using the Sign Test.

TF2 - TF1 (Group 1)

+ TF1 - TF2 (Group 2)

LO	p = 0.059	n = 15[1]	x = 4
HO	p = 0.402	n - 16	x = 9
T	p = 0.804(TT)	n = 16	x = 9

Table 7:8 Comparison of performance under project work compared with performance under lecturing, by means of the Sign Test

n = number of observations
x = number of observations where performance in project work is <u>higher</u> than performance in lectures.
p = probability of a random result as extreme as the observed result
TT = two-tailed test. Where not specified, a one-tailed test was used
1 = One subject showed no change, thus reducing the sample size to 15.

The table shows that the trend for performance on lower order objectives to be better under lectures very nearly achieves statistical significance (p = 0.059). There was, however, no observable result for higher order objectives or for total objectives.

d) Other Variables.

Finally, Tables 7:9 and 7:10 and 7:11 show the means and standard deviations of the student ratings of lectures and project work parts of the course. Table 7:9 (Lecture Evaluation) is based on 10 subjects from Group 1 who completed the rating scales at T2. Table 7:10 is based on 9 subjects (Group 2), who completed the Lecture Evaluation at T3. The nine subjects include the six whose cognitive style data and TF data were complete, and 3 others whose TF data was incomplete. Table 7:11 (Project Evaluation) is based on Group 1 (10 subjects) at T3.

The bearing of these tables on the external validity of the experiment is discussed below. It is sufficient, at this point, to note that Group 2, were more variable in their ratings than Group 1, and had a generally lower opinion of the lecture course. This serves to emphasise the fact that the groups were different (either because of different background, or because the lecturer treated them differently in some way), and that it would be

Evaluation of Lectures. Group n = 10

	6 - 1	m	s
Vocational relevance	good - bad	4.10	0.83
Interest	good - bad	5.30	0.64
Clarity of exposition	good - bad	4.60	0.92
Organisation of information	good - bad	5.20	0.60
visual presentation	good - bad	4.70	0.64
Delivery - audibility	good - bad	5.40	0.66
- speed	good - bad	4.80	1.08
Rapport with class	good - bad	4.80	0.98
Amount learned	a lot - nothing	4.70	0.46
Problem solving	improved - un-changed	2.90	0.94
Subject	easy - dif-ficult	4.30	1.19

Table 7:9 Evaluation of lectures by Group (n=10) at T2. Ratings on a 6-point scale, where 6 corresponds to adjective at the left. Means and standard deviations.

Evaluation of Lectures. Group 2 n = 9

	6 - 1	m	s
Vocational relevance	good - bad	3.88	1.39
Interest	good - bad	3.89	1.29
Clarity of exposition	good - bad	3.89	1.37
Organisation of information	good - bad	3.78	1.23
Visual presentation	good - bad	3.33	1.56
Delivery - audibility	good - bad	5.00	1.05
- speed	good - bad	4.22	1.47
Rapport with class	good - bad	4.22	1.13
Amount learned	a lot - nothing	3.44	1.26
Problem solving	improved - un-changed	2.89	1.20
Subject	easy - dif-ficult	3.78	1.40

Table 7:10 Evaluations of lectures by Group 2 (n=9) at T3 Ratings on a 6-point scale, where 6 corresponds to adjective at left. Means and standard deviations.

Evaluation of Project Work. Group 1 n = 10

	m	s
Interest	5.4	0.49
Memory	4.2	0.60
Comprehension	4.5	0.92
Motivation to work	4.8	0.60
Ability to use the information outside the course	4.8	0.75
Inter-relatedness of different parts of the course	4.5	1.02

Students asked to rate whether <u>this particular project</u> increases ------ decreases (6 ------ 1) the above qualities, as compared with the <u>typical</u> lecture course.

Table 7:11 Evaluation of project work by Group 1 (n=10) at T3
Ratings on a 6-point scale. Means and standard deviations.

preferable, if there was a choice, to analyse their data independently. A further comment is in order: in view of the apparent uniformity of the means of different rating scales, one might be tempted to think that an acquiescence factor was operating, or that the students were trying not to hurt the lecturer's feelings. The low ratings on the change in problem-solving ability, however, argue against an acquiescence factor, and the presentation of the rating scales to the students (which stresses honesty, and involved total anonymity) argues against the second possibility.

7:6 Discussion

The findings with regard to the main experimental hypotheses are very few. Firstly, Table 7:6 showed that only the CFT predicted relative performance on teaching methods at a statistically significant level, and that it did so on three criteria - lower order objectives, total objectives and change in Uses of Objects performance. Although one would expect almost as many significant findings in a matrix this size by change alone, the writer is inclined to accept these findings: they are, after all, hardly counter-intuitive. Of all the cognitive style dimensions reviewed in

Chapters 1 and 2, field dependence is probably the best researched, and has the greatest demonstrated reliability and generality. Field dependence is clearly implicated at a theoretical level with academic performance, and it is not surprising that this has been confirmed empirically. It is perhaps surprising that the difference between the teaching methods in fulfilling total educational objectives is significant at the p 0.01. On a sample size of only 16 students, this represents a trend which, if replicated on larger samples, would give a massive level of significance. This result at least indicated that field dependence is a variable which ought to be investigated in the educational situation.

It is interesting, also, that field dependence was found to predict the improvement in Uses of Objects scores at a significant level (p 0.025). Since Uses of Objects scores are independent from Tf - Total (unlike TF - LO and TF - HO, which are mathematically related), this adds confidence that the results are not a chance variation in the matrix. It is also interesting in that field dependence has been related both theoretically and empirically to 'creativity' (see Chapter 1): here is it shown to be related to 'induced creativity' as a function of teaching method.

The finding that authoritarian individuals tended towards better performance on project work is particularly curious in the context of the field-dependence findings. Normally, authoritarianism is thought to correlate negatively with field dependence scores, and consequently authoritarian individuals should perform better in lectures. However, this trend did not attain statistical significance, and it is probably safe to dismiss the result as error variation.

The most impressive result from the Table 7:7 is the prediction from the measure of authoritarianism, the Bal.F. More authoritarian individuals performed worse on the course as a whole: their performance was worse on all educational objectives (p = 0.01), but the main burden of this was carried by their lower performance on lower-order objectives (p = 0.005): a particularly impressive result on a sample as small as this. In view of these findings, it is probably acceptable to consider the poorer performance on higher order objectives (p = 0.07) as also significant. These results are in line with theoretical predictions, but not entirely: one can predict that authoritarian individuals, with their simple and more highly stereotyped cognitive systems, are more resistent to the intake and integration of novel information, but one might expect that their greater need for stucture and intolerance of ambiguity would make their performance poorer particularly on the higher order cognitive objectives. The higher objectives require a greater integration of the information into the pre-existing cognitive structures. However, the observed result is that the poorer performance comes largely from the lower order objectives. This apparent contradiction may simply be a result of sampling error: another possibility is that the performance on the higher-order objectives was so poor that differences between individuals high and low on authoritarianism are not perceptible.

Other results in Table 7:7, however, militate against this latter possibility. The Mod RCRT predicts performance on higher order objectives (p = 0.05) but not lower order objectives or total objectives. The result here was as predicted: individuals high on cognitive complexity perform better than individuals low on cognitive complexity. Similarly with the CFT, there is a trend for field independent individuals to perform better on higher order objectives, and <u>worse</u> on lower order objects than field dependent subjects. Using a one-tailed test on the former result, and a two-tailed test on the latter, neither trend is significant (respectively, p = 0.14, and p = 0.08). However, it is possible to argue that field independent subjects would do worse on lower order objectives: it could be that they are stimulated by tasks which involve the integration and analysis of new information, and other higher order cognitive functioning, but become bored with the lower level tasks of simple inferences, etc., involved in the attainment of lower order objectives. If this argument is acceptable, and a one-tailed test applied, the trend towards lower performance of field independent subjects on lower order objectives is significant ($p = 0.04$). However, the argument is tenuous, and it is perhaps symptomatic of the vagueness of such concepts as cognitive complexity that such tenuous arguments are possible. The only safe conclusion which can be drawn is that the possibility requires further investigation.

Much to the experimenter's surprise, the rather crude Reflection/Impulsivity score which was derived from the CFT was also found to predict performance on higher order objectives and performance on total objectives ($p = 0.04$) in both cases). Here, the result was in line with theoretical predictions: reflective individuals performed better than impulsive individuals. If this is observed to be the case in the present structured situation, where all of the course work was supervised, it can be hypothesised to be even more important in Departmental course work, where the higher contribution of self-initiated study, library work, and essay writing, would magnify the difference between reflectives and impulsives. As to the finding that the difference in performance is to greater extent carried by higher order objectives, and to a lesser extent by lower order objectives ($p = 0.08$), this is within the range of sampling error. An alternative possibility is that the reflection/impulsivity dimension interacts with the automatisation dimension, such that impulsive individuals have poorer scores only on tasks which are less automatised (i.e. require a greater conscious concentration, and a higher level of 'limited-capacity' processing): this explanation gains supporting evidence from the observed trend towards highly automatised individuals performing better on higher order objectives and total objectives than less automatised individuals. Unfortunately, these trends are not significant (respectively, $p = 0.16$, and $p = 0.10$), and thus the explanation is tenuous in the extreme.

Table 7:8 provides some evidence that lectures are more efficient than project work in promoting lower order educational objectives. In view of the inherent plausibility of this proposition, it is not unreasonable to accept it and to reject the null hypothesis, even though the results only reach the level of $p = 0.06$. It is worth remembering that this conclusion is based on the usage of a Sign test, which is based on nominal data and is consequently not very powerful. In the present case it was not possible to use a more powerful test, as the scores TF2 - TF1 involved a considerable number of ties. However, if the same result had been found with a more differentiated criterion, one could safely assume that they would have reached the conventional levels of statistical significance.

The difficulty with performing experiments in the educational situation is that there is almost certainly a loss of external validity. Even accepting all the conclusions outlined above, there is a distinct possibility that they will not generalise beyond the present experimental situation. To begin with, the content of the course was most unusual; it was explicitly concerned with presenting information which would improve the students' abilities to solve problems. Such information is marginally related to the information presented in Psychology departments, but the selection and organisation of the information was completely different. It is, of course, even less like the content of other departmental courses.

Secondly, the subjects were Complementary Studies students. The Complementary Studies programme is compulsory, but is not assessed for the Degree (that is to say, it is formally assessed, but in practice a student is failed only if he consistently refuses to attend). As general studies, the work is entirely in University time, in contrast to the usual departmental course, where students are expected to follow up references, write essays, etc., in their own time. Furthermore, Complementary Studies, is intended to add an element of general education: some students, however, have a resistance to general education which is independent of their performance on their own chosen course of study. This resistance is perhaps understandable considering the importance of the Degree result, which is only minimally based on general studies work. All of these factors suggest that the present experimental results will not generalise to the usual Departmental Studies, although it may be possible to predict some of the differences, as has been indicated above.

The present experiment has attempted to use two different teaching methods, but to some extent it is true to say that there are as many teaching methods as there are teachers. Some of the variations in different lecture methods and different project methods were indicated in the introduction, although these are additional to variations between teachers. While the investigation of such differences is a study in itself, the present experiment has tried to avoid the difficulty by specifying the teaching methods and the course content to a high degree. Furthermore, the assessment of the lecturer and of the project work by the students provides some evidence that the lecturer did not differ from the norm in any great degree. Tables 7:9, 7:10, and 7:11 presented these results. To take Group 1's ratings of lectures first (Table 7:9), all of the ratings bar one were above the midpoint of the 6-point scale. Mean Ratings of over 5 were given for "interest", "organisation of information", and "delivery-audibility". All the other ratings were between 4 and 5, with the exception of "problem solving", which was just under 3 (marginally nearer the "unchanged" end of the scale). Besides being extremely gratifying to an inexperienced lecturer, these ratings show the lecturer was no worse than, and possibly slightly better than, the norm, at least on the qualities rated. Table 7:10 contains the lecture ratings for Group 2: these were in all cases lower than the comparable ratings from Group 1, but they carry the same message.

All mean ratings, with the exception of "problem solving" were above the midpoint of the scale, with "delivery-audibility" at a mean of 5.0, and "delivery speed" and "rapport" at just over 4. The other ratings were between 3 and 4. It is interesting to note that Group 2 was more variable in its assessment of the lectures than Group 1, although one cannot know why.

The project evaluation (Table 7:11) served a different purpose: this type of project-work is extremely rare in academic, as opposed to practical, subjects. It was essentially a thinking exercise, concerned with the collection and collation of information. It would almost be characterised as a group essay, where the only end product was an essay plan. Consequently the Project Evaluation was aiming to investigate how effective the students thought this process to be. It asked the students to compare the project work with the typical lecture on a number of qualities. The mean ratings were above the mid-point of the scale on every quality. The highest mean rating was on 'interest' (5.4), with a "motivation to work" and "ability to use the information outside the course" joint second (4.8). These results are interesting, but perhaps not very reliable. In the first place, the students' perceptions that this project was better than lectures at improving "memory" and "comprehension" was not borne out by the above result that lectures are better on the lower order educational objectives. Similarly, the lack of any difference on higher order objectives does not confirm student assessment of "ability to use information outside the course" and "inter-relatedness of different parts of the course". Finally, although it is probable that "interest" and "motivation to work" were improved, this <u>may</u> be part of a Hawthorne effect.

Although the external validity of the experiment is low, the internal validity is high, at least, relative to what is normally possible in educational research. The few findings which the experiment has reported are relatively strong ones: what is disappointing is that more of the theoretical predictions were not confirmed. There are a number of possible reasons for this:

1. The sample was too small. This is a factor which has overshadowed the data analysis and results of the whole experiment. It has necessitated the summation of groups which should not have been combined, and has almost certainly masked the appearance of all but the strongest trends in the data. In one sense, this is no bad thing: a finding which only emerges at the 1% level of significance on 30-40 subjects is scarcely worth making. Even if it is of slight theoretical use, it is of no practical value at all.

2. The criterion variables were inadequate. The difficulties in the construction of the TF tests, and the low level of student performance, have already been discussed. They must certainly be a factor in the paucity of results. It is almost inconceivable that the students were as little affected by the course as the TF scores suggest, and it is likely that criterion tests which were more sensitive to the changes that were taking place, would have been more predictable from cognitive style dimensions. It is possible that the results which did emerge, did so in spite of these two factors, indicating that the relationships are particularly strong ones.

3. The experimental hypotheses were incorrect. Chapters 2 and 3 reviewed the cognitive style variables used and concluded that they had a good a priori predictive potential in education. Despite this, it is likely that in some cases the links with educational processes are tenuous or minimal, and that they would only emerge at a low level on large samples. To this extent it is likely that at least some of the experimental hypotheses are incorrect, and would not be supported even on a larger sample

and with more sensitive criterion measures.

Furthermore, the cognitive style literature abounds with empirical results which are apparently contradictory, and with related theoretical dimensions, the empirical measures of which do not correlate. This is usually the symptom of a "young" research area, in which the explanatory concepts need considerable empirical 'sharpening', and in which inter-relationships need realigning. Some of the cognitive style dimensions (e.g. cognitive complexity and concreteness vs abstractness, and, some would claim, field dependence) are vague and global in the extreme: when applied to a behavioural criterion as complex as educational performance, it is literally possible to rationalise almost any observed finding. Thus whilst there may be a grain of truth in the prediction from these dimensions to the educational situation, there are almost certain to be compensating 'masking' variables in the dimensions which, effectively prevent the emergence of empirical relationships. In this case, one can say that some of the experimental hypotheses are incorrect in that the conceptualisation and measurement of the predictor variables is "eccentric" (literally).

There is the further possibility that the failure of the cognitive style dimensions to predict lies not so much in their "eccentricity" of abstraction as in the dimensional nature of the abstraction. As Chapters 4 and 5 argued, it is possible that a dimension, which is "measured" at a given point in the life context and is assumed to generalise across all other life contexts, and on which every individual is capable of being represented - such a concept may be incapable of predicting a complex criterion in the educational situation.

There is some indication of this: the course was aimed at improving creative problem solving ability and ideational flexibility, and almost all of the subjects obtained higher scores on the Uses of Objects tests at the end of the course than they did at the beginning. Had the students of the course been a sub-sample of a larger group, the course would have succeeded in altering the students' observed cognitive style: there would have been a shift towards divergent thinking.

The same point applies to the use of the experiment as an investigative tool: it is possible that the causal relationships between cognitive functioning and educational performance are incapable of being forced into the experimental format. The use of the single predictor, and the summation of scores across individuals, may scarcely begin to take account of the variance in the observations. For instance, to follow up the point about the increase in Uses of Objects scores, it may very well be that cognitive styles are not "constitutional" dimensions, but that they are responses to specific situations. If this is the case, there is the possibility that an individual student will vary his cognitive style according to the demands of the teaching-learning situation and the subject matter. It may also be the case that education is implicitly aiming at changing the student's cognitive style: certainly the lecturer in the present study would be very pleased if the shift towards divergent thinking styles were real, and no mere measurement error in the convergent thinking/divergent thinking dimension.

If these considerations are a valid description of 'reality', then a non-dimensional, non-normative, non-experimental study of cognitive style and educational performance is required.

CHAPTER 8

FIELD STUDY : COGNITIVE STYLE AND ACADEMIC PERFORMANCE

8 : 1 Introduction, and Aims of the Study

Cognitive styles are usually thought of as normative dimensions of individual differences. Arguments have been presented to suggest that another conceptualisation, which takes into account structures and systems of cognitive content and hence idiosyncracy and interactions between variables, may be more useful in enabling us to understand the cognitive processes underlying academic performance. The present study is in direct contrast to that described in the last chapter. That study was an experiment, with a formal experimental design, relatively high internal validity, and in a highly controlled experimental situation - that is, highly controlled relative to what is usual and what is possible in educational research. The present study is a field study: it is not experimental in that there are no dependent variables. The internal validity of any conclusions is, logically, non-existent. The external validity is correspondingly high: the subjects of the study are students in a normal course in the Department of Education of the University of Aston, and the results can be expected to generalise to other University samples to a greater extent tha n the results of the previous experiment. Whereas the experiment used a number of conventional measures of cognitive style, which consequently imply a theoretical approach, and demand restricted typesof empirical conclusion, the field study samples a fairly wide selection of cognitive elements, and makes less stringent assumptions about the nature and structuring of these elements. The types of information which can come out of the study are consequently wider. Whereas the experiment was hypothetico-deductive, involving predictions from cognitive style and teaching methods to academic performance, and field study is largely inductive: it is investigating a number of aspects of the process of education, attempting to see whether structuring or systems of cognitive indices will act as "symptoms" of academic performance. The field study is therefore taxonomic, with the aim of describing and understanding the process at the time, rather than predicting it from ahead. It would be nice to present the two studies as a comparison of two research approaches and two theoretical approaches, in the best traditions of scientific confrontation. But they are not. Even ignoring differences in the aims of the two studies, and differences in the content, the experiment had certain defects which could be held to account for any paucity of results, and the field study is a priori unlikely to reach any substantive conclusions. With such a wide open and vague task, with no pretence of control and an almost infinite number of potential variables, it would be surprising indeed if the study showed that x, y, z, could be held to account for academic performance. The most that can be hoped for is the indication of problems which will be encountered if this approach is to be followed.

The aim of the present study can therefore be succintly stated as follows:

"To investigate individual differences in the structuring of some cognitive domains which may be hypothesised to underlie academic performance".

8:2 Design

Although not possessing dependent and independent variables, and comparison groups in the conventional sense, the logic of the study can be abstracted. It is presented in summary form in Diagram 8 : 1

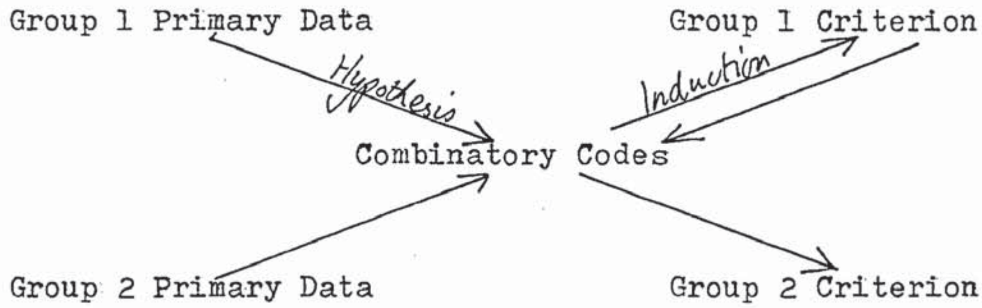


Diagram 8 : 1 Design of the Study.

The diagram shows that the study follows a 4-phase process:

1) Observations of aspects of cognition (hereafter called "primary data") in two independent groups.

2) Observation of criterion scores in two indepent groups.

3) The development of an account of systems, structures or interaction in the primary data of the first group, which will predict the criterion scores of that group. Such structures or systems will hereafter be called "combinatory codes" in that they specify, in operational terms, how the different scores in the primary data are to be combined.

4) The testing of the combinatory codes by applying them to the primary data and criterion scores of the second group.

It is necessary to comment on some aspects of this process.

8 : 2 : 1 Primary Data

Although it is stated that these are observations of aspects of cognition, the phrase "aspects of cognition" is taken in its widest sense. Rather than restrict the types or amount of information taken into account for the sake of theoretical tidyness, it was felt that it would be better to maximise the chances of predicting academic performance by using any information which might be relevant. Thus, one question asked the student to say what preference the University of Aston was on his UCCA form: although this would not normally be classified as an "aspect of cognition" it seemed quite likely that it would indicate something about cognition and about academic performance.

Thus the primary data attempted to sample a wide area of cognitive functioning: the actual amount of information collected, and the "depth" of the sampling of cognition was determined by the numbers of subjects to be investigated, and hence the amount of time and money which could be devoted to each subject.

8 : 2: 2 Criterion Scores

In the present study, these were the examination results at the end of the year. The previous discussion of examination results as research criteria concluded that they were usually assumed to be reliable, and psychologically pure, whereas they are in fact neither of these. The assumption about psychological purity is irrelevant to the present study, which can take into account (in principle at least) the possibility that students achieve a given examination mark in a variety of ways. The assumption of reliability is made by the present study, although it is known to be false: the over-riding consideration is that examination results are of very great importance to students and universities.

8 : 2 : 3 Combinatory Codes

These specify how the scores of students are combined in order to give a prediction of their examination results. A regression equation would be an example of a combinatory code. However, fewer presuppositions are made as to the structure of these codes. For instance it is possible that a high score on one of the primary indices will indicate something about performance, while a low score is irrelevant. Similarly it is possible that a combinatory code will apply to a sub-sample of the students, but not to others (provided, of course, that the sub-sample can be identified from the other indices): this would be an interaction effect between the primary indices. Thus the combinatory codes can take into account interactions, curvilinear relations, and logical relationships which are only minimally picked up by correlation and other normative analyses. Of course, the difficulty lies in knowing which codes are likely to predict to examination results. With any large number of primary indices, it is impossible to test out all of the possible inter-relationships between indices on all combinations of indices. This is the central problem of this type of research: if you do not know what you are looking for, how do you set about searching? This is the paradox of the Meno, although not in its fullest form. The corollary, "and how do you know when you have found it?" does not apply here: that is taken care of in the criterion test. There is no solution to this problem. Any mathematical method of searching for inter-relationships requires the specification of those relationships before the search. However, if there is no systematic search procedure which can be followed, there are a variety of heuristics which may be useful:

1) Intuition. The average student carries a fairly sophisticated model of the processes of education within him, as is witnessed by his ability to predict his own and other people's examination results with reasonable

accuracy. It may be possible to tap this model in hypothesising promising combinatory codes.

2) Induction By examining the scores of students who perform well in comparison with students who score poorly. Cluster analysis can be helpful here, by grouping together these students who have a similar pattern of scores. The task is then to infer a pattern, or code, which will differentiate between clusters.

This stage of the research process resembles the stage of generating possible solutions to a problem, before testing those solutions, and the same qualifications apply. In generating solutions, it is often considered best to "suspend evaluation", so that any possible solution is considered irrespective of its a priori likelihood. In the present case, this procedure would involve the serious consideration of a huge number of possibilities, but the "suspended evaluation" principle can be applied to the inductive process. In inferring codes to describe the clusters of high and low performers, it is possible that observed differences are being "rationalised" i.e. that they are random variations and are not criterial of performance. This is irrelevant at this stage of the process. The combinatory codes are subject to a final test on an independent group, and the methods used to generate the codes are logically independent of this final test. Thus, in the stage of generating combinatory codes, the normally repressed scientist can indulge himself in an orgy of unscientific thinking.

8 : 2 : 4 Testing of Combinatory Codes

In view of the methods of generating the combinatory codes, it is essential that there should be a formal testing phase, on completely independent data. This involves using the combinatory codes to generate a prediction of criterion performance for a different sample of primary data. The testing is by means of the usual normative statistical methods.

8 :2: 5 In addition to this part of the study, a small correlational study was performed, in which two of the conventional tests of cognitive style were used to predict examination results.

8 : 3 Subjects

The main part of the study used 27 subjects, 18 in Group 1 and 9 in Group 2. Group 1 subjects consisted of 15 females and 3 males, all of whom were in the 1st year of the Combined Honours Course, one part of which was Education. The total size of the first year Combined Honours Education group was 28, and 18 subjects of the present study comprised a sub-sample of this group.

Group 2 consisted of 4 females and 5 males; 9 subjects, in all, who were a sub-sample of 11 students in the Second Year Combined Honours Education groups. The selection of the sub-samples in both Groups 1 and 2 was inherent in the recruitment procedure, which involved both persuasion and financial inducement, but not compulsion: the sub-samples were therefore self-selected, and evidence will be presented in the results section to show that they differed from students in the Combined Honours Education Groups who were not included in the study.

There were certain differences between Groups 1 and 2 which should be noted. Firstly, Group 1 subjects were in the First year of the Combined Honours course, which required them to take 3 subject courses, whereas Group 2, in the second year, took only two subject courses. Secondly, Group 1 had a higher proportion of females than Group 2. Thirdly, Group 2 was visibly older and more mature than Group 1. Unfortunately, information with respect to age of subjects, which might corroborate this subjective judgement, was not collected.

The correlational study, which used two cognitive style tests to predict examination performance, was based on some of the above subjects, with a few additions. Group 1 consisted of 12 of the original 18 subjects together with 4 others from the first year Combined Honours Education Group, 16 subjects in all. Group 2 consisted of 7 of the 8 subjects above, together with 2 others, making 9 from the second year Combined Honours Education Group. Since this correlational study was included as a follow-up to the previous experiment, and was not intended to be comparable to the main part of the present study, this disparity between the observation samples is not important.

8 : 4 Measures

The aim with the primary observations was to sample as wide an area of cognition as possible, which might also be relevant to academic performance. The measures therefore attempted to cover the following domains:

1. Perceptions of the subject of study (Education)

   - topics
   - interest in topics and subject area
   - structuring of the subject area
   - importance of topics (with respect to examinations)
   - distribution of time and work on different parts of the subject area.

2. Perception of teaching-learning situations

   - interest in, and enjoyment of........
   - evaluation of the importance of.......
   - subjective criteria of these
   - use of different study habits.

3. Extra-Curricular activities

   - enjoyment of.......
   - contribution to intellectual development
   - subjective criteria of these

4. Social behaviour

   - social satisfaction
   - compatability with, or position in, social groupings.

5. Perturbations

   - existence of stresses of any type which might cause anxiety or disturbance of work.

6. Aims in Life

   - values which are important to the individual
   - level of importance of these values
   - relevance of education to these values

7. Vocation

   - certainty of vocation
   - relevance of education to the vocation.

These domains were sampled, with some slight shortfall from the original aims, by two measures, the Lifestyle Questionnaire (LSQ), and a scheduled interview procedure, called the Educational Environment Perceptions (EEP) interview.

The LSQ is presented in Appendix 2. It presented a number of questions designed to sample aspects of vocation, aims in life, perturbation and social behaviour, the answers to which were usually on a 10-point scale. It yielded the following information and scores:

Names of three preferred vocations, in order of preference.

1. Probable vocation (scored in terms of how high in the list of preferences this came).
2. Certainty of vocation (1 - 10).
3. Relevance of degree to vocation (1 - 10).
4. Relevance of education course to vocation (1 - 10).
5. Marriage ? (1 - 10).
6. Wealth ? (1 - 10).

7. Fame ? (1 - 10).

8. Popularity ? (1 - 10).

9. Service to individuals ? (1 - 10).

10. Service to mankind ? (1 - 10).
Degree relevance to

11. Marriage (1 - 10).

12. Wealth (1 - 10).

13. Fame (1 - 10).

14. Popularity (1 - 10).

15. Service to individuals (1 - 10).

16. Service to mankind (1 - 10).

17. Level of work satisfaction (1 - 10).

18. Preference for Aston on UCCA (1-6, or 7= clearing scheme).

Existence of problem (emotional, financial, social, health, organisational) (1 or 0).

19. Anxiety level (1 - 10).

20. Work disturbance by problem (1 - 10),

Names of three closest friends (on the course).

Names of three most distant acquaintances (on the course).

The information from these latter two were to be used for the construction of sociograms, which would give some measure of the individual's social position in the group. In the event, the failure to collect data from a number of members of the group made it impossible to construct completed sociograms).

Social satisfaction (0 - 3), comprising the questions:
"Is there sufficient opportunity to make friends?" (1 or 0)
"Are your fellow students friendly enough?" (1 or 0).
"Are they too friendly?" (1 or 0).

It was thought that a good a priori case could be made for each one of these scores to be related to examination performance. Also they have some inherent potential for combining into systems of variables. It will be noted that all of these scores were normative.

The Educational Environment Perceptions interview procedure aimed to sample the domains of perception of subject area, perceptions of teaching-learning situations, and perceptions of extra-curricular activities. Although an interview, it was fairly "objective", in the sense that the interviewer followed a schedule which allowed little scope for variation. The main function of the interviewer was to explain the nature of the tasks, to answer any questions, and to give encouragement. The EEP interview followed the prescribed schedule:

1. The subject listed "study activities" on index cards, one activity to each card.
2. When he had finished, he was given a checklist of study activities, and asked to write out any which he had missed.
3. The subject sorted the activities in terms of enjoyment. He did this by placing the cards on an imaginary ten-rung ladder, which went from "least enjoyable" to "most enjoyable".
4. The subject sorted the activities in terms of how much they trained him, to obtain good marks in examinations.
5. The subject listed leasure activities. As before, when he had finished listing them, he was presented with a check-list, and asked to add to his list any that he had overlooked.
6. The subject rated the leisure activities in terms of enjoyment.
7. The subject sorted the activities in terms of how much they contributed to his general development: "by this I mean how much they improve your skills and personality, your development into what you would like to be".

8. The subject listed the two or three subjects he was studying, and stated which he was most and least interested in. He also estimated the percentages of non-compulsory study time spent on the two or three.
9. The subject was asked to imagine himself a newly-appointed lecturer, employed to design and run a Department of Education course. The course should include all the course content, plus anything else which the subject thought relevant. The subject was asked to categorise the total course into 3-6 sub-courses. He was then asked to take each sub-course, and list (one per index card) the topics which it would comprise.
10. The topics were sorted according to interest.
11. The topics were sorted according to their importance in the present syllabus of the Department of Education.
12. The courses were sorted according to the amount of non-compulsory time which the subject had devoted to them.

It was originally intended that the EEP interview should be administered by students from a counselling course, and an "Interviewer's Kit" was made up for the purpose (see Appendix 2). The kit consists of the following:

1) interviewer's manual. This contains an explanation of the nature of the interview, and the interview schedule. This latter contained verbatim instructions to the subjects.
2) Envelope SA (Study activities) containing 30 cards
3) Checklist S (Study activities)
4) Scoresheet S (Study activities)
5) Envelope SQ (Study qualities) containing 6 cards
6) Envelope LA (Leisure activities) containing 30 cards

12

7. Checklist L (Leisure activities)
8. Scoresheet L (Leisure activities)
9. Envelope LQ (Leisure qualities) containing 30 cards
10. Envelope T (Topics) including 30 cards
11. Scoresheet SS (Subject of study)
12. Spare cards envelope (containing 10 cards).

The interview kit allowed for an extended interview procedure, in which subjects were asked to list the qualities of activities or topics which made them more or less enjoyable/interesting/important etc, and then sorted the activities or topics according to these qualities. Unfortunately the interview had to be restricted to one hour, for financial and administrative reasons, and pilot testing showed that these parts had to be cut out. The deletion of these parts of the interview procedure also necessitated minor changes in the order of presentation of various tasks: these changes will be apparent if the interviewer's Manual is compared with the description of the actual procedure used (above). Furthermore, again for administrative reasons, it proved impossible to use the counselling students as interviewers, and the writer conducted all of the interviews with an amended interviewer's kit. The full kit is presented here as it is sufficiently general to be used on different samples of students, with only minor modifications.

In addition to the idiosyncratic information included in the various listings, but regrettably not extended by the elicitation of constructs of judgement, the EEP yielded the following normative scores:

21) Study activities enjoyment/exam benefit compatibility

22) Leisure activities enjoyment/general benefit compatibility

23) Topic interest/syllabus importance compatibility

24) Topic interest/topic time compatibility

25) Topic time/syllabus importance compatibility

(the compatibility scores were Pearson product-moment correlation co-efficients)

Work bias (proportion of voluntary study time spent on education)

It is worth reporting some intuitive observations of how these two measures would work out in practice. Firstly, there were no problems with either the administration or the scoring of the LSQ, except of course that the sociogram questions are of very little use unless the whole of a group is sampled. The EEP interview worked fairly well, with the following provisoes:

1. It may be better to ask the subject to sort the cards into rank order, rather than onto a 10-rung ladder. Most subjects did this anyway, and then collapsed the ranks into the 10 required categories. It seems that rank ordering is an activity with considerable intuitive validity, and such scores would allow the use of a rank-order correlation co-efficient, which is almost certainly more appropriate to the data. The use of a ten-rung ladder, inevitably meant that there were too many tied ranks to allow the use of a rank-order correlation coefficient.

2. It would be wise to specify even the shortened interview as $1^{1}/_{4}$ hours or 1.1/2 hours in length. The one hour timetable means that there is insufficient time for students who are slow at listing, or for encouraging and persuading those students who show a reluctance to list more than a very few items. The one hour timetable also allows insufficient time for a thorough study

of the subject area, and its differentiation into topics.

2. The listing of topics is not in a form which allows any appreciation of the subjects' comprehension of the subject area, or of his perception of the logical structure of the area. Different forms of procedure are required in order to elicit information of this type. One possibility is to ask students to draw a "logical tree" of the area.

In addition to the LSQ and the EEP interview, all subjects completed the Wren Study Habits Inventory. The score from this measure was treated as if it were a subsidiary score in the LSQ.

8:4:2 Conventional Cognitive Style Measures

The Closure Flexibility (Concealed Figures) Test, and the Balanced F Scale, as used in the previous experiment, were administered to related samples of students in the combined honours Education groups. The intention was to follow up the findings of the experiment, that scores from these measures predicted overall course performance, by using them as predictors of examination performance in a more typical setting. As in the previous experiment, a score which may indicate the degree of reflection vs impulsivity was also derived from the CFT.

8:4:3 Criterion Measure

The criterion used was the end of year grading in Education. This was the average of the overall assessments for Psychology, Philosophy, and Sociology. Each of these scores was based on a final examination and two essays which were performed during the year. The examination was given a 50 percent weighting, and each of the essays 25 percent. The overall 'examination' mark was therefore based 50 percent on essay marks and 50 percent on written examination, and 33⅓ percent on each of Psychology, Philosophy, and Sociology. These marks were "raw scores" i.e. they were in the same state as prior to the examiners' conference.

8:5 Procedure

The investigator attended one of the usual Education lectures for each of the groups. A brief explanation was given as to the nature of the investigation, and LSQ's were given out. Appointments were made with a few of the subjects for interviews, and the payment rate (50p for the interview) was mentioned.

Further appointments for interviews were made at ensuing lectures, when the investigator also distributed LSQ's to students who had previously been absent.

The interviews were conducted according to the schedule described above, and were terminated with the payment of the fee. All of the interviews took place in the second half of the Spring Term, 1973.

Towards the end of the Spring Term, the investigator took over a lecture for each group, when the cognitive style measures and the Study Habits Inventory were administered. Any student who was absent from this lecture, but who had taken part in the EEP interview, was given a copy of the Study Habit Inventory, and asked to return it in their own time.

Finally the examination results were those which appeared in June, 1973, based on examinations set in June 1973, and essays written throughout the 1972/3 sessions.

8:6 Analysis and Results

8:6:1 Sampling

It was stated above that the subjects were recruited by means of persuasion and financial inducement, delivered at the end of scheduled lectures. This procedure allows a measure of self-selection in two ways : firstly, students who habitually do not attend lectures are less likely to be recruited, and secondly, students who are more resistant to persuasion or financial inducement are less likely to be recruited. One could make a plausible connection between either of these factors and academic performance, and there is a certain amount of evidence in the examination results to support these connections. For instance, in Group 2, the two students who were not included in the study obtained the bottom two examination results ($p \leqslant 0.025$, Mann-Whitney U Test, one-tailed test). In Group 1, the students not included in the study also performed worse, at a level which approached, but did not reach the 0.05 level of significance. (Mann Whitney U test; Observed U = 42, U = 41 at $p \leqslant 0.05$).

Thus the sampling was unrepresentative in a very important respect. This is obviously a factor which needs to be taken into account in any study of academic performance where it is not feasible to make participation in the study compulsory. It may also have implications for the previous experiment, where there was a particularly high experimental mortality rate.

It was also intuitively observable that Group 2 differed from Group 1 in being more mature, and hence they were likely to be more involved with their subject of study. This was confirmed in the examination results, where Group 2 performed better ($p \leqslant 0.01$, Mann Whitney U Test, one-tailed test), although it is possible that the difference is a result of the incomparability of any psychological variable.

8:6:2 Primary Data

The distributions of scores on the LSQ and the EEP interview are presented in tables 8:1, 8:2, and 8:3. These show, respectively, the scores for Group 1, Group 2, and Groups 1 and 2.

A brief glance at the Tables will show that many of the variables are only differentiated to all small extent, and are discontinuous, and some variables are distinctly skewed. Strictly speaking, therefore, the data is not suitable for analysis using parametric statistics.

As mentioned above, the information on most, and least, preferred

SCORE

VARIABLES	0	1	2	3	4	5	6	7	8	9	10	Median
1		0	3	1	14							4
2		2	1	1	1	0	4	3	2	2	2	6·5
3		1	0	0	3	5	1	2	2	3	1	5·5
4		3	2	0	1	0	4	2	1	2	3	6
5		0	1	1	0	2	0	0	4	3	7	9
6		2	0	1	3	3	3	2	4	0	0	5·5
7		12	1	2	0	2	1	0	0	0	0	1
8		1	0	1	1	2	0	5	2	3	3	7
9		0	1	1	1	4	3	2	2	2	2	6
10		1	3	1	0	2	4	1	3	1	2	6
11		11	3	2	1	0	0	1	0	0	0	1
12		3	2	3	2	2	2	3	0	1	0	4
13		13	3	1	0	1	0	0	0	0	0	1
14		15	2	0	0	0	0	1	0	0	0	1
15		10	2	0	0	4	0	0	0	1	1	1
16		10	2	1	0	1	2	1	0	1	0	1
17		2	0	1	0	4	4	1	5	1	0	6
18		1	0	0	2	0	5	10				7
19	8	1	1	4	2	1	1	0	0	0	0	1·5
20	8	2	1	2	2	1	1	0	1	0	0	1
	← NEGATIVE					POSITIVE →						
	1.0–0·8	0·79–0·6	0·59–0·4	0·39–0·2	0·19–0	0·01–0·2	0·21–0·4	0·41–0·6	0·61–0·8	0·81–1.0		
21	1	2	1	2	4	6	2	0	0	0		–·05
22	0	0	0	0	4	1	1	5	5	2		·54
23	1	1	1	1	4	6	3	0	1	0		·03
24	0	0	0	1	4	2	3	3	4	1		·35
25	0	1	0	1	2	0	4	1	4	5		·61
	0–10	11–20	21–30	31–40	41–50	51–60	61–70	71–80	81–90	91–100		
26	1	1	5	9	1	0	0	0	1	0		35
	← NEGATIVE			POSITIVE →								
	60–41	40–21	20–1	0–19	20–39	40–59	60–79	80–99				
27	1	5	5	2	2	2	1	0				–18
	25–29	30–34	35–39	40–44	45–49	50–54	55–59	60–64	65–69			
28	1	0	0	0	1	3	7	5	1			57

TABLE 8:1. FREQUENCY DISTRIBUTIONS OF LSQ AND EEP SCORES. GROUP 1. N = 18.

SCORE

VARIABLES	0	1	2	3	4	5	6	7	8	9	10	Median
1		0	1	3	5							4
2		0	0	0	2	0	0	0	2	3	2	9
3		0	1	1	0	3	3	0	1	0	0	5
4		0	0	1	1	0	1	4	1	1	0	7
5		0	0	0	0	0	0	1	0	6	2	9
6		1	1	2	0	0	1	1	0	2	1	6
7		7	0	0	0	1	0	1	0	0	0	1
8		0	1	2	1	0	0	2	1	0	2	7
9		0	0	0	0	1	0	4	1	2	1	7
10		1	0	1	0	0	2	4	1	0	0	7
11		7	2	0	0	0	0	0	0	0	0	1
12		1	1	0	0	2	3	0	0	2	0	6
13		6	1	0	0	0	0	0	0	1	1	1
14		8	0	0	1	0	0	0	0	0	0	1
15		4	1	1	1	1	0	0	0	1	0	2
16		4	2	2	0	1	0	0	0	0	0	2
17		1	1	2	0	2	2	0	1	0	0	5
18		3	1	0	0	1	2	2				5
19	3	2	1	0	1	0	0	1	1	0	0	1
20	3	0	1	0	1	3	0	0	1	0	0	4
← NEGATIVE POSITIVE →	1.0–0.8	0.79–0.6	0.59–0.4	0.39–0.2	0.19–0	0.01–0.2	0.21–0.4	0.41–0.6	0.61–0.8	0.81–1.0		
21	1	0	1	4	1	1	0	1	0	0		−.26
22	0	0	0	0	:	1	1	3	2	1		.49
23	0	0	1	0	2	0	3	0	2	1		.26
24	0	0	0	0	0	3	3	1	1	1		.26
25	0	0	0	0	1	0	1	4	2	1		.48
	0–10	11–20	21–30	31–40	41–50	51–60	61–70	71–80	81–90	91–100		
26	0	0	2	3	1	3	0	0	0	0		40
← NEGATIVE POSITIVE →	60–41	40–21	20–1	0–19	20–39	40–59	60–79	80–99				
27	0	3	3	1	0	1	0	1				−17
	25–29	30–34	35–39	40–44	45–49	50–54	55–59	60–64	65–69			
28	0	0	0	0	0	0	0	7	2			61

TABLE 8:2. FREQUENCY DISTRIBUTIONS OF LSQ AND EEP SCORES. GROUP 2. N = 9.

SCORE

VARIABLES	0	1	2	3	4	5	6	7	8	9	10	Median
1		0	4	4	19							4
2		2	1	1	3	0	4	3	4	5	4	7
3		1	1	1	3	8	4	2	3	3	1	5
4		3	2	1	2	0	5	6	2	3	3	7
5		0	1	1	0	2	0	1	4	9	9	9
6		3	1	3	3	3	4	3	4	2	1	6
7		19	1	2	0	3	1	1	0	0	0	1
8		1	1	3	2	2	0	7	3	3	5	7
9		0	1	1	1	5	3	6	3	4	3	7
10		2	3	2	0	2	6	5	4	1	2	6
11		18	5	2	1	0	0	1	0	0	0	1
12		4	3	3	2	4	5	3	0	3	0	5
13		19	4	1	0	1	0	0	0	1	1	1
14		23	2	0	1	0	0	1	0	0	0	1
15		14	3	1	1	5	0	0	0	2	1	1
16		14	4	3	0	2	2	1	0	1	0	1
17		3	1	3	0	6	6	1	6	1	0	6
18		4	1	0	2	1	7	12				6
19	11	3	2	4	3	1	1	1	1	0	0	1
20	11	2	2	2	3	4	1	0	2	0	0	2
	← NEGATIVE POSITIVE →											
	1.0 — 0.8	0.79 — 0.6	0.59 — 0.4	0.39 — 0.2	0.19 — 0	0.01 — 0.2	0.21 — 0.4	0.41 — 0.6	0.61 — 0.8	0.81 — 1.0		
21	2	2	2	6	5	7	2	1	0	0		-.16
22	0	0	0	0	5	2	2	8	7	3		0.52
23	1	1	2	1	6	6	6	0	3	1		0.05
24	0	0	0	1	4	5	6	4	5	2		0.34
25	0	1	0	1	3	0	5	5	6	6		0.51
	0-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100		
26	1	1	7	12	2	3	0	0	1	0		.40
	← NEGATIVE POSITIVE →											
	60-41	41-21	20-1	0-19	20-39	40-59	60-79	80-99				
27	1	8	8	3	2	3	1	1				-18
	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69			
28	1	0	0	0	1	3	7	12	3			61

TABLE 8:3. FREQUENCY DISTRIBUTIONS OF LSQ AND EEP SCORES. GROUPS 1 AND 2 COMBINED. N = 27.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1																												
2	41*																											
3	39*																											
4																												
5			-27	43*																								
6																												
7					24	50**																						
8						48**	30																					
9		39*	-48**	29	52**	26																						
10			-27	37*		-36*		-40*	62**																			
11		-43*		-27	-35		-21		-24																			
12						26	44*			-28																		
13		29				35	42*		20			55**																
14				-42*	-58**				-25		49**																	
15	-25	-56**									49**	26																
16		-21	49**	-24					-23		23	30		29	71**													
17								26		22	42*			-25														
18					-27	-37*	-65**			27	-50**	-62**																
19				41*								-40*				-22		27										
20				30								-22	21		-28	-25	-21		81**									
21	36*	25	23																									
22					24				29	39*			-33		22		21			-40*	25							
23			-26						-25	-21	26			21			-44*			21								
24	24									-23			-22	23	30	48**				-33			34					
25			-29		-29	24		-23			31	-25			-21	-30	-29			22			55**	-23				
26		28	-25	-22																20	24		41*					
27											-31			-40*			21		-28	-40*				20				
28			-36*			-21	-21	-28	24	34									-25		-26		24					

TABLE 8:4. INTERCORRELATION MATRIX OF PRIMARY VARIABLES. SUMMARY. GROUP 1 + GROUP 2. N = 27

$r > 0.36$, $p \leqslant 0.05$ = *

$r > 0.47$, $p \leqslant 0.01$ = **

	*	**
BY CHANCE, EXPECTED FREQUENCY =	19	4
OBSERVED FREQUENCY =	36	17

acquaintances could not be used in the construction of sociograms, and has been excluded from the analyses.

The medians of the scores are also presented in the tables. It will be seen that Groups 1 and 2 differ only slightly. Excluding examination results which have already been mentioned, there are only 3 scores on which the medians differ by more than two points, the Group 2 medians on "vocational certainty" and on "degree relevance to wealth" are high, and Group 1 median on "preference for Aston" is higher (indicating the Group 1 students were more likely to have preferred other Universities or to have come in on the UCCA clearing scheme). These differences are in the direction of supporting the intuitive observation of the two groups, that Group 2 was older and more goal-directed.

Table 8:4 presents a summary of the inter-correlations matrix of primary scores, based on Pearson Product Moment correlation coefficients. As mentioned above, the data is not really suitable for parametric statistics, but the large number of tied scores also makes it difficult to use a rank correlation coefficient. Thus, as this is a preliminary stage of the analysis, designed to be informative rather than to reach definite conclusions, it was felt that a parametric correlation coefficient would be acceptable. The deficiencies of the data should, however, be borne in mind when examining the result.

An examination first of the correlations with Examination performance (variable 28) reveals only one which is significant: Vocational Certainty is negatively correlated with performance ($p \leqslant 0.05$), which is to say that those students who have a clear idea of what they want to do for a career tend to obtain lower examination results. This is a most surprising finding, completely contrary to hypothesis, and for this reason it should be accepted only with caution.

Moving further up the Table, "Study Habits" is found to correlate negatively with "Degree relevance to Popularity" and work Disturbance" ($p \leqslant 0.05$ for both). The latter result is understandable, but the former seems arbitary, and particularly suspect in view of the heavily skewed distribution for that variable. Variable 23 (Topic preference/ Topic importance compatibility) is correlated significantly with Variable 26 (work bias) ($p \leqslant 0.05$) : this also is understandable - the more the students' interests fit in with the Department of Education's emphasis, the more the time devoted to Education relative to other courses. Variable 23 is also correlated with Variable 25 (Topic time/topic importance compatibility) ($p \leqslant 0.05$).

Variable 24 (Topic preferences/topic time compatibility)(approaching $p = 0.05$) and Variable 17 (work satisfaction) ($p \leqslant 0.05$). The former two correlations would appear to be boosted by artefact, in that they are based on correlations between three sortings of the same cards, correlations which are then used as raw data. The latter correlation, between "Topic preferences/topic importance compatibility" and "Work satisfaction", is again, according to hypothesis.

Variable 24 (Topic preference/topic time compatibility) is correlated at the 0.01 level with Variable 16 (degree relevance to service to mankind) a result which is particularly difficult to rationalise.

Variable 22 (Leisure activity preferences/Leisure activity contributions to general benefit compatibility) is seen to be correlated negatively with "work disturbance" ($p \leqslant 0.05$) and positively with Variable 10 (importance of service to mankind) ($p \leqslant 0.05$). Variable 22 was intended to approximate a discrepancy between perceived self and ideal self: where there is little discrepancy between perceived and ideal self, the individual will prefer and enjoy those leisure activities which contribute to his preferred self-image. If this interpretation is not too fanciful, then both a high level of work disturbance and a low level of interest in serving man-in-general, might be symptoms of a less integrated "self". However, with the present data, these interpretations must be viewed as interesting ideas rather than as an empirical result.

Variable 21 is found to correlate with Variable 1 (p 0.05), but this can almost certainly be attributed to the lack of differentiation in the latter (the total scale is only 1 - 4).

Variables 19 and 20 (Anxiety level and work disturbance) not surprisingly correlate at an extremely high level (0.81), and Variable 19 also correlates significantly with variables 12 and 4 (degree relevance to wealth) and (course vocation relevance) respectively: in the former case the correlation is negative. It is difficult to see any rationale for these.

Variable 18 is found to correlate negatively at the 0.01 level with variable 7, 11 and 12. These are all a result of the lack of differentiation in variable 18, together with the extreme skewness of the other variables.

There are a considerable number of significant inter-correlations amongst the lifestyle values, and the ratings of the importance of the Degree to those values (variables 5 - 16). All of these significant correlations are suspect because of the skewness of the distribution of scores, as can be seen in Table 8:3.

However, it is worth mentioning the correlations with Variables 2, 3, and 4 (respectively, "vocational certainty", "Degree-vocation relevance", and "course-vocation relevance"). "Vocational certainty" is correlated with "lifestyle importance of service to individuals" ($p < 0.05$), and negatively with /degree relevance to service to individuals" ($p < 0.01$). Thus those individuals who are sure of their vocation are more likely to want to be of some service, but the attainment of a degree is less likely to be of relevance. Furthermore, individuals whose degree is relevant to their vocation (variable 3) are less likely to want to be of service to individuals ($p < 0.01$), and their degree is also likely to be relevant to their desire be of service to mankind. It would be possible to rationalise these results, but the hypothesised connections would be tenuous, and probably more than the data can stand.

8:6:3 Combinatory Codes

a) By hypothesis.

It is fairly easy to think of ways in which the variables might be combined in order to yield potential predictors of academic performance. In fact, it is possible to think of many more codes than it is feasible, or worthwhile to test. Inevitably, the selection presented below and tested in the present situation is somewhat arbitary, and in particular the arithmetical operations used in the codes are totally arbitary. This arbitariness, is no doubt reflected in the results; when the scores derived from these codes were correlated with examination performance, none of the correlations were significant, except someof those based on Variable 2 which yielded lower correlations with examination scores than did Variable 2. Nevertheless, for the sake of completeness, the following a priori codes were calculated.

Secondary Indices (based on the primary variables):

S1 Vocational Motivation (V1 + V2)

S2 Vocational Motivation for the degree (V2 x V3)

S3 Vocational Motivation to coursework (V2 x V4)

S4 Vocational Motivation to the education course (V3 x V4)

S5 Self Definition $\frac{1}{2}\left(V2 + \frac{(V5 + V6 + V7 + V8 + V9 = V10)}{6}\right)$

S6 Total degree relevance $\frac{1}{2}\left(V3 + \frac{(V11 + V12 + V13 + V14 + V15 + V16)}{6}\right)$

S7 Degree lifestyle effect
$1/6\left((V5 \times V11)+(V6 \times V12)+(V7 \times V13)+(V8 \times V14)+(V9 \times V15)+(V10 \times V16)\right)$

S8 Work level self concept V17 x S6

Tertiary Indices (derived from Secondary Indices).

T1 $\frac{S1 + S4}{2}$

T2 S1 x V3

T3 $S1 \times V4$

T4 $\frac{S2 + S3}{2}$

T5 $\frac{S4 \times S1}{10}$

T6 $\frac{S7 + S2}{2}$

T7 $\frac{S7 + S4}{2}$

b) By Induction.

Part of the aim of the present study was to allow for the possibility that the variables, or systems of variables, which predict academic performance differ with individual students or groups of students. One way of investigating this would be to look at each student individually, postulate which variables determine his academic performance, and then combine students who were similar into groups. However, this procedure is totally impractical: there are too many variables to hold in mind simultaneously, and there is no criterion for selecting which items of information known about the student are relevant to his academic performance. Another, more practicable, method of investigation is to start off by grouping the students into "types", where some of the types perform at a high level and others at a lower level. It may then be possible to derive the variables, or systems of variables, which differentiate between the "types", and which predict academic performance within a "type."

By hand, and with the 28 variables used in the present study, this method would be as impractical as the previous. However, methods of "numerical taxonomy" or "cluster analysis", which have been largely developed and used in biology, and are only recently being applied in psychology and education, allow the harnessing of the computer's superior sorting abilities. Cluster analytic procedures sort "cases" (in the present case, subjects) into groups or types on the basis of their similarity or dissimilarity in their scores. Thus if the subjects scores on the 28 variables were plotted in a 28-dimensional hyperspace, each subject would be represented by a point. With most clustering procedures, the two subjects who were nearest together in the hyperspace would then be the first to "cluster".

The present study used a suite of clustering programmes at the University of Aston. This suite was imported from the University of Bradford, who in turn imported it from the University of St. Andrews, where it is called "Clustan 1A" (Wishart, 1969). Inevitably, the movement of the suite of programs from one computer operating system to another has resulted in modifications: further more, the program at the University of Aston Computer Centre is still in an experimental state at the time of writing, and is not on general release, nor is it covered by a User's Manual. While the general functions performed by the Aston version of Clustan are the same as for the St. Andrew's version,the names of some of the sub-programs are different, and some subsidiary functions have been omitted. It is necessary to go into such detail about the pedigree of the programs used, as it was found at a late stage of the investigation that the Aston version of Clustan does not output the means and standard deviations of the hierarchical clusters on the 28 variables. This makes the interpretation of the clusters more difficult.

The basis of the clustering procedures is the similarity matrix and the k-linkage lists. The similarity matrix specifies the distance of each individual from each other individual on the 28 variables. The k-linkage lists are the lists of the k nearest neighbours to each case, based on the similarity matrix (in the present study , k = 5). The k-linkage lists are the basis of the iterative procedures used to calculate the clusterings. Both the similarity matrix andthe k-linkage lists are dependent on the nature of the distance criterion used, and on whether it is applied to raw or standard scores. The cluster suite has options for any of 40 similarity criteria to be used: in the absence of any knowledge about how these different criteria may be more or less compatible with cognitive structuring, the present study has used the simplest - "squared Euclidean Distance". The criterion was used by Entwhistle and Brennan (1971), and is based on a multi-dimensional deviation from Pythagoras'theorem. In the present study, the criterion was applied to standard scores as is the recommended procedure. Had this not been done, the clustering would have been biased by those variables with a larger variance, and there are considerable differences between variances of the different primary variables.

The way in which different subjects cluster together is also dependent on the clustering method. In the present study, five different clustering methods were used:

1. "Natural clusters". This is produced by program ATMO in the Aston version of clusters, and by program MODE in the original version. Program ATMO calculates the average A (I) of the 2k smallest distance coefficients for each individual (I), thus giving an estimate of the density of the space near each individual. Small A(I) values correspond to high density space. Individuals are then ordered in terms of their A(I) values, and the individual with the smallest A(I) (i.e. who is in the most densely populated area of the space) is the first introduced to the clustering procedure. During subsequent cycles, the density threshold R is raised to the next A(I) value, and those individuals are progressively introduced and become "dense" (i.e. join the cluster). At each cycle, the individual introduced may lead to one of four courses of action. First, he may be separated from all other dense points by a distance exceeding R, in which case this individual initiates a new cluster nucleus. Secondly, he may be within the distance R of one or more points from one cluster, in which case he joins that cluster. Thirdly, he may be within the distance R of points in two or more clusters, in which case

the clusters are fused. Or fourthly, the threshold A(I) value may be more than the smallest distance between dense points belonging to two different clusters: here again, these two clusters are fused. The program outputs the classification of individuals into clusters, and the means and standard deviations of the clusters on the variables, for each grouping obtained immediately before a cluster fusion.

The criterion of the "natural" clustering method is therefore based on the density of the hyperspace. In contrast, the other three clustering procedures used are based purely on similarity, where the entire sample of n individuals is taken to represent n clusters, and the two most similar clusters are 'fused' at each cycle until all individuals are contained in a single cluster. The output from these procedures is in terms of the order in which the clusters are fused, and the "fusion co-efficient" at which each fusion takes place. The output is presented in the form of a dendrogram table which can be drawn graphically. It is on these procedures that the Aston version of Clustan does not output means and standard deviations for the clusters: thus the program indicates which individuals cluster together, but does not give any information about their status on the variables, which makes it particularly difficult to interpret what the clusters 'mean'.

This clustering program, as mentioned above, fuses the two most similar clusters. In the present case, it was used with three different fusion criteria:

2. Hierarchical Clustering: Single linkage. The similarity between clusters is defined as the highest single similarity between individuals, one from each cluster. This criterion is also known as Nearest Neighbour, and Minimum Method (Johnson, 1967). The Single Linkage method has been reported as having a tendency to produce long, straggling, clusters in the form of a chain, (Sokal, 1966).

3. Hierarchical Clustering: Complete Linkage. Also known as Furthest Neighbour, and Maximum Method. It is defined as the smallest single

similarity between two individuals, one from each cluster. This criterion tends to produce spherical clusters, but is also reported as giving "irregular" results (Wishart, 1969).

4. Hierarchical Clustering: Average Linkage. Otherwise known as group average. The average of all the similarity co-efficients for pairs of individuals, one from each cluster. This criterion is more regular in the results it gives, in that it takes account of group structure. It also tends to produce spherical clusters. However, Johnson (1967) points out that this criterion, unlike the two previous criteria, and the one below is not invariant over monotonic transformations of the variables. This is particularly important in psychology, where it is not usually wise to invest the data with more than ordinal properties.

5. Hierarchical Clustering: Median. This criterion approximates to the average linkage, except that it uses the median similarity co-efficient, and does, of course, have the property of monotonic invariance.

As with factor analysis, the output to a cluster analysis depends on the input. In the present case the input is a number of variables which arbitarily sample a variety of cognitive domains, and consequently the cluster solutions are likely to have little value per se. This would be true even if the clusterings were robust (i.e. if the same clusters were produced by the different clustering methods), but the clusters derived in the present study are far from robust (indicating a relatively homogeneous spread of points throughout the hyperspace), which should double the reluctance to draw theoretical conclusions.

But the aim in applying cluster analysis to the present data is not to draw theoretical typological conclusions. Rather, cluster analysis is being used as a preliminary step, which will hopefully facilitate the drawing of inductive conclusions with regard to prediction of academic performance. The indicative conclusions and their subsequent testing on a second group are the important phases. With this in mind, the criterion for evaluating the cluster solutions can be specified as the degree to which they involve the splitting of the sample of individuals into individuals, and groups of individuals, of high and low performance. If the cluster analysis does result in such groupings, then there is a possibility of inferring combinatory codes: if the codes predict successfully to a second sample, the initial cluster solution is validated. If, on the other hand, the clusters produced do not differ on academic performance, they are of no use in deriving combinatory codes.

Five different cluster analyses were performed, using different samples and different variables. The first preliminary analysis used the total sample of subjects and the total number of primary variables. Other analyses used only the Correlated variables, and Group 1 only, students above the median on examination performance, and students below the mean on examination performance. All analyses used all five clustering methods. The correlated variables were basically those variables found to correlate with examination performance (on Group 1 and Group 2 data) at a level of 0.2 or above, with the addition of "Lifestyle importance of marriage" (which was included for completeness, as all other lifestyle values had correlated) and three other variables which were of interest in their own right. The list of correlated variables was therefore as follows:

3. Degree-vocation relevance.
5. Lifestyle importance of marriage
6. Lifestyle importance of wealth
7. Lifestyle importance of fame
8. Lifestyle importance of popularity
9. Lifestyle importance of service to individuals
10. Lifestyle importance of service to mankin d
17. Work satisfaction
18. Preference for Aston
19. Anxiety-Level
21. Study activity preference/exam benefit compatability
23. Topic preference/Topic importance compatibility
27. Study habits

28. Examination score

The results of the cluster analyses, and the attempt to derive combinatory codes from them were most disappointing. In view of their lack of value, and in particular their length, they will not be reported in full. Instead, a summary of the main points, and the most promising (or rather, least useless) results is presented below.

Analysis 1 Primary Variables (Nv = 28) Groups 1 and 2 (Ns = 27)

1. "Natural" clustering. This produced only two clusters whose means on examination performance differed by less than 2 points.

2. Hierarchical Clustering. Nearest Neighbour. The dendrogram output, which summarised the order in which the clusters were collapsed, produced what was essentially one long chain cluster. This means, in practical terms, that any cluster solution taken (e.g. a 6 cluster solution), gives one large cluster plus a number of individuals (one cluster plus 5 individuals). Furthermore, the chaining bore no apparent relationship to examination performance.

3. Hierarchical Clustering. Furthest Neighbour. The dendrogram produced by the Furthest Neighbour criterion was slightly more encouraging, there was less chaining, thus allowing the selection of a solution with a number of clusters of individuals. Clusters containing small groups of individuals are better for present purposes than large groups, since they are obviously more likely to have a consistent level of examination performance. The nine-cluster solution and the six-cluster solution were selected as being the most promising. The 6-cluster solution gave 2 clusters of one individual, one of 3 individuals, one of 5 individuals, one of 8 individuals, and one of 9 individuals. However, the differences in performance were not great, and there was a large variation in the performance within clusters. The nine cluster solution gave three clusters of 1 individual, three of 2 individuals, one of 3, one of 6, and one of 9 individuals. There was more consistency of performance within clusters and a slightly better differentiation in the performance of different clusters.

It was mentioned earlier that the Aston version of Clustan failed to output the means and s.d.'s of the clusters produced by the hierarchical clustering procedure. This makes it very difficult to interpret the clusters and derive combinatory codes from the cluster solutions - as this task requires a knowledge of how the clusters differ from each other and are similar to each other. It was not possible for the author to calculate the means and s.d.'s by hand, partly because of the sheer size of the task, (over 1,000 such calculations at a minimum), but mainly because the different clustering procedures mean that the arithmetical mean of a cluster is not the mean of the individuals in that cluster. Thus the calculations of the means of the clusters is a complicated procedure which the writer has not the time to solve.

In the absence of this program output, the writer had to revert to the "string and chewing gum" traditions of an earlier psychology. The subjects were rank ordered on all the variables, and the clusters were then evaluated by visual inspection. Thus the cluster is listed as containing a number of individuals: each variable was taken, and an impression was obtained of the level of scores of the individuals on the variable, and the consistency of the scores within the cluster. Unfortunately, impressions of "level" and "consistency" noted down on a 5-point scale of low-high, are but poor substitutes for means and standard deviations.

The nine-cluster solution was examined in this way on the 28 variables, but it was not possible to obtain a coherent description of any one cluster. Since this was a preliminary analysis on the total sample of subjects, no attempt was made to derive any combinatory codes.

4. Hiererchical Clustering. Group Average. This criterion produced a dendrogram which was intermediate to the nearest neighbour and furthest neighbour structures. The chaining was pronounced, but not as badly as in the former method. Because of this, all the cluster solutions involved one large cluster together with a number of very small ones. The large cluster was not consistent on examination performance.

5. Hierarchical clustering. Median. This criterion produced a very similar dendrogram to the former. It was equally difficult to obtain moderate-sized clusters with consistent examination scores.

Analysis 2 Correlated Variables (Nv = 14) Group 1 only (Ns = 18)

1. "Natural" clustering. This produced two separate solutions. The first solution involved two clusters, one of three individuals and one of 15 individuals. The means of the two clusters differed by less than one point. The second solution also involved two clusters, one of 2 individuals (not included in the previous 3 individuals cluster) and one of 16 individuals. The two-individual cluster contained one individual of poor performance and one individual of moderate performance. Any conclusions drawn from the nature of these clusters can therefore have little relevance to examination performance.

2. Hierarchical clustering. On this sample, both the Nearest Neighbour criterion and the Median criterion produced chain clusters which are not suitable for present purposes. The Furthest Neighbour and Group Average criteria both produced cluster hierarchies which were satisfactory and very similar. The Furthest Neighbour dendrogram was therefore analysed in depth, as this criterion does not break the assumption of monotone invariance. The dendrogram is shown in figure 8:2.

The 11 cluster solution was selected as the most satisfactory in that it involved anumber of small clusters, with relatively low intra-cluster variance on examination performance. The 11 clusters included 5 clusters of one individual each, 5 clusters of two individuals, and one cluster of 3 individuals. Of the 5 one-individual clusters, four of them had low performance, and one had high performance. Of the remaining clusters, one had low performance, one had moderate-low performance, three had moderate-high performance, and one was inconsistent (high and low performance in the same cluster).

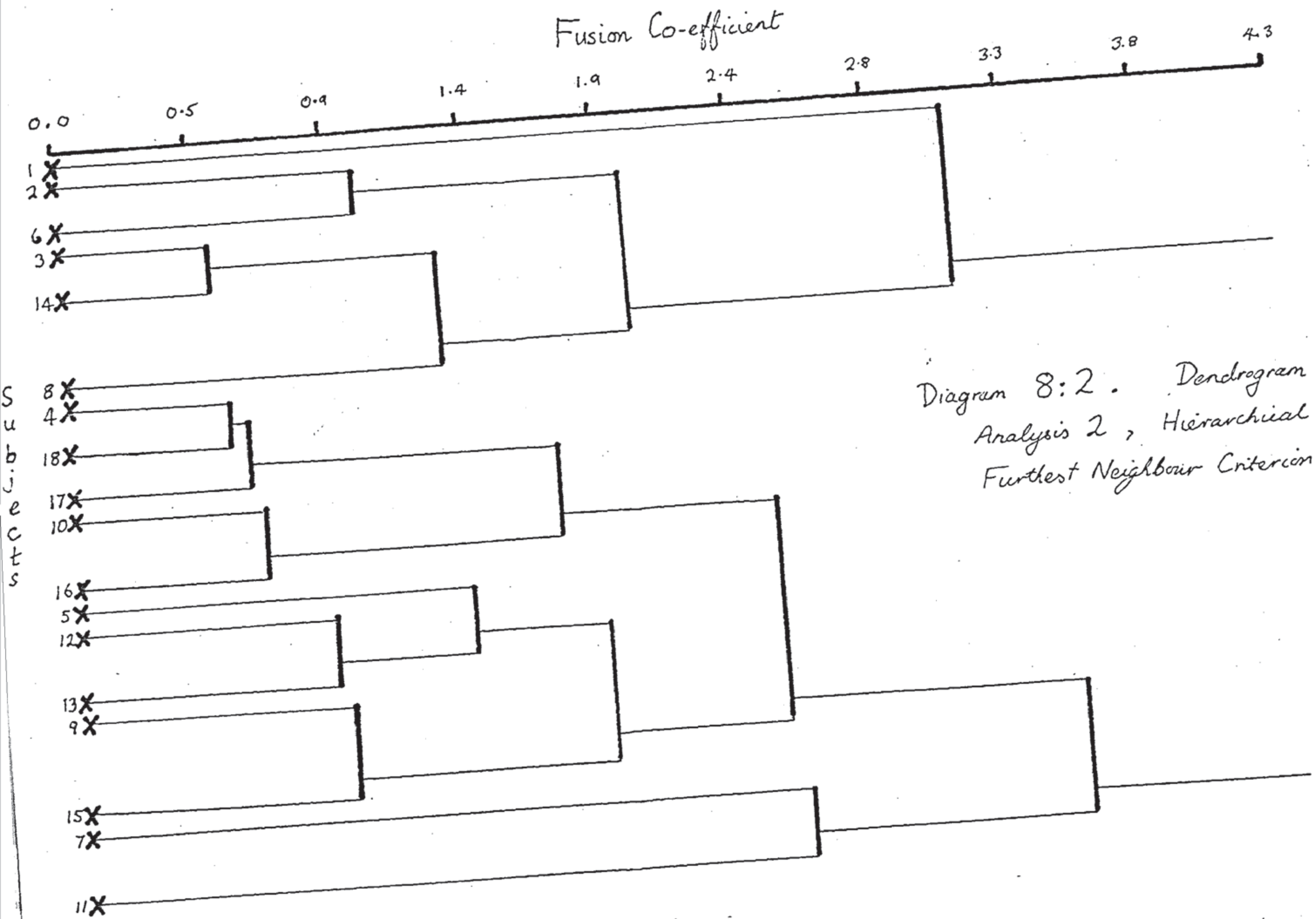


Diagram 8:2. Dendrogram from Analysis 2, Hierarchial Clustering, Furthest Neighbour Criterion.

The description of these clusters on the 14 correlated variables is presented below. Where a variable is not mentioned for a particular cluster, the scores on that variable were inconsistent. It is regrettable that these descriptions cannot contain any numeric information.

Cluster 1 Very poor performance. This student had a high anxiety level and low work satisfaction. He* felt that his degree was highly relevant to his vocation, but he had come into Aston on the Clearing Scheme. With regard to the lifestyle values, he had a high need for marriage, wealth, fame, and popularity, and a low need to serve individuals and mankind. Study habits were very poor, his preference for study activities was highly compatible with his perception of their importance, and his topic preference/topic importance compatibility score was moderate.

Cluster 2 This cluster contained two individuals who had a Moderate-High performance in the examinations. Their expressed need for all the lifestyle values was moderate to low, as was their study habits, study activity compatibility and work satisfaction. Anxiety level was moderate to high, as was topic preference/topic importance compatibility. The degree vocation relevance was low-moderate.

Cluster 3 Again containing two individuals, who were moderate-low on examination performance. This cluster had a low anxiety level, a low-moderate work satisfaction, and a low-moderate study habits score. With regard to lifestyle values, need for marriage, fame and service to individuals were low, and need for service to mankind was moderate-low: need for popularity was moderate and need for wealth high. Degree-vocation relevance was moderate-high, as was study, activity compatibility, and topic preference/importance compatibility. These students came to Aston on the Clearing Scheme.

*The pronoun "he" is used in all descriptions of individual students regardless of their sex.

Cluster 4 One individual who had a high examination performance. He had poor study habits, a low work satisfaction, and a moderate anxiety level. On the lifestyle values, he had a low need for fame, popularity, and service to individuals, a high need for service to mankind, and a moderate need for wealth and marriage. His degree vocation relevance was high, and Aston was the last choice on his UCCA form. Both his study activity compatibility and his topic preference/topic importance compatibility were high.

Cluster 5 Containing 3 individuals of moderate-high examination performance. Degree-vocation relevance and anxiety level were both low to moderate. Need for marriage, popularity and service to individuals were high, need for wealth moderate, and need for fame low. Study activities compatibility was moderate, and these students had Aston as the sixth choice on the UCCA form, or entered in the Clearing Scheme.

Cluster 6 Two individuals with high examination performance. These individuals had low anxiety, high work satisfaction and good study habits. Their topic preference/importance compatibility was moderate, their study activity compatibility variable, and their degree was not perceived to be relevant to their vocation. They had a low need for wealth, fame, and popularity and a fairly high need for service to individuals, and mankind and came to Aston via the Clearing Scheme.

Cluster 7 One individual with a low examination performance. Good study habits, low anxiety and high work satisfaction, and Aston fourth position on the UCCA form. His degree was moderately relevant to his vocation, and both his study activity compatibility and his topic preference/importance compatibility were low. With regard to lifestyle values, his need for service to mankind was low, his need for service to individuals, and for marriage and wealth were moderate, and his need for fame high.

Cluster 8 This contained two individuals, one of whom had a high examination performance, while the other was low. In view of this the general characteristics of the cluster were of little interest; suffice it to say that they differed on anxiety, work satisfaction, and topic preference/importance compatibility. The student with the high examination mark had low scores on the former two and high scores on the latter, with the relations reversed for the poor performer.

Cluster 9 Two individuals with a moderate examination performance. They had a high degree-vocation relevance, moderate-high study habits, moderate anxiety and low-moderate work satisfaction. Need for wealth and fame was low, and need for popularity and service to mankind was moderate-high. Study activities compatibility was moderate and topic preference/importance was low. These students have Aston sixth on their UCCA forms.

Cluster 10 One individual with a poor examination performance. Low anxiety, high work satisfaction and a moderate degree-vocation relevance characterised this student. Aston had been his first preference on his UCCA form. His study habits were moderate, and he placed importance on all of the lifestyle values. Study activity compatibility was moderate, and topic preference/importance compatibility was low.

Cluster 11 One student with a poor examination performance. Here again, a low anxiety level, a high level of work satisfaction, and good study habits, make this individual similar to a previous high performance cluster, (Cluster 6). Degree-vocation relevance was high, as was need for marriage, fame and popularity. Need for wealth was moderate, and need for service to individuals and mankind was low. Study activities compatibility was high and topic preference/importance compatibility was moderate. This student put Aston fourth on his UCCA form.

These clusters were also examined in terms of their status on the other non-correlated, primary variables, and on the secondary, and tertiary indices hypothesised above (although, of course, these did not serve as a basis for the clustering.) Even using all of these variables, the attempt to construct any form of combinatory code was totally unsuccessful. It will be apparent from the descriptions of the clusters that the patterns of scores for high and low performance clusters appear to be equally arbitary and unpredictable: the addition of the other variables to the descriptions does not reduce these characteristics in the slightest. From the point of view of very simple relationships, only one was noticed: anybody who put more than the lowest possible score on "lifestyle importance of fame" obtained a poor examination mark. This result did not predict for Group 2, and in any case, is likely to be either random, or a function of Group 1.

Furthermore, when the top three and bottom three examination scores were taken, and the position of the students on the other variables was examined, there was no indication in the data as to why one student might perform better than the other. In view of the lack of intuitive hypotheses which could be applied, the search for more complex patterns, or systems of variables, became almost impossible, and totally fruitless.

Analysis 3 and 4 Correlated Variables (Nv = 14)

Subjects from the total sample who were above the median examination performance, (Ns = 13) or who were below the median, and including the median (Ns = 14).

As analyses 3 and 4 were performed on the total data sample, there is no question of deriving combinatory codes and applying them to an independent group. The aim of these analyses was to "force" the clusterings to fit into the High-performance - low performance typology, and thus to provide information for a typological analysis on a wider sample base. Unfortunately, the total lack of success in the search for correlates or combinatory codes, accociated with academic performance, means that there are no inductive hypotheses with which to temper the arbitary nature of the cluster analysis input. The position of the clusters produced by Analysis 3 and 4 on the primary variables has been examined: they appear to be fully as arbitrary and unpredictable as the clusters reported from Analysis 2, and there is consequently little point in reporting them.

8:6:3 Conventional Cognitive Styles

The three scores derived from the two conventional cognitive style tests were used to predict examination results be means of the Mann Whitney U Test. The tests were carried out on the data for Group 1 only, and then repeated on the total sample. The results were as follows:

Group 1 only ($n1 = 8$, $n2 = 8$)

CFT	$U = 26$
CFT - reflection/impulsivity score	$U = 18$ $p = \leqslant 0.08$ one tail
Balanced F scale	$U = 31$

Group 1 & Group 2 ($n1 = 13$, $n2 = 12$)

CFT	$U = 44$ $p \leqslant 0.05$ one tail
CFT reflection/impulsivity score	$U = 25$
Balanced F Scale	$U = 62$

The only measure to predict examination performance in Group 1 was the reflection/impulsivity score derived from the CFT. This only predicted at the level of $p = 0.08$, which would normally be considered an unacceptable level of prediction. However, this reslt does confirm the result observed in the experiment, that more reflective individuals tend to perform better. Bearing in mind the sample sizes, neither result should be taken as a firm conclusion, but together they constitute a very powerful argument for further research into a score which has not to the writer's knowledge, previously been used.

The reflection-impulsivity finding was not replicated on the sample considered as a whole. However the CFT score predicted at the level of ($p \leqslant 0.05$), indicating that differentiated, or field independent subjects perform better on examinations. This is an interesting finding, in that the experiment found field dependence to predict relative performance under different teaching methods, but not overall course performance. The present finding should not be taken as a firm conclusion, especially in view of the fact that it emerged only when group 2 scores were added. Group 2 were higher on field independence, and also performed better in examinations, than group 1, and so the finding may be the result of differential sampling. Here again, the results of the experiment and this study considered together constitute a powerful argument for further research into the measures of field dependence in the educational situation.

The massive level of prediction produced by the Balanced F scale in the experiment was not replicated in the present study. In fact there was scarcely any trend in the data, so this lack cannot be ascribed to small samples. This suggests that the result must be specific to the course work, the students, or some other aspect of the experiment.

8:7 Discussion

There can be few studies in which the report of substantive results has been as barren as the present. The most immediate conclusion which springs to mind is that the approach adopted here is inappropriate, and the arguments presented in the theoretical section are spurious. However, the writer has reconsidered these arguments in detail, and can find no fault with the logic. As to the presuppositions of the argument, it is to be expected that any global reasoning will be based on a selection of the relevant information, as the pool of relevant information is too large for any one thinker. Thus the selection is likely to be biased, both by the experiences and development of the thinker, and by possible subconscious processes within the thinker. Bearing these biases in mind, the arguments still seem to apply to the writer's own cognitive domain of "psychological knowledge."

In another, more particular sense, the approach may very well be inappropriate to the research problem tackled here. The aim of predicting academic performance from the patterns or systems of elements of cognitive content within the individual does presuppose that academic performance is also an element in cognition. It is possible that academic performance as measured by examination results is a) meaningless and random, b) non-random, but with no meaning to the individual, or c) meaningful to the individual, but relevant to cognitive domains not considered here. Evidence was presented in Chapter 5 that examination results is general appear to have a sizeable random factor, and that predictions above a level of 0.6-0.7 cannot be expected. Suffice it to say that predictions at the level of 0.6-0.7 would have been entirely satisfactory to the present writer. Secondly, the non-random element may be psychologically meaningless. It may be that students have no concept of what makes a good examination performance: they are merely interested or not interested, work hard or not so hard, etc., and the assessment of their performance is an artefactual and external judgement which has no meaning until after the event. This would imply that students cannot predict the examination performance of themselves or their friends. The writer knows of no research which has examined the predictive abilities of the students themselves, although it is probable that such research exists. But informal questioning of students, and memories of sitting examinations, suggest that students can predict at least a grade rank ordering with a 50-60 per cent accuracy. Although possibly predicting their own grade wrongly, it seems likely that they can predict with fair accuracy that A will obtain the same grade as themselves, and B will do much better, whereas C will fail or scrape a pass. If this is so, then there must be some representation of academic performance levels in cognition.

Questioning of students has further revealed that this prediction is based on judgements such as "how easy or difficult they, and other students find particular courses", and performance, in the shape of examination results, are quite important to most students, and it is possible that they build up a separate judgement domain related to degree results, but which is entirely separated from other domains associated with study habits and activities, perceptions of the subject area, etc. In this connection, the writer recalls predicting that a fellow student would obtain a good degree despite an incomprehension of some of the most basic concepts in psychology: he had a glittering verbosity which was particularly difficult to penetrate. Evidence such as this, is, however, no more than suggestive.

If this is the case, then the present study was inappropriate in that it was tapping the wrong domains of cognition. It can also be added that a study which successfully tapped the appropriate domain and predicted academic performance would add little to our knowledge of performance: it would only be using the model which the student had already built up unconsciously, and would perhaps reveal more about the mechanics of examinations, than the processes of educational development.

Despite the lack of substantive results, the present study is not entirely wasted. There are a number of methodological issues which are thrown into relief by the performance of the study, but which are less likely to be predicted in advance. The research method applied has placed a very heavy emphasis on inductive procedures. Inductive procedures are difficult and subjective at the best of times, and provide a much larger scope for imagination and the lack of imagination than does hypothesis testing (this comparison does not of course apply to the generation of hypotheses). The writer did not therefore embark on the present study with excessive optimism. However, in a number of respects, the study made the inductive procedures more difficult:

1. Number of variables. When involved in induction at a subjective level, any increase in the number of variables makes the many regularities far more difficult to observe. The Biological taxonomists were characterised as only being able to take into account two or three variables at any one time (Sokal, 1966), and the writer now understands this limitation very fully.

2. Number of cases. The same comment applies to the number of cases examined. The present study, using 27 cases, and 28 variables, was foredoomed to failure in terms of subjective induction.

3. Despite the excess of variables and cases, the writer feels that the main defect of the present study was the lack of prior and specific hypothesis. This lack of hypotheses was deliberate, and was the result of an unwillingness to make any more structural assumptions about cognition than were absolutely necessary. Furthermore, academic performance is a global measure and any attempt to predict it must involve a fairly global sampling of cognition. The present sampling was fairly global, and consequently fairly superficial. Lifestyle values, anxiety level, preference for Aston, and study activities are poles apart, and there was no sampling of elements, which might suggest how, if at all, they might be connected. Thus the study sampled a number of dispersed cognitive elements, which were unlikely to be related in systems because of their dispersion. Thus it may be that the study failed to achieve any results because the lack of prior hypotheses was supplemented by a lack of potential hypotheses.

The study also placed some reliance on cluster analytic techniques. There is no doubt that cluster analysis is a useful supplement to factor analysis, but it may be totally inappropriate to the present type of study. Cluster analysis is a purely taxonomic technique, producing clusters on the basis of similarities only. However, in attempting to map cognitive content in the individual, it seems to the writer that systems and structures (in the Piagetian sense) are likely to be more powerful analytic concepts than "types of similar scorers". Cluster analysis is as insensitive to these more complex relationships as are the normative correlational techniques. It seems that the importance of cluster analysis is as a preliminary tool. Given a number of global indices that are known to be of major heuristic significance, cluster analysis will help to show which groupings of cases should be considered in a more penetrating analysis. But where the heuristic significance of the indices is not known, as in the present study, the results from a cluster analysis are likely to be either meaningless or misleading.

In conclusion, the in-depth study of cognitive content in the individual may be, as the author believes, the only way to arrive at a detailed understanding of complex cognitive activities, but it must also be uncompromisingly molecular. The aim of producing general predictions of global criteria is either unrealistic, or far in the future. The present study would have been more useful had it taken a more specific research criterion (such as a carefully prepared measure of performance in response to a particular course), a very small number of subjects (for instance, five or six), and a more intensive analysis of particular domains of content. The research criterion would need to be more differentiated than a simple "level of performance" measure: an improved version of the measure used in the experiment might be suitable. And the analysis of cognitive domains would have to use more informative and more time-consuming techniques, such as repertory grids and logical trees.

CHAPTER 9

REVIEW

9:1 Introduction

It is always difficult to map the logical structure of a research report onto the essentially intuitive and circumstantial structure of a research project. In the present case it is more difficult than usual. Thus the writer feels that the logical structure of the present thesis is satisfactory, but he is aware that it has left certain questions as to the "motivation" of the research - as to why the research has dealt with certain aspects of things but not others - these questions are left unanswered. If such questions have occurred to the reader, a brief account of the temporal and circumstantial development of the research project may help to answer them.

The research began as a small study within the context of an SSRC - sponsored project into "the use of project work in Colleges of Further Education". The major study in this project was what might be called "action research", and involved correlating a small number of individual difference measures with tutor's ratings on project work performance, based on a large sample of students from some 25 Colleges of Further Education. The conscious aim of this study was to involve the college tutors in the research, and so to stimulate them to consider and systematise their use of project work. As a result of this, and as a result of the sampling, the study had negligible internal validity, and was not expected to yield any strong empirical results (see Small, 1973). The minor study, carried out by the writer and reported in Chapter 7, was intended to be complementary to the major study, being methodologically fairly tight, and giving some hope of strong empirical conclusions. It aimed to use experimental controls where none were possible in the major study.

The experiment was consequently researched and piloted in 1971, and performed in February - May 1972. It was intended to be a preliminary study narrowing down the field of cognitive style measures worth investigating in this context, and to be followed by a more detailed investigation of particular cognitive style measures. However, simultaneously with the piloting and the running of the experiment, the writer was becoming dissatisfied with the use of experiments in general, and with normative tests. The writer's "feelings" from this time were later elaborated into the logical arguments which were presented in Chapter 4.

The SSRC contract ended in September 1972, and the writer became free to follow up these "feelings" from that date. The timetable of research dictated that a study should be designed by December 1972, and run by March 1973, but the "feelings" had not crystallised into a theoretical position until April 1973. Thus the writer was in the position of having to design a study which would fulfil the theoretical requirements he would delineate in six months' time. This is a somewhat hazardous way of doing research, but is the way which is dictated when the more mundane considerations of time and money are allowed to intrude.

Thus the temporal logic of the thesis might be:

Section 1. Chapters 1, 2, 3, 5, 7. Cognitive style dimensions: literature survey and experiment.

Section 2. Chapters 4, 6, 8 Non-dimensional approach: theoretical survey and field study.

However, the writer decided to follow the marginally more elegant structure of a theoretical comparison between conventional approaches to cognitive style and his own insights, followed by an empirical comparison.

Perhaps because the research followed two approaches, neither empirical study can be thought of as more than a pilot study. It is now necessary to review, in turn, these two studies.

9:2 The Experiment

In a sense it is not necessary to criticize the experiment in that the second part of the thesis (Chapters 4, 5, 6, 8) constitutes a reasoned contradiction of the presuppositions of the experiment. Had the present theoretical position been held two years previously, the experiment would never have been performed.

However, if the overall strategy is accepted for a moment, there are other weaknesses (which have been pointed out to me and which I acknowledge) which are worth mentioning:

1. The data analysis relied exclusively on non-parametric statistics, on the grounds that these made less rigid assumptions about the sampling distributions of psychological characteristics. In doing this, the vastly superior analytic power of parametric statistics was also sacrificed. There is also a modern trend towards stating a specific probability of occurrence of a result, and then considering that probability in the context of all relevant details of sampling etc., rather than relying on a rigid "level of significance". This trend appears to the writer to be an advance, in precision, and in making explicit the considerations which were hidden under the blanket result of "no significant differences" "$p \leqslant 0.05$": although it should be borne in mind that

a skilled researcher would have looked beyond the significance levels if he were doing anything more than "catching up on the journals". The abandonment of the conventional "levels of significance" does also make it less necessary to obey distribution assumptions of tests: disobedience becomes what it should be - a practice which may be worthwhile in certain circumstances, and then an additional consideration when evaluating the results. In this context, the decision to use non-parametric tests in the experiment was wrong.

2. The cognitive style dimensions used were selected to form a "battery" of tests. The "test battery" approach in psychometrics has recently come under heavy criticism for a variety of reasons:

(a) It is an "easy way out": if one includes a wide variety of tests, one of them is sure to correlate, but such a correlation gives very little information even about that particular test, and no information about the tests which do not correlate.

(b) The factor analysis which so often follows the empirical study, is more likely, in the absence of clear theoretical hypotheses, to obscure the interpretation of results than to clarify it.

(c) The step-by-step collection of so many "ad hoc" correlations does almost nothing to help the construction of theories.

The writer does not believe that these objections apply in the present case. The experiment began from the theoretical position that all of the cognitive style dimensions used have some potential predictive validity in the academic situation (Chapters 2 and 3), and from the empirical position that little actual research had been done to investigate this potential (Chapter 5) the aim of the experiment was therefore to see which dimensions actually did predict academic performance. Furthermore, the "battery" was more apparent than real: logically, the experiment involved using a number of tests separately on different combinations of subjects, rather than using all the tests in combination on the same subjects. There was thus no possibility of a multiple regression.

(d) This type of study rests on the notion of predictive validity. In its narrow sense, i.e. that of predicting, in advance, by means of correlation co-efficient, this concept is inadequate. It involves the assumption of linear causal relationship, and suggests the philosophical assumption that academic performance is determined by a person's cognitive style and not by the educational process. In fact, the criticisms of the experimental method (Chapter 4) can be read also as a criticism of this idea of predictive validity, and this fault in the first study is an aspect of the approach which the writer has later rejected.

It should be noted that the writer does not reject the broader interpretation of predictive validity, which includes "prediction" after the event (understanding), and which is not tied to linear relationships. Of course it is a moot point as to whether it is better to evolve non-linear statistical techniques, or to attempt to convert non-linear relationships into linear statistics before a statistical analysis. The writer opted for the latter in the theoretical notion of a hierarchy of predictive indices.

9:3 The Field Study

The field study suffered from many faults, both major and minor. Some of these faults can be traced to the temporal necessities of the study, as was mentioned above, but it must be admitted that the writer can find no excuse for the others. Most of the criticisms of the study have been dealt with in detail in the previous Chapter, but it may be worthwhile to recapitulate them here, and add others which have since been pointed out.

Firstly, to consider the field study in its context in the research project, the main criticism is that it did not meet the research specifications delineated in Chapters 4 and 6. Chapter 4 presented theoretical arguments to the effect that both the analytic concepts of cognitive psychology and the dimensions of cognitive style rest on assumptions about the ability to describe cognitive functioning in normative terms. It was suggested that these assumptions were at best untested, and at worst, false, and that an approach which would take account of idiosyncratic cognitive elements should be attempted. Chapter 6 suggested that such an approach would need to reflect certain characteristics of cognition:

1) Cognitive elements are internally referenced

2) Cognition is systemic or structural

3) The importance of cognitive context

4) The importance of cognitive content

Chapter 6 also proposed that a "Method of Structured Observation" would conform to these characteristics. The field study did not, however, conform to these characteristics. Some circumstantial factors relating to this shortfall were mentioned at the beginning of the chapter. The nature of the shortfall was as follows:

1) Internal referencing of cognitive elements. Although many of the primary data were normative observations, the use of combination codes applied intra-individually made some allowance for internal-referencing. However, the most important observation was the examination results - the criterion score - and this was externally referenced. One can make out a good case that examination results have no internal reference i.e. they have no cognitive representation other than as "information" after the event. In retrospect it seems silly to attempt to construct a~~s~~n internally-referenced structure to connect internally-referenced elements to a single externally-referenced element.

2) Cognition is systemic and possibly structural. This, also, was catered for in the use of combinatory codes, although the codes allowed for systems, but not structures (in the Piagetian sense). The failing was of course that the hypothetical systems were ad hoc, and that induction proved impracticable. In retrospect, again, it seems

silly to attempt to predict academic performance without a prior hypothetical explanation of that performance, but the acceptance of this point leads the researcher into an interesting dilemma, which is better dealt with below.

3) The importance of cognitive context. This is to emphasize the importance of measuring the structure of relevant domains. The field study took care to sample from a wide variety of domains, but in the failure of the combinatory codes, also failed to give an account of domain structure.

4) The importance of cognitive content. Very little was done on this point, largely because of the time-consuming nature of such observations coupled with the lack of research funds.

A retrospective analysis of the situation when the field study was being set up suggests that the researcher was in a dilemma, to which there is perhaps no solution. The researcher's aim was to attempt a systemic account of academic performance. This can be contrasted both with piecemeal studies (both experimental and correlational) and with multi-variate regression studies. Academic performance is a global index, and this has two consequences:

1) It is not possible to make strong hypothetical predictions in that almost any prediction can be rationalised. This is one reason why the a priori combinatory codes were so weak.

2) A global index demands a global sampling of cognition (which, incidentally, weakens theoretical connections even more), which demands in turn a large

number of measures and subjects. This, in turn, requires a large amount of money.

The field study attempted to make do with a weak theoretical position, bolstered up with induction, and with a small number of subjects. The field study failed. The dilemma for the researcher is what to do when faced with an aim which is apparently unachievable: on the one hand he can discard the aim, or on the other hand he can continue with a high risk study but resigned to failure. One can argue that of all research, Ph.D. research tends to be the least valuable, and is therefore "expendable" on high risk projects. For the present author, there was little choice: having worked on academic performance for two years, it was simply not conceivable to change in the third.

9:4 An Alternative Study

The failure of the field study leaves a void in the thesis. It remains to indicate a more suitable study which, if performed, would have more chance of substantiating the theoretical arguments presented.

Perhaps one of the most unsatisfactory aspects of the field study was the criterion. Academic performance, as indicated by examination results, seems to be too global, too unreliable, and psychologically irrelevant, as a criterion for the study of cognitive functioning. Examinations, however, largely consist of writing essays, albeit rather special essays. It may therefore be possible to obtain a more realistic criterion in essay writing.

With this in mind, examinations also include a number of chance factors, such as ability or luck in question spotting and sampling the syllabus, and task-irrelevant factors, such as the ability to handle stress. It may be possible to avoid these by using essay writing as a response to a seminar situation. There is a plausible link between these two types of essay writing, viz:

Essay writing (seminars) + stress resistance + motivational factors + random error = essay writing (examinations).

Such a link could easily be investigated, although the situation would no doubt turn out to be more complex than this. However, essay writing is still likely to be an unreliable criterion because of the unreliability of marking essays. To avoid this, it may be better to use essay marking by the student, as a criterion of his academic progress. One can put up a plausible argument that essay marking should be a valid predictor of essay writing: to perceive the merit of any aspect of an essay, presumably the marker should be potentially capable of producing that aspect himself. (To be fair, the counter-argument, that it is possible to perceive a spelling mistake without knowing the correct spelling, would also bear investigation). And one could also put up a plausible argument that essay marking is a <u>more valid</u> measure of academic progress than conventional examinations. However this involves consideration of educational aims and objectives, and is beyond the scope of the present thesis.

<u>9:4:1 Aim of the Study</u>

It will be remembered that the study is primarily taxonomic in nature, and is intended to discover something about the units of cognitive functioning, and to validate a meta-theoretical position. Within this context, the aim of the study is:

To predict an individual's marking of essays from measurements of relevant cognitive domains.

This stated aim is obviously too vague, and must be considered in conjunction with operational definitions presented below.

9:4:2 The Criterion

Essay marking. Rather than using 'natural' essays, which have an "unknown" content, it may be better to provide specimen essays which have a calculated content. There are a huge number of variables on which essays can be differentiated, so let us select, quite arbitrarily, just three: factual content, analysis and synthesis. For the purposes of this study, each of these variables could be binary: an essay could be said to possess merit or lack merit on each of the three variables.

By factual content, one is referring to the accuracy of the factual content, rather than to the "quantity of facts" used by the essay.

Analysis can be loosely described as the presentation of different arguments or points of view.

Synthesis would involve the subsumption of the analytic arguments within a super-ordinate argument, thus giving integration to the essay.

Eight essays, covering each combination of merit or lack of merit in the three variables would therefore be prepared, and checked for reliability with "skilled" judges. It would also be desirable to pilot them with a sample of subjects to ensure "unreliability" or individual differences. They would then be presented to the study sample as the culmination of a series of seminars on the topic. Each subject would be required to mark the presence or absence of merit, in terms of the three variables.

It is possible that there would be no individual differences between subjects on the presence or absence of a particular merit in essays. If this were so, it would be necessary to prepare a second set of essays which were similar on the three variables, and which were not marked reliably by skilled judges. If this were necessary, the study would still be satisfactory in terms of its cognitive style aspect, but would not indicate much about "objective" and academic performance.

9:4:3 The Combinatory Codes

Before working out the primary observations, it is necessary to specify the method by which a criterion prediction could be evolved. It would not appear to be possible to use quasi-mathematical formulae, either at this stage of the study, or after the primary observations have been specified. Instead, the combinatory codes can take the form of a series of postulates, which can be used to transform the primary data into a prediction. This, of course, is similar to the classical hypothetico-deductive method, although it differs in that the primary observations differ with each subject.

Let us call the essays which are presented to the subject, the "input structures". Let us also identify the primary observations with the subject himself, and call them the "existing structures".

The postulates are then as follows:

1. Factual Input Structures. A fact presented in an essay will be accepted unless directly contradicted in the existing structures.

Academic facts tend to be highly reliable. It would therefore seem to be a reasonable assumption that a subject will accept as true a fact - even a counter-intuitive fact - unless it is known to be inaccurate.

2. Analytic and Synthetic Input Structures. Let us suggest that these both consist of a) a core argument or idea, and b) the evidence or arguments presented in support. Respectively, these will be called the "core structure" and the "supporting structures". The postulate is therefore that an analytic structure or a synthetic structure will be judged as meritorious <u>either</u> if the core structure is consistent with existing structures

<u>or</u> if <u>both</u> the core and supporting structures are consistent <u>and</u> the supporting structures are consistent with existing structures.

3. Consistency. Two structures are <u>in</u>consistent if their psychological implications are mutually contradictory.

This is an adaptation of Festinger's (1957) definition of cognitive dissonance.

It should be noted that the postulate specifies psychological implications: these would tend to correspond to logical implications, although differences would occur as a result of both logical error, and limited cognitive context (principally, an implication may be logically appropriate, but psychologically out of context - "that's a different thing entirely").

This postulate is obviously the crucial one. Although it relies heavily on the investigator's judgement in comparing the essays with the sample observations of the individual's cognitive structures. It is to be hoped that the judgement is sufficiently precise as to be capable of objective application, and in any case, this could be easily tested.

<u>9:4:4 Primary Observations</u>

Now that the combinatory codes have been specified, it is possible to see what sort of information is needed from the primary observations.

Firstly, there is no problem with the factual elements. In order to establish whether an individual will appreciate factual accuracy in an essay, it is necessary to investigate his knowledge of these facts. This can be done by means of a forced-choice questionnaire which contains the facts to be presented in the essays, plus a number of others from the surrounding domain. Applying postulate No. 1, it is possible to see which inaccurate "facts" in the essays will be perceived by a given individual, and, by counting the number of hits, to predict which essays will be judged as worse.

The analytic and synthetic structures are by no means as easy. Here, there is a need to obtain some representation of the over-all cognitive domain which surrounds the essay topic, and to obtain it in a way which is reasonably independent of the observer, and is acceptable as an input to the consistency postulate. It is not possible to specify these measures in detail, without some pilot studies, but it is possible to indicate the types of measures which should be piloted.

Let us suppose that there had been a series of psychology seminars on the over-all topic of "motivation", and that the criterion essays were on some particular aspect of motivation. In order to score the subject's appreciation of analytic and synthetic merit, it is now necessary to obtain some idea of the subject's cognitive domain relating to motivation - the organization of the core structures, and their associated elements, and the pattern of implications and the limits of implications (or cognitive context) within the domain.

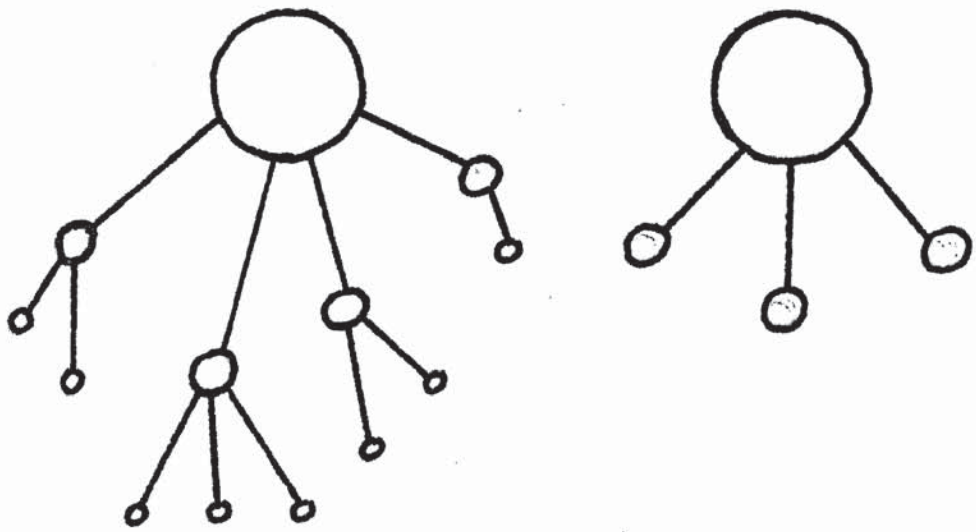
<u>The Core Structures</u>. Although a crude way of doing things, it may be possible simply to ask the subject what these are:

"If you were asked to write a book on motivation, what would your chapter headings be?"

One would hope that this question would elicit a relatively small number of super-ordinate categories, and that the titles would reflect the subject's categorisation of the domain. Of course, there is no guarantee that these are "core structures", or in any way more essential than other structures. But it might b e adequate as a starting point.

The titles elicited might be ambiguous, and it is therefore necessary to explore further the contents of the book? This might be done by taking each chapter in turn and listing the contents, and taking each 'content' and listing the subsidiary contents. If time, money and the subject's patience allowed, it would be desirable to go down as far as particular 'facts' (e.g. an experimental conclusion).

It would be possible to represent the information elicited in terms of a super-ordinate/subordinate map, in the form of a tree diagram:



It would be important to emphasize to the subject that the elements in the map would overlap, and that he should not mind repeating the same element in different places. This is to avoid stimulating an attempt at synthesis, and rendering the measuring procedure even more reactive than it is already likely to be.

Implications and Cognitive

The theoretical analysis in Chapter 4 would suggest that these natural groupings of elements elicited from the subjects are, in fact, abstractions from task schemata (or, in behaviouristic terms, habit hierarchies). Thus one would predict that implications and inferences would be more easily made within a domain, or sub-domain, than between them. This can be easily checked during the interview procedure by requiring the subject to make inferences:

> "If you were asked to participate in a psychological debate, and to propose/oppose the motion that (.....................), what arguments and evidence would you quote?"

The parentheses should be filled in with an appropriate issue which can be clearly located in one of the subject's sub-domains. The arguments and evidence elicited could also be traced to the domains of origin, with or without the subject's help.
A more direct test of implications and cognitive context would be to provide the subject with a 'specimen conclusion' (for instance, a fictitious experimental report of a startling nature), and to ask what other parts of the book would be affected. By using several of these, from different sub-domains, it should be possible to gain an impression of the limits of implication. Unfortunately, it is difficult to see how it could be much more precise than an impression: an

attempt to force such a sketchy representation of cognitive organisation into arithmetical formulae is likely to obscure more than it would clarify.

Finally, another indication of cognitive implications, which might also give evidence relative to synthetic structures, would be a test of forced implications. This is essentially a forced compliance technique (Osborne, 1953) applied to the core structures and to their elements:

"If I were to tell you that (............) and (..........) are directly related, can you tell me how?"

Parenthesis to be filled in with elements from the initial lists produced by the subject.

9:4:5 The Criterion Prediction

The investigator has in front of him the criterion essays which will be presented to the subject, and a number of indicators about the subject's cognitive domains, in the form of the primary observations. His task is to compare the subject in turn with each essay, and by applying the combinatory code postulates, to predict whether the merit or lack of merit of the essay on each variable will be appreciated.

9:5 CONCLUSIONS

This outline study would fulfil the requirements specified in Chapter 6 for an examplar of the Method of Structured Observation. Cognitive elements are internally-referenced, and this supposition is reflected in all aspects of this study. Cognition is structural: the present proposal has paid lip-service to the idea, if little more. In the writer's view, it is not possible to exemplify the characteristics of idiosyncratic structures which make them structures, without investigating in more detail the "mechanics" of cognitive functioning. The postulates of the combinatory code attempt to do this. The emphasis of the present study on both cognitive content and cognitive context, needs no further elaboration.

The study is weakest in the recording of the primary observations, and in the application of the consistency postulate. One would hope that both procedures could be applied with a moderate amount of objectivity, even at the present state of under-development: however there is little doubt that pilot research would both reveal problems and suggest improvements. Further specification of these procedures without pilot research would be fatuous.

A sceptic might comment that the proposed study is a living contradiction to the principle of parsimony in science: at its worst, it is using three or four hours of observation to predict an unimportant and irrelevant skill, and the observations are unlikely to predict anything else. This is a valid if short-sighted viewpoint. The study is intended to be tasconomic and is therefore also tautologous. Its value, if any, would be in validating the research procedure, confirming the combinatory codes (which are generalisable), and stimulating insights into cognitive structure and functioning.

This might be a beginning to the daunting, yet exciting, task of mapping cognitive content, and examining its relevance to everyday behaviour.

There are a number of measures already available. Semantic differential techniques, (Osgood, Suci, and Tannenbaum, 1957), and rating scales of all types, produce much valuable information. The Repertory Grid Technique developed by Kelly (1955), and modifications to this proposed by Hinkle (1965), Bannister and Fransella (1966), Ravenette, (1969), and Riedel (1970), have indicated one way of investigating idiosyncratic cognitive elements and deriving normative indices from them. These techniques have also been used to investigate possible structural connections between idiosyncratic elements: thus Hinkle's (1965) "Implications Grid" (or Impgrid) measures the extent to which one construct implies another and vice versa. This development seems to be particularly important in that it represents a break from the domain of the psychology of judgement, and the inter-relating of all types of constructs. In many ways these developments in Repertory Grid Technique represent a systematisation of earlier phenomenological methods, for instnce, Zajonc's (1960) procedure, and Hay's (1958) measures of trait implication and trait similarity. These developments are very promising, but the application of normative indices derived from them to particular behaviour is still by means of intuitive hypothesis. Thus Cartwright and Lerner (1963) found that improvement in psychotherapy was related to the amount of empathy between client and clinician, when empathy was defined as the similarity between the patient's self-description and the therapist's description of the patient. This is an interesting validation of Repertory Grid Technique, and a confirmation of a common-sense hypothesis which ought to have been (but might very well not have been) true. But the result derives its significance from the psychologists' intuitive knowledge of the psychotherapy situation: it is scientifically incomplete until one can derive a measure of the "nature" of the psychotherapist and client, and what they mean to each other in the context of the life process. Until this can be done, it is not possible to state why empathy should benefit this relationship, and possibly the teacher-student relationship, but possibly not the Sergeant-Major-Private relationship. In the same way, it is now possible to measure a student's degree of differentiation in the domain of perception of subject area and perception of future occupation. But it is also necessary to measure the significance of these domains in the student's life functioning, and how they relate to each other. What exactly is a knowledge of Thorndike's Law of Effect and in what way is it different from a knowledge of McClelland's Achievement Motive or a knowledge of how one's car functions? In the jargon of the field study presented above, Repertory Grid, and other, techniques, are very promising primary variables, which have been shown to relate to a number of criteria. What is lacking is a combinatory code which gives these variables significance at a theoretical level.

A theory of cognitive content has two separate aspects: a description of the structures of content, or of how a given element relates to other elements, and a theory of the nature of these connections.

With regard to the latter, there have been a number of conceptualisations in terms of implication, similarity, association, etc., some of which are mentioned above. The description of structures of content, however, is a far more difficult proposition which has hardly been attempted. To begin with, while different people from the same culture may have similar cognitive domains, the content and structuring of those domains is likely to be idiosyncratic. Secondly, the basic unit of description and its anchorage in objectivity, becomes a difficult problem once one leaves S-R psychology and empiricist philosophy. It is this second problem which Piaget's (1971) Biology and Knowledge tries to solve, by using the basic unit of "structure" and anchoring structures in internalised operations. Piaget's work, however, has dealt with certain very general structures (the structures of intelligence), and although he notes the contexts from which these structures develop, he has not attempted a conceptualisation of these contexts or of their cognitive counterparts. It is surely the contexts of action, and the way that these contexts are categorised by the individual that form the basis of knowledge and cognitive content.

In the writer's opinion, the problem of cognitive content is the most exciting challenge and the most important problem now facing psychology. Any movement towards its solution is a movement towards giving psychology the external validity and "human relevance" it is often criticised of lacking. The problem is also of inestimable importance to education: a theory of the nature and structure of content is obviously the pre-requisite to the study of its acquisition, and to individual differences in its acquisition. It is also the key to the design of a curriculum which efficiently fulfills educational aims and objectives, and to assessment procedures which measure the achievement of these aims and objectives.

REFERENCES

ABELSON, R.P., ARONSON, E., McGUIRE, W.J., NEWCOMB, T.M., ROSENBERG, M.J., and TANNENBAUM, P.H., (1968). Theories of Cognitive consistency: a Source-book. Rand McNally: Chicago.

ABERCROMBIE, M.L.J., HUNT, S. and STRINGER, P. (1968). Selection and Academic performance of students in a University School of Architecture. London: Society for Research into High Education Monographs.

ADEVAI, G., SILVERMAN, A.J. and McGOUGH, W.E., (1968). MMPI findings in field-dependent and field-independent subjects. Percept.Mot.Skills, 26, 3-8.

ADORNO, T.W., FRENKEL-BRUNSWIK, E., LEVINSON, D.J., and SANFORD, R.N., (1950). The Authoritarian Personality. New York. Harper and Row.

ALLARD, M., and CARLSON, E.R. The Generality of Cognitive Complexity. J.Soc.Psychology, 59. 73-5.

ALLPORT, G.W., (1961). Pattern and Growth in Personality. Holt, Reinhart, and Winton.

ARONSON, E., (1968), Dissonance Theory: Progress and Problems in the., ABELSON, R.P., ARONSON, E., McGUIRE, W.J., NEWCOMB, T.M., ROSENBERG, M.J., and TANNENBAUM, P.M. (Eds.) Theories of Cognitive Consistency: a sourcebook. Rand McNally.

ASCH, S.E., (1952), Social Psychology, Prentice Hall.

ATKINSON, R.C., and SHIFFRIN, R.M., (1968) and K.W.Spence and J.T.Spence (Eds.): the Psychology of Learning and Motivation: Advances in Research and Theory. Vol.2., Academic Press, New York and London.

AUSTWICK, K., (1960). G.C.E. to B.A. Universities Quarterly, 15, 67-71.

AVERBACH, E., and CORIELL, A.S., (1961). Short term memory in vision. Bell.Syst.Tech. J., 40, 309-328.

AVERBACH, E., and SPERLING, G., (1961). Short-term storage of information in vision. In C. Cherry (Ed.) Information Theory: Proceedings of the Fourth London Symposium. London: Butterworth.

BADDELEY, A.D., (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. Quart. J., exp. Psychol. 18. 302-309.

BADDELEY, A.D., (1966). Short term for word sequences as a function of acoustic, semantic, and formal similarity. Quart.J. exp. Psychol. 18, 362-5.

BADDELEY, A.D., (1970). Effects of acoustic and semantic similarity on short-term paired-associate learning. Brit.J.Psychol, 61, 335-343.

BADDELEY, A.D., and PATTERSON, K., (1971). The relation between long-term and short-term memory. Brit.Med.Bu....., 27, 237-242.

BAGG, D.C. (1968). The correlation of G.C.E. 'A' level grades with University Examinations in Chemical Engineering. Bri.Journ.Educ.Psych. 38, 194-7.

BAGG, D.C., (1970). 'A' levels and University Performance. Nature, 225, 1105-1108.

BAILEY, W., HUSTMEYER, F.E., and KRISTOFFERSON, A., (1961). Alcoholism, brain damage and perceptual dependence., Quart.J.Stud.Alcoh., 22, 387-93.

BANKS, C., KARDAK, V.S., JONES, E.M., and LUCAS, C.J., (1970). The relationship between mental health, academic performance and cognitive test scores among chemistry students. Brit. Journ.Educ.Psych. 40, 74-79.

BANNISTER, D., and FRANSELLA, F., (1966). A grid test of schizophrenic thought disorder, Brit. J.Soc.Clin.Psychol.,5, 95-102.

BARKER E.M., (1958). Authoritarianism of the political right, centre, and left; Unpublished Doctoral Dissertation, Teachers' College, Columbia Univ.

BASS, B., (1955). Authoritarianism or acquiescence: J.Abnorm. soc.psychol. 51, 611-23.

BARTLETT, F.C., (1932). Remembering. Cambridge University Press.

BARTLETT, F., (1958). Thinking, George Allen and Unwin, London.

BAUMAN, G. (1951). The stability of the individual's mode of perception, and of perception-personality relationship. Unpublished Doctoral Dissertation. New York University.

BEACH, L.R., (1960). Sociability and Academic achievement in various types of learning situations. J.Educ.Psych. 51, 208-12.

BEARD, R.M., and BLIGH, D.A., (1971). Research into teaching methods in higher education, Soc. for Research into Higher Education.

BENDIG, A.W. and SPRAGUE, J.L. (1954). The Guilford-Zimmermann Temperament Survey as a predictor of achievement level and achievement fluctuation. Journ.of App.Psy., 38, 409-413.

BENDIG, A.W. (1960)., Extraversion, neuroticism and student achievement in introductory psychol. Mournof Educ.Research, 52, 263-267.

BENNETT, S.W., (1973). Divergent thinking abilities - a validation study. Br. J.Educ.Psychol. 43, 1-7.

BIERI, J., (1955). Cognitive Complexity - simplicity and predictive behaviour. J.Abnorm. Soc.Psychol. 51, 263-8.

BEIRI, J. (1960). Parental identification, acceptance of authority and within-sex differences in cognitive behaviour. J.Abnorm.Soc.Psych. 60. 76-79.

BIERI, J., ATKINS, A.L., BRIAR, S., LOBECK, R., MILLER, H., and TRIPODI, T., (1966). Clinical and Social Judgement, Wiley.

BIERI, J., and BLACKER, E., (1956). The Generality of Cognitive complexity in the perception of people and inkblots, J.Abnorm. Soc.Psychol., 53, 112-17.

BIERI, J., BRADBURN, W.M., and GALINSKY, M.D., (1958). Sex differences in perceptual behaviour. J.Pers. 26, 1-12.

BIERI, J., and MESSERLEY, S. (1957). Differences in perceptual and cognitive behaviour as a function of experience type. J.Consult. Psychol. 21, (3) 217-221.

BLIGH, D.A., (1971). What's the Use of Lectures? University of London, University Teaching Methods Unit.

BLOOM, B.S., ENGELHART, M.D., FURST, E.J., HILL, W.H., and KRATHWOHL, D.R., (1956). Taxonomy of Educational Objectives, Handbook 1, Cognitive Domain. Longmans.

BLOOMBERG, M. (1967). An enquiry into the relationship between field independence-dependence and creativity. J.Psychol. 67, 1237-140.

BONARIUS, J.C.J., (1965). Research in the Personal construct of George A. Kelley, Role Construct Repertory Test and Basic Theory. In B.A. Maher (Ed.), Progress in Experimental Personality Research, Academic Press.

BORING, E.G. (1961) A History of Introspection. In Psychologist at Large. New York. Basic Books.

BOUSFIELD, W.A., (1953) The occurrence of clustering/the recall of randomly arranged associates. J.Gen,Psychol.49, 229-40.

BOWER, G.H., (1970). Organisational factors in memory, Cognitive Psychol. 1, 18-46.

BOWER, G.H. (1971). In Gregg, L., (Ed.) Cognition in Learning and Memory, Wiley, N.Y.

BOWER, G.H., CLARK, M.C., LESGOLD, A.M. and WINZENZ, D., (1969). Hierarchical Retrieval Schemes in recall of categorized word lists. J.Verb.Learan.Verb.Beha. 8,323-343.

BRANSFORD, J.D., and FRANKS, J.J. (1971). The abstraction of linguistic ideas. Cog.Psychol. 2, 331-350.

BREHM, M.L., BACK, K.W., and BOGDONOFF, M.D., A Physiological effect of cognitive dissonance under stress and deprivation. J.Abnorm.Soc.Psychol., 69, 303-310.

BRIDGMAN, P.W., (1969). The Way Things Are.(New York, Basic Books).

BROADBENT, D.E. (1958). Perception and Communication. London. Pergamon Press.

BROADBENT, D.E. (1970). In Defence of Empirical Psychology, Proc.R.Soc.B., 175, 333.

BROTMARKLE, R.A. (1933). What the Bernreuter Personality Inventory does not measure. Journ.App.Psychol. 17.

BROVERMAN, D.M. (1960 a) Dimensions of Cognitive Style. J. Person. 28, 169-85.

BROVERMAN, D.M., ~~xxx~~ (1960 b) Cognitive style and int individual variation in abilities, J.Person, 28, 240-256.

BROVERMAN, D.M., and LAZARUS,R.S. (1958). Individual differences sn task performance under conditions of cognitive interference, J.Person, 26, 94-105.

BRUNER, J.S. (1957) On Perceptual readiness. Psychol. Rev., 64, 123-52.

BRUNER, J.S., GOODNOW, J.J., and AUSTIN, G.A. (1956). A study of thinking. Wiley.

BRUNER, J.S., and GOODMAN, C.C. (1947). Value and Need as organising factors in Perception. J.Abnorm.Soc.Pwsychol. 42, 33-49.

BRUNER, J.S., and POSTMAN, L. (1947). Emotional selectivity in perception and reaction. J.Personal. 16.

BRUNER, J.S., and POSTMAN, L. (1948). An Approach to Social Perception, in W. Dennis (Ed.) Current Trends in Soc.Psychology, Pittsburgh, Univ.Pittsburgh Press.

BRUNER, J.S., and POSTMAN, L. (1949). On the perception of incongruity, J.Pers. 18, 206-23.

BRUNER, J.S., and RODRIGUES, J.S. (1953). Some Determinants of Apparent size. J.Abnorm.Soc.Psychol. 48, 17-24.

BRUNER, J.S., and TAJFEL, H. (1961). Cognitive Risk and Environmental Change, J.Abnorm.Soc.Psychol. 62,231-41.

BURDICK, H.A. and BURNESS, A.J. (1958). A test of "strain towards symmetry" theories, J.Abnorm.Soc.Psychol. 57, 367-369.

BURGESS, E. (1956). Personality factors in over-and under-achievement in engineering. Journ.Ed.Psychol. 47.

BYRNE, D., (1961). The repression-sensitisation scale: rationale, reliability and validity, J.Person, 29, 334-349.

BYRNE, D. (1964), Repression-sensitisation as a dimension of personality, in B.A. Maher (Ed.) Progress in experimental personality research, Vol.1, Academic press.

CAINE, T.M. (1970) in Boorer, D., A question of attitudes, Appendix 1, pps.54-64, London, Kings Fund Hospital, Centre.

CAINE, T.M., and LEIGH, R. (1972). Conservatism in relation to psychiatric treatment, Brit.J.Soc.Clin.Psychol. 11, 52-6.

CAINE, T.M., WIJESINGHE, B., and WOOD, R.R. (1973). Personality and Psychiatric treatment expectancies, Brit.J.Pshcyiatr. 122, 87-8.

CAMPBELL, B.M. (1960). Assumed similarity, perceived sociometric balance, and social influence. Unpublished Ph.D.Dissertation, Univ.of Colorado.

CAMPBELL, D.J., and STANLEY, J.C. (1963) Experimental and quasi-experimental designs for research on teaching. In N.L. Gage (Ed.) Handbook of research on teaching. pgs.171-246 - Rand-McNally.

CARMICHAEL, L.C., HOGAN, H,P., and WALTERS, A.A. (1932). An Experimental Study of the effect of language on the reproduction of visually-perceived form. J.Exp.Psychol. 15, 73-86.

CARR, J.E. (1965). The role of Conceptual organisation in Interpersonal discrimination, J.Psycho. 59, 159-76

CARTER, L.F. and SCOOLER, K., (1949). Value, need and other factors in perception. Psychol.Rev. 56, 200-207.

CARTWRIGHT, R.D., and LERNER, B., (1963). Empathy, need to change and improvement with psychotherapy. J.Consult.Psychol. 27, 138-44.

CATTELL, R.B. (1957 a) Personality and Motivation Structure and Measurement. London: Harrap.

CATTELL, R.B. (1965) The Scientific Analysis of Personality. London: Penguin Books.

CATTELL, R.B., BELOFF, H., and COAN, R.W. (1958), Handbook for the ImP.A.T. High School Personality Questionnaire. Illinois: Institute of Personality and Ability Testing.

CHRISTIE, R., and LINDAUER, F., (1963) Personality Structure, Am.Rev. of Psychol. Vol.1, 201-230.

CICIRELLI, V.G. (1965) Form of the relationship between creativity I.Q. and Academic achievement. J.Educ.Psychol. 56, 303-308.

CLARK, M.H., and BISHOP, G.H. (1954) Responses from an association area secondarily activated from optic cortex, J.Neurophysiol. 17, 271-77.

CLINE, V.B., TICHARDS, J.M., and NEEDHAM, W.E., (1963). Creativity Tests and achievement in high school science. J.Appl.Psychol. 47, 184-89.

COHEN, J., (1959) The factorial structure of the WISC at ages 7-6, 10-6 and 13-6. J.Consult.Psychol. 21, 283-290.

COHEN, L. and CHILD, D. (1969). Some sociological and psychological factors in University failure. Durham Research Review, 6. 22, 365-372.

COHN, T.S. (1952). Is the F-Scale indirect? J.Abn.Soc. Psychol. 47, 332.

CONRAD, R. (1964) Acoustic confusions in immediate memory. Brit.J.Psychol. 55, 75-84.

COX, R. (1967) Examinations and higher education. Universities Quarterly, 21, 292-340.

CRAIK, R.I.M. (1968). Two components in free recall, J.Verb. Learn.verb.behav. 7, 996-1004.

CRAIK, F.I.M. (1971) Primary memory, Brit.Med.Bull, 27, 232-236.

CRAIK, F.I.M. and LEVY, B.A. (1970) Semantic and acoustic information in primary memory. J.Exp.Psychol. 86, 77-82.

CRANO, W.D., and SCHRODER, H.M. (1967). Complexity of Attitude, structure and processes orconflict reduction. J.Personal.Soc.Psychol. 5, 110-114.

CROCKETT, W.H. (1965) Cognitive Complexity and impression formation, in B.A.Maher (Ed.) Progress in Experimental Personality Research, Vol.2. Academic Press.

CROPLEY, A.J. (1967) Creativity. Longmans, Green.

CROPLEY, A.J. (1967) Divergent thinking and science specialists, Nature 215, 5101, 671-672.

CROPLEY, A.J. and FIELD, T.W. (1968) Intellectual Style and High School Science, Nature, 217, 5135, 1211-1212.

CRUTCHFIELD, R.S. (1955) Conformity and character, Aner. Psychologist. 10, 191-198.

CRUTCHFIELD, R.S. WOODWORTH, D.G., and ALBRECHT, R.E.,(1958) Perceptual Performance and the effective person, U.S.A.F.

DALE, B. (1935) The use of mental tests with university women students. Brit.Journ.Ed.Psychol. 5, 59-75;

DALE, R.R. (1954) From School to University. London: Routledge and Kegan Paul.

DALE, R.R., (1958) University Standards. Universities Quarterly, 13, 186-195.

DANA, R.H., and GOOCHER, B. (1959) Embedded-figures and personality. Percept. mot. skills, 9, 99-102.

DEESE, J. (1969) Influence of inter-item associative strength upon immediate free recall. Psychol. Reports, 5, 305-312.

DEESE, J. (1969) Conceptual categories in the study of content, In G. Gerbner (Ed.) communication and content. Wiley.

DEMENT, W. and KLEITMAN, (1957) The relations of eye movements during sleep to dream activity. An objective method for the study of deaming. J.exp.Psychol. 53, 339-346.

DeWIT, F. (1955) The measurement of values by means of analysis of judgements. Unpublished Doctor Dissertation Univ.of Illinois.

DEUTSCH, J., and DEUTSCH, D., (1963) Attention: some theoretical considerations. Psychol. Rev. 70, 80-90.

DIXON, N.F., (1958) Apparent changes in the visual threshold as a function of subliminat stimulation. Quart. J.Exp.Psychol. 10, 211-219.

DOTY, R.W. (1958) Potentials evoked in cat cerebral cortex by diffuse and by punctiform photic stimuli. J.Neurophsyio. 21, 437-64.

EAGLE, M. and ORTOF, E., (1967) J. Verb. Learn Ver. Behav. 6, 226-

EGGLESTON, J.F. and KELLY, P.J. (1970) The assessment of project work in 'A' Level Biology, Educ.Res.12, 225-229.

ELLIOTT, R. (1961) Interrelationships among measures of field-dependence, ability and personality traits. J.Abnorm. Soc.Psychol. 63, 27-36.

ENTWHISTLE, N.J. (1971) Personality and academic attainment, Brit.J.Educ. Psychol. 42, 137-151.

ENTWHISTLE N.J. and ENTWHISTLE, D (1970) The relationship between personality, study methods, and academic performance. Brit.Journ.Ed.Psychol. 4, 2, 132-143.

ENTWHISTLE, N.J and BRENNAN, T. (1971) The academic performance of students, II - Types of successful student. Brit.J.Educ.Psychol., 41 268-276.

ENTWHISTLE, N.J. and PERCY, K.A., and NISBET, J.B. (1971) Educational objectives and Academic Performance in Higher Education. Unpublished Report, Dept. of Educational Research, Univ.of Lancaster.

EVANS, F.J. (1967) Field dependence and the Maudsley Personality Inventory Percept.Mot.Skills, 24, 526.

EYSENCK, H.J. (1947) Student selection by means of psychological tests: a critical survey. Brit.Joun.Ed. Psychol. 4, 2, 132-143.

FELKNOR, C. and HARVEY, O.J. (1964) Some Cognitive determinants of concept formation and concept attainment, Technical Report No.10, Contract Nonr. 1147 (07), University of Colorado.

FESTINGER, L. (1957) A theory of cognitive Dissonance. Evanston, IU: Row, Peterson.

FIEBERT, M. (1967) Sex differences in cognitive styles. Percept. Mot. Skills. 24, 1277-1278.

FIELD, T.W. & POOLE, M.E. (1971) Intellectual style and achievement of Arts & Science undergraduates, Bri. Journ. Ed. Psychol. 41, 338-41.

FILLENBAUM, S., & RAPPOPORT, A. (1971) Structures in the subjective Lexion; an experimental approach to the study of semantic fields. Academic Press.

FISHBEIN, G.M. (1963) Perceptual modes and asthmatic symptoms: an application of Witkin's hypothesis. J. Consult. Psychol. 27, 54-8.

FLAVELL, J.H. (1963) The Developmental Psychology of Jean Piaget. Van Nostrand, London.

FORGUS, R.H. (1966) Perception. McGraw-Hill, London & New York.

FORSTER, M. (1959) An Audit of Academic Performance - A study of the predictions of success in the Queen's University, Belfast. Belfast: Queen's University.

FOULKES, D., & FOULKES, S.M. Self-Concept, dogmatism, and tolerance of trait inconsistency. J. Person. Soc. Psychol. 2, 104-111.

FRENCH, V. (1947, 1948) The structure of sentiments. J. Personal. 15, 247-82, 16, 78-100, 209-44.

FRICK, J.W. & KEENER, H.E. (1956) A validation study of college achievement. Journal of Applied Psychology 40, 251-52.

FURNEAUX, W.D. (1956) Report to the Imperial College of Science and Technology. Quoted by Lynn (1959).

FURNEAUX, W.D. (1961) The Chosen Few. London: Oxford University Press.

GAIER, E.L. (1952) Personality and Students' thinking. School Review 60, 404-411.

GARDINER, J.M. & THOMPSON, G.P. (1973) Modality and recency effects in critical filled-delay free recall, Proceedings of the Annual Conference of the Brit. Psychol. Soc. April, 1974, p. 15.

GARDNER, R.W. (1953) Cognitive styles in categorising behaviour. J. Person. 22, 214-33.

GARDNER, R.W., JACKSON, D.M., & MESSICK, S.J. (1960) Personality, Organisation in cognitive controls and intellectual abilities. Psychol. Issues, 2, No. 4, whole No. 8.

GARDNER, R.W. & HOLZMAN, P.S. & KLEIN, G.S. LINTON, H.B. & SPENCE, D.P. (1959) Cognitive Control: a study of individual consistencies in cognitive behaviour. Psychol. Issues, 1, 4, and 8.

GARDNER, R.W. & SCHOEN, R.A. (1962) Differentiation, and abstraction in concept formation. Psychol. Mon. 76, no. 41, whole no. 560.

GETZELS, J.W. & JACKSON, P.W. (1962) Creativity and Intelligence, Wiley, New York.

GIBSON, C. & ABELSON, R.P. (1965) The subjective use of inductive evidence, J. Person. Soc. Psychol. 2, 301-10.

GLANZER, M. & CUNITZ, R.R. (1966) Two storage mechanisms in free recall. J. Verb. Learn. Verb. Beahviour.

GLANZER, M. & MEINZER, A. (1967) The effects of intra-list activity. J. Ver. Learn. Verb. Behav. 6, 928-

GLASS, D.G., (1968) Individual differences and the resolution of cognitive inconsistencies in Abelson, R.P. Aronson, E. McGuire, W.J. Newcomb, T.M. Rosenber, M.J., and Tannenbaum, P.H. (Ed). Theories of Cognitive Consistency: A sourcebook. Rand McNally.

GLIXMAN, A.F. (1965) Categorising behaviour as a function of meaning domain. J. Personal. Soc. Psychol. 2, 370-77.

GOEDINGHAUS, C.H. (1954) Studies ofthe Relationship between temperament and Academic achievement. M.A. Thesis. University of Southern California.

GOLDSTEIN, A.G. (1955) An experimental study of depth perception from the viewpoint of the sensory-tonic field theory of perception. Micro-filmed PhD Thesis. Clark University.

GOODENOUGH, D.R. & KARP, S.A. (1961) Field dependence and intellectual functioning. J. Abnorm. Soc. Psychol. 63, 2, 241-246.

GORDON, B. (1953) An experimental study of dependence-independence in a social and a laboratory setting. Unpublished Doctoral Dissertation Univ. of Southern California.

GORDON, O. BRAYER, R. & TIKOFSY, R. (1961) Personality variables and the perception of embedded figures. Percept. Mot. Skills, 12, 195-202.

GOUGH, H,G. & SANFORD, R.M. (1952) Rigidity as a psychological variable. Unpublished manuscript, University of California, Institute of Personality Assessment & Research.

GOULD, E.M. & M'COMISKY, J.G. (1958) Attainment level in leaving certificate and academic performance at university. Brit. Journ. Ed. Psychol. 28,2,129-134

GRACE, H.A. (1957) Personality factors and college attrition. Peabody Journal of Education, 35, 1.

GREGORY, R.L. (1966) Eye and brain. World University Library, London.

GRUEN, A. (1955) The relation of dancing experience and personality to perception. Psychol. Monog. 69, No.14.

GRUEN, A. (1957) A critique and re-evaluation of Witkin's perception and perception-personality work. J.Gen. Psychol. 56, 73-93.

GRUNEBERG, M.M. COLWILL, S.J. WINFROW, P. & Woods, R.W. (1970) Accoustic Confusion in long term memory. Acta Psychol. 32, 394-398.

GUILFORD, J.P. (1950) Creativity. Amer. Psychol. 5, 444.

HADDON, F.A. & LYTTON, H. (1968) Teaching approach and development of divergent thinking abilities in primary schools. Br.J.Educ.Psycho. 38, 171-180.

HAHN, H. & SINGER, E. (1949) An investigation of some factors related to success or failure of school of commerce students. Journ.Ed.Psych. 40.

HAMILTON, V. (1968) Scholastic and non-scholastic correlates of university students' academic performance. Paper presented at the Fourth Annual Conference of the Society for Research into Higher Education.

HARONIAN, F. & SUGARMAN, A.A. (1967) Fixed and mobile field-dependence. Review of studies relevant to Werner's dimension. Psychol.Reps. 21, 41-57.

HARTOG, P. & RHODES, E.C. (1935) An examination of examinations. London, Mcmillan.

HARVEY, O.J. (1962) Personality factors in resolution of conceptual incongruities. Sociometry, 25, 336-52.

HARVEY, O.J. (1963) Cognitive determinants of role playing. Technical Report, University of Colorado. No. 3.

HARVEY, O.J. (1964) Some cognitive determinants of influenceability, Sociometry, 27, 208-21

HARVEY, O.J. (1966) System structure, flexibility and creativity, in Harvey, O.J. (Ed) Experience, structure, and adaptability. Springer NY.

HARVEY, O.J. (1967) Conceptual Systems and Attitude change, in Sherif, C.W. & Sherif, M. (Eds) Attitude, Ego-Involvement, and change. Wiley pp 205-18, 222-6.

HARVEY, O.J. HUNT, D.E. & SCHRODER, H.M. (1961) Conceptual systems and personality development. Wiley.

HARVEY, O.J. REICH, J. & WYER, R.S. (1966) Affective and Personality determinants of differentiation. Unpublished Manuscript, Univ. Colorado.

HASAN, P. & BUTCHER, H.J. (1966) Creativity and Intelligence: a partial replication with Scottish children of Getzels and Jackon's study. Br. J. Psychol. 57, 129-135.

HART, I. (1957) Maternal child-rearing practices and authoritarian ideology. J. Abnorm.Soc.Psychol.55, 232-7.

HAYS, W.L. (1958) An approach to the study of trait implication and trait similarity in Taguiri, R. & Petrillo, L. Person. Perception & Interpersonal Behaviour. Stanford Univ. Press, 289-299.

HEAD, SirH. (1920) Studies in Neurology. Oxford.

HEBB, D. (1949) The organisation of Behaviour. London, Wiley.

HEIDER, J.F. (1961) Quart. J. Exper. Psychol. 13. 229.

HELLKAMP, D.T. & MARR, J.N. (1965) Dogmatism and field dependency. Percept. Mot. Skills. 20, 1046-48.

HENLE, M. (1962) On the relation between logic and thinking. Psych. Rev. 69, 366-78.

HIMMELWEIT, H.T. (1963). Student selection-implications derived from two student selection enquiries. Sociological Review. Monograph 7, 79-98.

HIMMELWEIT H.T. & SUMMERFIELD, A. (1951) Student selection: an experimental investigation. Brit.Journ.Psychol. 1, 328-342 and 2, 340-353.

HINKLE, D.N. (1965) The change of personal constructs from the viewpoint of a theory of implications. Unpublished PhD thesis, Ohio State Univ.

HOFFMEISTER, J.K. (1965) Conceptual determinants of strength and certainty of beliefs, Unpublished manuscript. Univ. of Colorado.

HOLLAND, J.L. (1959) The prediction of College grades from the California Psychological Inventory and the Scholastic Aptitude Test. J.Ed.Psych. 50, 135 245-254.

HOLLAND, J.L. (1960) The prediction of college grades from personality and aptitude variables. Journal Ed. Psych. 51, 245-254.

HOLLINGWORTH, H.S. (1913) Experimental studies in judgement. Arch. Psych. 29, 44-52.

HOLMES, F.J. (1960) Predicting academic success in a general college curriculum. Illinois: Institute of Personality and Ability Testing. Information Bulletin 4.

HOLZMAN, P.S. (1954) The relation of assimilation tendencies in visual auditory and kinaesthetic time-error to cognitive attitudes of levelling and sharpening. J. Person. 22, 375-94.

HOLZMAN, P.S. & KLEIN, G.S. (1951) Perceptual attitudes of 'leveling' and 'sharpening', relation to individual differences in time error, Amer.Psychol. 6, 257.

HONIGFELD, G. & SPIGEL, L.M. (1960) Achievement Motivation and field independence. J. Consult. Psychol. 24, 550-551.

HOPE, K. (1968) Methods of multivariate analysis. Univ. London Press.

HOVLAND, C.I. JANIS, I.J. & KELLEY, H.H. (1953) Communication and Persuasion. Yale Univ. Press.

HUBEL D.H. & WIESEL, T.N. (1959) Perceptive fields of single neurones in the cat's striate cortex. J. Physiol. (London) 148, 574-91.

HUBEL, D.H. & WIESEL, T.N. (1961) Integrative action in the cat's lateral geniculate body. J. Physiol (London), 155, 385-98.

HUBEL, D.H. & WIESEL, T.N. (1962) Receptive fields, binocular interaction and functional architecture in the cats' visual cortex. J. Physiol (London) 160, 106-54.

HUBEL, D.H. & WIESEL, T.N. (1963) Shape and arrangement of columns in cat's striate cortex. J. Physiol. (London) 165, 559-68.

HUBEL, D.H. & WIESEL, T.N. (1965) Receptive fields and functional architecture in the non-striate areas (18 & 19) of the cat. J. Neurophysiol, 28, 229-289.

HUDSON, L. (1966) Contrary Imaginations: A Psychological study of the English schoolboy. Penguin Books.

HUNT, R.A. (1967) Self and other semantic concepts in relation to choice of a vocation. J. Appl. Psychol. 5, 242-246.

JACKSON, D.M. (1955) Stability in resistance to field forces, Unpublished Doctoral Dissertation, Purdue Univ.

JACKSON, D.N. (1956) A short form of Witkin's Embedded Figures Test. J. Abnorm. Soc. Psychol. 53, 254-255.

JACKSON, D.N. (1957) Intellectual Ability and mode of perception. J. Consult. Psychol. 21, 458.

JACKSON, D.N. MESSICK, S.J. & MYERS, C.T. (1964) Evaluation of group and individual forms of embedded figures measures of field-independence. Educ. Psychol. Measmt. 24, (2) 177-192.

JACKSON, D.N. MESSICK, S.J. & SOLLEY, C.M. (1957) A multidimensional scaling approach to the perception of personality. J. Psychol. 44, 311-318.

JOHNSON, S.E. (1967) Hierarchical clustering schemes. Psychometrika, 62, 241.

JOHNSON, D.T. NEVILLE, C.W. & WORKMAN, S.M. (1969) Field independence and the sixteen personality factor questionnaire. Percept. Mot. Skills 28, 670.

JONES & McPHERSON (1972) Predicting first year university achievement. Times Educational Supplement, 12.5.72 & 19.5.72.

KAGAN J(1966) Developmental studies in reflection and analysis, in Kidd A.M. & Rivoire, J.L. (Eds) Perceptual Development in Children, Univ. of London Press.

KAGAN, J. & MOSS, M.A. (1962) From birth to Maturity, a study in Psychological development. Wiley.

KAGAN, J. & ROSMAN, B.L. (1964) Cardiac and respiratory correlates of attention and an analytic attitude. J. Exp. Child Psychol. 1, 50-63.

KAGAN, J. & MOSS, M.A. & SIGEL, I.E. (1963) Psychological significance of styles of conceptualisation in basic cognitive processes in children Mongr. Soc. Res. Child Devel. 28, No. 2.

KAGAN, J. ROSSMAN, B,L. DAY, D. ALBERT, J. & PHILLIPS, W. (1964) Information processing in the child. Psychol. Monogr. 78, No. 578,

KARLINS, M. & LAMM, H. (1967) Information search as a function of conceptual structure in a complex problem-solving task. J. Personal. Soc. Psychol. 5, 456-9.

KARP, S.A. & KONSTADT, N.L. (1965) Alcoholism and Psychological differentiation: long range effects of heavy drinking on field dependence. J. Nerv. Ment. Dis. 140, 412-6.

KARP, S.A. POSTER, D. & GOODMAN, A. (1963) Differentiation in alcoholic women. J. Personal. 31 386-93.

KELLEY, G.A. (1955) The psychology of personal construct. Norton.

KELMAN, H.C. & BARCLAY, J. (1963) The F.scale as a measure of breadth of perspective. J. Abnorm. Soc. Psychol. 67, 608-

KELSALL, R.K. (1963) University selection in relation to subsequent performance. A critical appraisal of the British evidence. In Halmos P (Ed) The Sociol. Rev. Monog. 7, 99-115, Univ. of Keele.

KELVIN, R.R. LUCAS, C.J. & OJHA, A.B. (1965) Personality mental health and academic performance. Brit. Journ. Soc. & Clin. Psych. 4, 244-253.

KESSEL, F.S. (1972) Imagery: a dimension of mind rediscovered. Brit. J. Psychol. 63, 2, 149-162.

KIEFERLE, D.A. & SECHREST, L. (1961) Effects in alterations in personal constructs, J. Psychol. Studies, 12, 173-8.

KINTSCH, W. (1970) Learning, Memory and conceptual processes. Wiley, N.Y.

KINTSCH, W. & BUSCHKE, H. (1969) Homophones and synonyms in short term memory. J. Exp. Psychol. 80, 402-407.

KISS, G.R. (1972) Long term memory: a state-space approach. Brit. J. Psych. 63, 327-341.

KLINE, P. & GALE, T. (1971) Extraversion, neuroticism and performance in a psychology examination. Bulletin of Brit. Psych. Soc. 24, 83, 151.

KOENIG, K.E. & McKEACHIE, W.J. (1959) Personality and independent study. J.Educ. Psychol. 50, 132-134.

KONSTADT, N. & FORMAN, E. (1965) Field dependence and external directedness. J. Pers. Soc. Psych. 1, 490-93.

KRECH, D. & CRITCHFIELD, R.S. (1948) Theory and problems of social psychology. McGraw-Hill.

KUFFLER, S.W. (1953) Discharge patterns and functional organisation of mammalian retina. J. Neurophysiol. Vol. 16 pp 37-68.

LAVIN, D.E. (1965) The prediction of academic performance. Russell Sage Foundation, and New York, Wiley.

LAZARUS, R.S. et al (1957) Personality and psychological stress. J. Person. 25, 559-557.

LEAGUE, B.J. & JACKSON, D.M. (1961) Activity and passivity as correlates of field independence. Percept. Mot. Skills, 12 291-298.

LEE, C.L. KAGAN, J. & RABSON, A. (1963) The influence of a preference for analytic categorisation upon concept acquisition. Child Dev. 34, 433.

LEE, R.E. & WARR, P.B. (1969) The development and standardisation of a balanced F scale. J. Gen. Psych. 81, 109-129.

LEFFORD, A. (1946) The influence of emotional subject matter on logical reasoning. J. Gen. Psychol. 34, 127-51.

LEVENTHAL, H. (1957) Cognitive processes and interpersonal predictions. J. Abnorm. Soc. Psych. 55, 176-180.

LEVENTHAL, H. & SINGER, D.L. (1964) Cognitive Complexity, impression formation and impression change. J. Pers. 32, 210-226.

LEVY, B.A. & MURDOCK, B.B. (1968) The effects of delayed auditory feed-back and intralist similarity in short term memory. J. Verb. Learn. Ver. Behav. 7, 887-894.

LEWIN, K. (1936) Principles of Topological Psychology. McGraw-Hill.

LEWIN, K. (1951) Field Theory in Social Science, Harper.

LEWIS, H.B. GOODENOUGH, D.R. SHAPIRO, A. & SLESER, I.(1966) Individual differences in dream recall. J. Abnorm. Psychol. 71, 52-9.

LINTON, H,B. (1952) Relation between mode of perception and tendency to conform. Unpublished Doctoral Dissertation, Yale University.

LINTON, H.B. (1955) Dependence on external influence, correlates in perception, attitudes and judgement. J. Abnorm. Soc. Psych. 51, 502-507.

LLOYD, B.B. (1972) Perception and cognition: a cross-cultural perspective; Penguin Books.

LOCKE, M. (1958) The relationship between the academic performance of university students and various cognitive and personality variables. Unpublished report cited by Rushton (1969).

LOCKE, E.A. (1962) The development of criteria of student achievement Educational and Psychological measurement. 23, 2, 299-307.

LUCAS, C.J. KELVIN, R.P. & OJHA, A.B. (1966) Mental health and student wastage. Brit. Journ. of Psychiatry 112, 277-284.

LURIA, A.P. (1966) Higher Cortical Functions in Man. Tavistock, London.

LYNN, R. & GORDON, J.E. (1961) The relation of neuroticism and extra-version to intelligence and attainment. Brit. Journal of Educ. Psych. 31, 194-203

LYTTON, H. & COTTON, A.C. Divergent thinking abilities in secondary schools, Br. J. Educ. Psych. 39, 188-190.

MACLAY, I. (1968) A random sample of university undergraduates. Universities Quarterly, 23, 80-94.

MACKINNON, W.J. & CENTERS R. (1956) Authoritarianism and urban stratifications. Amer. J. Soc. 61, 610-620.

MCANDLESS, B.R. (1949) The Rorschach as a predictor of academic success. J. Applied Psych. 33, 43-50.

McCLELLAND, D.C. & ATKINSON, J.W. (1948) The projective expression of needs. The effect of the different intensities of the hunger drive on perception. J. Psych. 25, 205-222.

McCULLOCH, J.W. & FOREN, R.J.W. & HITCH, P.J. (1969) Motivation and academic performance as factors in the selection of social work students J. Educ. Research 2, 207-11.

McCULLOUGH, C. & VAN ATTA, E.L. (1960) Experimental evaluation of teaching programmes utilising a block of individual work, unpublished MS, Oberlin College, Ohio.

McGUINIES, E. (1949) Emotionality and perceptual defence. Psych Rev. 56 244-251.

McKEACHIE, W.J. (1961) Motivation, teaching methods and college learning. In Jones, M,R. (Ed Nebraska Symposium on Motivation pp 111-141. Lincoln Univ. of Nebraska Press.

McKINNON, D.W. (1962) The Nature of creative talent. Am. Psych. 17, 484-495.

McLEISH, J. (1968) The lecture method, Cambridge Institute of Education.

MCQUARY, J.P. (1953) Some relationships between non-intellectual characteristics and academic achievement. J.Ed.Res. 4, 44, 215-28.

McWHINNIE, H.J. (1967) Some relationship between creativity and perception in sixth grade children. Percept. Mot. Skills, 25, 979-980.

MAIER, N.R.F. (1930) Reasoning in humans. I On direction. J.Comp. Psych. 10, 115-43.

MAIER, N.R.F. (1931) Reasoning in humans II The solution of a problem and its appearance in consciousness. J. Comp. Psych. 12, 181-94.

MANDLER, G. (1967) in Spence, K.W. & Spence, J.T. (ed) The psychology of learning and motivation: advances in research and theory. Vol. I p 327. Academic Press, NY & London.

MANDLER, G. (1971) Paper presented at a conference on organisation in memory, Univ. of Pittsburgh. 4-5 March 1971.

MARLOWE, D.(1958) Some psychological correlates of field independence J. Consult. Psychol. 22, 334.

MORRIS, P. (1964) The Experience of Higher Education. London: Routledge & Kegan Paul.

MARRS, C.L. (1955) Categorising behaviour as elicited by a variety of stimuli, unpublished Master's thesis. Univ. of Kansas.

MASSARO, D.W. (1970) Perceptual auditory images. J. Exp. Psych. 85, 411-417.

MAYO, C.W. & CROCKETT, W.M. (1964) Cognitive complexity and primacy-recency effects in impression formation. J. Abnorm. Psych. 68, 335-338.

MELTON, A.W. (1963) Implications of short term memory for a general theory of memory. J. Veb. Learn. Veb. Behav. 2, 1-21.

MERLEAU-PONTY M. de (1942) The structure of behaviour. London. Methuen.

MESSICK, S. & DAMARIN, F. (1964) Cognitive styles and memory for faces. J. Abnorm. Soc. Psych. 69, 313-318.

MESSICK, S. & FREDERICKSEN, J. (1958) Ability, acquiescence, and "authoritarianism", Psych. Rep. 4, 687-697.

MESSICK, S. & FRITZKY, F.J. (1963) Dimensions of analytic attitude in cognition and personality. J. Pers. 31, 346-370.

MEWHORT, D.J.K. (1966) Sequential redundancy and letter spacing as determinants of tachistoscopic recognition. Can.J. Psych. 20, 435.

MILLER G.A. (1956) The magical number seven, plus or minus two. Psych. Rev. 63, 91.

MILLER, G.A. (1969) A psychological method to investigate verbal concepts J. Mathematical Psychol. 6, 169-71.

MILLER, G. GALLANTER, E. & PRIBRAM, K. (1960) Plans and the structure of behaviour. New York. Holt.

MILLER G. & ROKEACH, M. (1968) Individual differences and tolerance for inconsistency in Abelson et al. (1968).

MILLER, G. & SELFRIDGE, J. (1950) Verbal context and the recall of meaningful material. Am.J.Psych. 63, 176.

MINARD, J.G. (1965) Response-bias interpretation of perceptual defence. Psych. Rev. 72, 74-88.

MOGAR, R.E. (1960) Three versions of the F Scale and performance on the semantic differential, J. Abnorm. Soc. Psych. 60, 262-5.

MONALTO, F.D. (1946) An application of the Group Rorschach technique to the problem of achievement in college. J. Clin. Psych. 2, 254-280.

MORTON, J. (1969) Interaction of information in word recognition. Psych. Rev. 76, 2.

MORRELL, F. (1967) Electrical signs of sensory coding. The neurosciences a study program, Rockefeller Univ. Press. pp 452-69.

MUNROE, R.L. (1942) An experiment in large scale testing by a modification of the Rorschach method. J. Psych. 13, 229-263.

MUNROE, R.L. (1945) Prediction of the adjustment and academic performance of college students. Applied Psych. Mono. 7.

MURDOCK, B.B. (1961) The retention of individual items. J. Exp. Psych. 62, 618.

MURDOCK, B.B. (1962) The serial position effect of free recall. J.Exp. Psych. 64, 482-88.

MURDOCK, B.B. (1967) Recent developments in short term memory. Brit. J. Psychol. 58, 421-433.

MURDOCK, B.B. (1972) in Bower, G.H. & Spence, J.T. (eds) The Psychology of Learning and Motivation. Vol. 5. Academic Press, N.Y.

MUSSEN, P.H. & KAGAN, J. (1958) Group Conformity & Perceptions of Parents. Child Devt. 29, 57-60.

NEISSER, U. (1967) Cognitive Psychology. New York. Appleton-Century-Crofts.

NEMZEK, C.L. (1938) The value of the Bernreuter Personality Inventory for direct and differential prediction of academic success as measured by teachers' marks. Journal of Applied Psych. 22, 576-586.

NIDORF, J. (1961) Individual differences in impression formation. Unpublished Doctoral Thesis. Clark University.

NISBET, J. & BUCHAN, J. (1959) The long term follow-up of assessments at age eleven. British J. Ed.Psych. 29. 1, 1-8.

NORMAN, D.A. (1969) Memory and Attention: an introduction to human information processing. Wiley, N.Y.

NORMAN, D.A. (1970) (Ed) Models of human memory, Academic Press, N.Y. and London.

NUTTALL, D.L. (1972) Convergent and Divergent Thinking. In Butcher, H.J. & Pont, H.B. (eds) Educational Research in Britain, Vol. 3, London UCP.

O'DONOVAN, D. (1965) Rating extremity: pathology or meaningfulness? Psych. Rev. 72, 358-372.

OHNMACHT, F.W. (1966) Note on the validity of the Sixteen Personality Factor Questionnaire measure of field independence. Percept.Motor Skills 27, 564.

OSBORN, A.F. (1953) Applied imagination. Scribner.

OSBORNE, R.T. & SANDERS, W.B. (1949) Multiple choice Rorschach responses of college achievers and non-achievers. Education and Psychological Measurement. 9, 685-691.

OSGOOD, C.E. (1968) Interpersonal verbs and interpersonal behaviour, Technical Report No. 64, Institute of Communications research, Urbana, Univ. of Illinois.

OSGOOD, C.E. SUCI, G.J. & TANNENBAUM, P.H. (1957) The measurement of meaning. Univ. of Illinois Press.

OTSUKA, R. & HASSLER, R. (1962) Uber Aufbau und Gliederung der corticales Sehspare bei der Katze. Arch. Psychiatr. Nervenkrantz 203, 212-34.

OWENS, W.C. & JOHNSON, W.C. (1949) Some measured personality traits of collegiate under-achievers. J. Educ. Psych. 40, 41-46.

PAIVIO, A. (1969) Mental Imagery in associative learning and memory. Psych. Rev. 76, 241-263.

PAIVIO, A. & CSAPO, K. (1969) Concrete image and verbal memory codes. J. Exp. Psych. 80, 279-285.

PARDES, H. & KARP, S.A. (1965) Field dependence in obese women. Psychoson. Med. 27, 238-44.

PASTORE, N. (1949) Need as a determinant of Perception. J. Psych. 28, 457-475.

PEAK, H. (1958) Psychological structure and person perception in R. Tagiuri and L. Petrullo, (Eds). Person perception and inter-personal behaviour. Stanford Univ. Press.

PEAK, H. MUNEY, B. & CLAY, M. (1960) Opposites structures, defences and attitudes, Psychol. Monogr. 748.

PEMBERTON, C.L. (1952) The closure factors related to other cognitive processes. Psychometrika, 17, 267-288.

PETCH, J.A. (1959) Before admission. Liverpool Univ.Press.

PETERSON, H. GUILFORD, J.P.M. & HOEPFNER, R. (1963) Determination of 'structure of intellect' abilities involved in ninth grade algebra and general mathematics. Rep. No. 3 Apt. Res. Proj. Univ. of S. Calif.

PETTIGREW, T.F. (1958) The measurement and correlates of category width as a cognitive variable. J. Personal. 26, 532-44.

PETRIE, A. (1948) The selection of medical students. Lancet 255.

PIAGET, J. (1955) Recherches sur le developpement des perceptions XXII Arch. Psychol. Geneve 35, 1-24.

PIAGET, J. (1971) Biology and Knowledge. Edinburgh Univ. Press.

PIAGET, J. ALBERTINI, B. & ROSSI, M. (1944) Arch. Psychol. Geneve, 30 95-138.

PIAGET, J. & LAMBERCIER, P. (1942-58) "Recherches.." series. Arch. Psychol. Geneve. 1942-58.

PIAGET, J. VIN-BANG, & MATALON, B. (1958) Note on the law of the temporal maximum of some optico-geometric illusions. Amer. J. Psych. 71, 277, 282.

PILKINGTON, G.W. & HARRISON, G.J. (1967) The relative value of two high level intelligence tests, advanced level and first year university examinations marks for predicting degree classification. Brit.J. Ed. Psych. 37, 3 382-389.

PILLINER, A.E.G. (1960) Prediction of success in the Arts Faculty of the University of Edinburgh by Marks obtained at the Scottish Leaving Certificate Examination. Edinburgh University Department of Education.

POLANYI, M. (1961) Personal knowing. Routledge & Kegan Paul. London.

POLLACK, M. KAHN, R.L. KARP, S.A. & FINK, M. (1960) Individual differences in the perception of the upright in hospitalised psychiatric patients, paper read at the Eastern Psychological Assoc. N.Y.

POLLEY, E.H. & DIRKES, J.M. (1963) The visual cortical (geniculo-cortical) area of the cat brain and its projections. Anat. Record. 145 p 345.

POLYAK, S. (1927) An experimental study of the associtiion callosal, and projection fibres of the cerebral cortex of the cat. J.Comp. Neurol, 44, 197-258.

POPPER, K.R. (1959) The Logic of Scientific Discovery. Hutchinson, London.

POPPER, K.R. (1972) Objective Knowledge, Oxford Univ. Press.

POSTMAN, L. (1955) Witkin et al (1954) Personality through perception. Psych. Bull, 52, 79-83.

RAVENETTE, A.T. (1969) The situations grid: a further development in grid techniques with children, unpublished manuscript.

RAYMOND, B. (1969) J. Verb. Learn. Verb. Behav. 8, 567-

ROBERTS, A.H. & ROKEACH, M. (1956) Anomic, authoritarianism and prejudice: a replication. Amer. J. Sociol, 61, 355-8.

ROGERS, G.W. (1960) Lecture Listening skills: their nature and relation to achievement. Dissertation Abstracts 20, 4165.

RICHARDS, J.P.G. & WILSON, A.J.C. (1961) 'A' level and pass degree in Physics. Universities Quarterly, 15, 389-391.

RICHARDS, P.N. & BOLTON, N. (1971) Divergent thinking, mathematical ability and type of mathematics teaching in junior school children. Br. J. Educ. Psych. 41, 32-37.

RIEDEL, W.W. (1970) An investigation of personal constructs through non-verbal tasks. J. Abnorm. Psych. 76, 173-9.

ROKEACH, M. (1954) The nature and meaning of dogmatism. Psych. Rev. 61, 194-204.

ROKEACH, M. (1960) The open and closed mind. New York: Basic Books.

ROSENTHAL, R. (1966) Experimenter Effects in Behavioural Science N.Y. Appleton-Century-Crofts.

RUDIN, S.A. & STAGNER, R. (1953) Figure-ground phenomena in the perception of physical and social stimuli. J. Psych. 55, 213-25.

RUST, R.M. & RYAN, F.J. (1953) The relationship of some Rorschach variables to academic behaviour. Journal of Educational Psychol. 29, 573-580.

RYLE, A., & LUNGHI, M. (1968) A psychometric study of academic difficulty and psychiatric illness in students. Brit. J. of Psych. 114, 57-63.

SACKS, J.S. (1967) Recognition memory for syntactic and semantic aspects of connected discourse. Percept. Psychophys. 2, 437-443.

SAUNDERS, M. et al (1969) Report of the Commission on Teaching in Higher Education, London, National Union of Students.

SAVAGE, R.D. (1962) Personality Factors and academic performance. Brit. J. Educ. Psych. 32, 251-253.

SCHILLER, P.H. (1965) Monoptic and dioptic visual masking by patterns and flashes. J. Exp. Psych. 69, 193-199.

SCHOFIELD, W. (1953) A study of medical students with the M.M.P.I. Personality and academic success. J. Applied Psych. 37, 47-52.

SCHONBAR, R.A. (1964) Some dimensions of sensitivity of recallers and non-recallers of dreams, in Dream Research and Theory, Symposium presented at the Postgraduate Centre for Mental Health.

SCHRODER, H.M. DRIVER, J.J. & STREUFERT, S. (1967) Human Information Processing. Wiley.

SCOTT, W.A. (1962) Cognitive structure and social structure, in Washburne, N.F. (Ed) Decisions, Values, and Groups. Vol. 2, Pergamon.

SCOTT, W.A. (1963) Conceptualising and measuring Structural properties of cognition, in Harvey, O.J. (ed) Motivation and Social Interaction: Cognitive Determinants, Ronald Press.

SCOTT, W.A. (1965) Psychological Structure and Social Correlates of International Images, in Kelman, H.C. (ed) International Behaviour, Holt, Reinhart & Winston.

SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION (1936) The Prognostic Value of the University Entrance Examination. Univ. of London Press.

SECHREST, L. & JACKSON, D.N. (1961) J. Personal. 29, 167-82.

SHAEFFNER, E.A. & BAYLEY, N. (1963) Maternal behaviour, child behaviour, and their intercorrelations from infancy through adolescence. Monogr. Soc.Res.Child Devel. 28, No. 3.

SHALLICE, T. & WARRINGTON, E.K. (1970) Independent functioning of verbal memory stores; a neuropsychology study. Quart. J. exp. Psychol. 22 261-273.

SHAPIRO, R.J. (1968) Creative Research Scientists. Psychologica Africana Monagr. 4.

SHERWIN, E. & CHILD, D. (1970) Predicting the performance of under-graduate chemistry students. Education in Chemistry, 7 (4) 156-158.

SHIFFRIN, & ATKINSON (1969) Storage and retrieval processes in long term memory. Psychol. Rev. 76, 2, 179-193.

SHOUKSMITH, G. (1970) Intelligence, creativity and cognitive style, Batsford, London.

SIEBER, J.E. & LANZETTA, J.T. (1964) Conflict and conceptual structure as determinants of decision-making behaviour. J. Person. 32, 622-41.

SIEGEL, S. (1955) Non-Parametric Statistics for the Behavioural sciences. McGraw-Hill, New York.

SLOANE, H.N. (1959) The generality and construct validity of equivalence range, Unpublished Doctoral Dissertation, Pensylvania State Univ.

SMAIL, D.J. (1970) Neurotic symptoms, personality, and personal constructs. Brit. J. Psychiatr. 117, 645-8.

SMALL, M.A. (1973) Final Report of Investigation into the Use of Student Project of Colleges of Further Education, Social Science Research Council, April, 1973.

SMITHERS, A. (1970a) Some factors in lecturing. Educ. Rev. 22, 141-150.

SMITHERS, A. (1970b) Open-Mindedness and the University Curriculum, J. Curriculum Studies, 2, 73-77.

SMITHERS, A.G. & BATCOCK, A. (1970) Success and failure among social scientists and health scientists at a technological university. Brit. Journ. of Educ. Psych. 40, 144-153.

SOKAL, R.R. (1966) Numerical Taxonomy, Scientific American, Dec. 1966, 106-116.

SOKAL, R.R. & SNEATH, P.H. (1963) Principles of Numerical Taxonomy.

SPERLING, G. (1960) The information available in brief visual presentations. Psych. Monog. 74, 11.

SPERLING, G. (1967) Successive approximations to a model for short term memory. Acta Psych. 27, 285-293.

SPINELLI, D.N. (1966) Visual receptive fields in the cat's retina; complications. Science 152, 1768-9.

SPOTTS, J.V. & MACKLER, B. (1967) Relationship of field-dependent and field independent cognitive styles to creativity test performance. Percept. Mot. Skills. 24, 239-268.

STAGNER, R. (1961) Psychology of personality. McGraw-Hill.

STEINER, I.D. (1954) Ethnocentrism and Tolerance of trait inconsistency. J. Abnorm. Soc. Psychol. 49, 349-354.

STEINER, I.D. (1964) Galvanic skin resistance and responses to interpersonal disagreements. Unpublished progress report, United States Health Service Grant M-4460, Univ. of Illinois.

STEINER, I.D. & JOHNSON, H.H. (1963) Authoritarianism and tolerance of train inconsistency. J. Abnorm. Soc. Psychol. 67, 388-391.

STEINER, I.D. & McDIARMID, C.G. Two kinds of assumed similarity between opposites. J. Abnorm. Soc. Psychol. 555, 140-142.

STERNE, L. (1760) The Life and Opinions of Tristram Shandy; Gentlemen, Folio Society, London, 1970.

STONE, D.R. & GANLING, G.R. (1956) A study of scholastic achievement related to personality as measure by the MMPI Journal of Educational Research, 50, 155-156.

STREUFERT, S. & CASTORE, C.H. (1971) Information search and the effects of failure: a test of complexity theory, J. Exper. Soc. Psychol. 7, 125-143.

STREUFERT, S. KLIGER, S.C. CASTORE, C.H. & DRIVER, M.J. (1967) Tactical and negotiations game for analysis of decision integration across decision areas. Psychol. Rep. 20, 155-70.

STREUFERT, S. SUEDFELD, P. & DRIVER, M.J. (1965) Conceptual structure, information search, and decision making. J. Personal. Soc. Psychol. 2. 736-40.

STRONG, E.K. (1943) Vocational Interests of Men and Women. Stanford Univ. Press.

STRONG, E.K. (1955) Vocational Interests 18 years after college. Minneapolis, University of Minnesota Press.

Student's Society Committee of the Royal Dental Hospital School of Dentra Surgery (1966). Report on Opinion Poll on the lecture courses. Supplement to Extract 33, Part I.

SUEDFELD, P.M. & STREUFERT, S. (1966) Information search as a function of conceptual and environmental complexity. Psychonomic Science. 4, 351-2.

SUPER, D.E. (1942) The Bernreuter Personality Inventory. Psychol. Bull. 39. 94-125.

SUPER, D.E. (1957) Vocational Development: a framework for Research, New York, Teachers College Press. Columbia University.

SUPNICK, J. (1964) Differences in univalence and extremity of impressions of others as a function of personal and situational variables. Unpublished Master's thesis. Clark University.

TAJFEL, H. RICHARDSON, A. & EVERSTINE, L. (1964) Individual consistencies in categorising. A study in judgemental behaviour. J. Person. 32 90-108.

TALBOT, S.A. & MARSHALL, W.H. (1941) Physiological studies on neural mechanisms of visual localisation and discrimination. Am. J. Ophthalmol. 24, 1255-64.

TAYLOR, J.M. (1956) A comparison of delusional and hallucinatory individuals using field dependency as a measure. Unpublished Doctoral Dissertation, Pardue Univ.

THOMPSON, J.B. (1973) An interview approach to the assessment of academic achievement motivation. Proceedings of the Annual Conference of the British Psychological Society. April 1973.

THOMPSON, J.M. WOOLSEY, C.M. & TALBOT, S.A. (1950) Visual areas I and II or cerebral cortex of rabbit. J. Neurophysiol, 13, 277-88.

TORGERSON, W.S. (1958) Theory and Methods of Scaling, Wiley.

TREISMAN, A. (1961) Attention and Speech, Unpublished doctoral thesis, Univ. of Oxford.

TRIPODI, T. & BIERI, J. (1963) Cognitive complexity as a function of own and provided constructs. Psych. Rep. 13, 26.

TRIPODI, T. & BIERI, J. (1964) Information transmission in clinical judgements as a function of stimulus dimensionality and cognitive complexity. J. Person. 32, 119-137.

TULVING, E. (1962) Subjective organisation in free recall. Psych. Rev. 69, 344-354.

TULVING, E. & COLOTLA, V.A. (1970) Free recall of tri-lingual lists. Cog. Psych. 1, 76-78.

TULVING, E. & MADIGAN, S.A. (1970) Memory and verbal learning. An. Rev. Psych. 21, 437-477.

TULVING, E. & PEARLSTONE, Z. (1966) Availability vs accessibility of information in memory for words. J. Verb. Learn. Verb. Behav. 5, 381.

TURVEY, M.T. (1967) Repetition and the preperceptual information store. J. Exp. Psych. 74, 289-293.

TURVEY, M.T. & KRAVETZ, S. (1970) Retrieval from iconic memory with shape as the selection criterion. Percept. Psych. 8, 171-172.

ULEHLA, Z.J. (1961) Individual differences in information yields of raters. Unpublished Master's thesis, Univ. of Colorado.

UNDERWOOD, B.J. (1949) Proactive inhibition as a function of time and degree of prior learning.

VALENTINE, C.W. (1961) The use of a new reasoning test for the selection of university and training college students. Brit.J. Ed. Psych. 31, 1, 227-231.

VANNOY, J.S. (1965) Generality of cognitive complexity as a personality construct. J. Personal. Soc. Psych. 2, 385-396.

VASTOLA, E.F. (1961) A direct pathway from lateral geniculate body to the association cortex. J. Neurophysiol, 24, 469-87.

VAUGHT, G.M. (1965) The relationship of roleidentification and ego strength to sex differences in the Rod and Frame Test. J. Pers. 33, 271-283.

VICK, O.C. & JACKSON, D.M. (1967) Cognitive styles in the schemating process; a critical evaluation, Educational and Psychological Measurement, 27, 267-86.

VON FIANDT, K. (1966) The World of Perception. Homewood, I.U. Dorsey.

VROOM, V.H. (1957) Effect of design on estimation of size of coins, Cand. J. Psych. 89-92.

WALLOCK, M.H. & KOGAN, N. (1965) Modes of thinking in Young Children. New York, Holt Rinehart & Winston.

WANKOWSKI, J. (1968) Some aspects of motivation in success and failure at university. Paper presented at the Fourth Annual Conference of the Society for Research into Higher Education.

WAGNER, S. (1964) Some aspects of a research program based on an organismic-developmental approach to cognition: Experiments and Theory, J. Amer. Academy of Child Psychiatry. 3, 193-212, 226-230.

WARBURTON, F.W. (1952) The selection of university students. Manchester Univ. Press.

WARBURTON, F.W. (1963) Analytic Methods of factor rotation. Brit. J. of Statistical Psych.

WARD, J.N. (1965) Group study versus lecture-demonstration method in a physical science introduction for general education college students. J. Exp. Psych. 24, 197-210.

WASON, P.C. (1969) Structural simplicity and psychological complexity; some thoughts on a novel problem. Bull. Brit. Psych. Soc. 22, 281-284.

WASON, P.C. (1969) Regression in Reasoning? Br.J. Psychol. 60, 471-480.

WASON, P.C. & JOHNSON LAIRD, P.N. (1972) Psychology of Reasoning: structure and content, London, Batsford.

WAUGH, N.C. & NORMAN, D.A. (1965) Primary Memory. Psych. Rev. 72, 89-104.

WEISS, R.J. SEGAL, B.E. & SOKOL, R. (1965) Epidemiology of emotional disturbance ina men's college. Journal of Nervous Disease. 141, 2.

WERNER, H. (1940) Comparative Psychology of Mental Development. Harper.

WERNER, H. & WAPNER, S. (1949) Symbol formation: an organismic-developmental approach to Language and the expression of thought. New York, Wiley.

WERNER, H. & WAPNER, S. (1956) The Innsbruk studies on distorted visual fields in relation to an organismic theory of perception. Psych. Rev. 62, 130-138.

WERNER, H. & WAPNER, S. (1956) Sensory-tonic field theory of perception basic concepts and experiments. Revista di Psicologia, 50, 315-337.

WERTHEIM, J. & MEDNICK, S.A. (1958) The achievement motive and field independence. J. Consult. Psych. 22, 38.

WHITE, B.J. ALTER, R.B. & RARDIN, M. (1965) Authoritarianism, dogmatism, and the use of conceptual categories. J.Per.Soc.Psych. 2, 293-295.

WHITE, B.J. & HARVEY, O.J. (1965) Effects of personality and own stand on judgement and production of statements about a central issue, J. Exp. Soc. Psych. 1, 334-347.

WIGGINS, B.R. (1968) Personality Structure, Ann. Rev. of Psych. 293-350.

WILLIAMS, E.M. (1950) An investigation of the value of higher school certificate results in predicting performance in first-year examinations. Brit. J. Ed. Psych. 20, 83-98.

WILSON, J.D. (1971) Predicting levels of 1st year university performance. Brit. J. Ed. Psych.

WINESTINE, M.C. (1964) Twinship and psychological differentiation. Unpublished Doctoral dissertation, New York, Univ.

WISHART, D. (1969) Clustan IA: User's Manual. Univ. of St. Andrew's Computing Laboratory.

WITHERSPOON, P. & MELBERG, M.E. (1959) The relationship between grade point averages and sectional scores on the Guilford-Zimmerman Temperamental Survey. Educational and Psychological measurement, 19, 673-674.

WITKIN, H.A. (1950) Individual differences in ease of perception of embedded figures. J. Pers. 19 1-15.

WITKIN, H.A. (1965) Psychological differentiation and forms of pathology. J. Abnorm. Psych. 70, 317-336.

WITKIN, H.A. DYK, R.B. FATERSON, H.F. GOODENOUGH, D.R. & KARP, S.A. (1962) Psychological differentiation. New York: Wiley.

WITKIN, H.A. LEWIS, H.B. HERTZMAN, M. MACHOVER, K. MEISSNER, P.B. & WARNER, S. (1954) Personality through perception. New York: Harper.

WRIGHT, J.M. (1968) Selection in visual immediate memory. Quart. J. Exp.Psych. 20, 62.

WULF, F. (1938) Tendencies in figural variation, in Ellis, W.D. (ed) a Source Book of Gestalt Psychology, Routledge & Kegan Paul.

WYER, J.A. (1964) Assessment and correlates of cognitive differentiation and integration. J. Personal. 32, 495-509.

YOUNG, H.H. (1959) A test of Witkin's field dependence hypothesis. J. Abnorm. Soc. Psych. 59, 188-192.

ZAJONC, R.B. (1960) The process of cognitive tuning in communication, J. Abnorm. Soc. Psych. 61, 159-67.

ZIMBARDO, P.G. & DWORKIN, L. (1964) A technique for the objective measurement of behaviour influenced by delayed auditory feedback. Unpublished Manuscript. New York University.

ZUCKMAN, L. (1957) Hysteric compulsive factors in perceptual organisation. Unpublished Doctoral Dissertation, New School for Social Research.

Appendix I

Experiment - Measures

Truth-Functional Tests for the Measurement of the Attainment of Educational Objectives in the Course On Human Thinking

T - F Tests 1 and 2 are developed specifically for this course from the Truth-Functional Test suggested by Bligh (1971). Each test has questions sampling from six levels of educational objective, as classified on Bloom's taxonomy of cognitive objectives. The sampling is as follows :

Type		Number of questions
3	Recognition of key point/principle	8
4	Memory of simply stated relationships	8
5	Making of simple inferences	8
6	Application	5
7	Analytic Thought	2
8	Synthesis	2

The student responds to the test by indicating agreement, disagreement, or don't know, to a number of statements. If he answers correctly, he scores 2 points. Incorrect answers score zero, and "don't know" scores 1 point. The maximum possible score and the "guessing score" for each "type" are as follows :

Type	3	4	5	6	7	8	Total
Maximum	16	16	16	10	4	28*	90
Guessing	8	8	8	5	2	14*	45

* Two questions, each with 7 points

T-F 1 and T-F 2 are presented as filled in by the "ideal" student. The number at the extreme left indicates the "type" of educational objective which the question samples

Appendix II

Field Study - Measures

BALANCED F SCALE

R.E. Lee and P.B. Warr

Name.............................. Sex ....

Age............................... Occupation ....

INSTRUCTIONS

Please read through the statements listed below. You will probably agree with some of them, and disagree with others. You can signify this agreement or disagreement by putting a tick in the appropriate division of the scale.

Example. Agree very much — Disagree very much

It is virtually impossible to go to a film without seeing violence. : : : : : : : :

If you wholeheartedly agree with this statement, you should tick the left-hand division. If you agree only moderately, tick the next division to the right. If you are indifferent, but if pushed, would disagree, you should check the third division from the right.

The statements below are similar to the example: please read through each statement and try to give an honest opinion which is your own, rather than putting what you think the marker would like to hear.

Agree very much. — Disagree very Much

1. The minds of today's youth are being hopelessly corrupted by the wrong kind of literature. : : : : : : : :

2. The worst danger to Britain during the last 50 years has come from foreign ideas and agitations. : : : : : : : :

3. As young people grow up, they ought to try to carry out some of their rebellious ideas and not be content to get over them and settle down. : : : : : : : :

4. One of the greatest threats to the true British way of life is for us to resort to the use of force. : : : : : : : :

5. Most censorship of books or films is a violation of free speech and should be abolished. : : : : : : : :

6. Unless something drastic is done the world is going to be destroyed one of these days by nuclear explosion or fallout. : : : : : : : :

7. It is only natural and right for each person to think that his family is better than any other. : : : : : : : :

8. There is a divine purpose in the operations of the universe. : : : : : : : :

9. Science declines when it confines itself to the solution of immediate practical problems. : : : : : : : :

10. No person who could ever think of hurting his parents should be permitted in the society of normal decent people. : : : : : : : :

11. We should be grateful for leaders who tell us exactly what to do and how to do it. : : : : : : : :

12. Sex crimes, such as rape and attacks on children, deserve more than mere imprisonment: such criminals ought to be publicly whipped or worse. : : : : : : : :

13. The church has outgrown its usefullness and should be radically reformed or done away with. : : : : : : : :

14. Disobedience to the government is sometimes justified. : : : : : : : :

15. It usually helps the child in later years if he is forced to conform to his parents' ideas. : : : : : : : :

16. It is the duty of a citizen to criticize or censure his country whenever he considers it to be wrong. : : : : : : : :

17. Few weaknesses or difficulties can hold us back if we have enough will power. : : : : : : : :

18. The poor will always be with us. : : : : : : : :

19. What a youth needs most is the flexibility to work and fight for what he considers right personally even though it might not be best for his family and country. : : : : : : : :

20. A world government with effective military strength is one way in which world peace might be achieved. : : : : : : : :

21. Members of religious sects who refuse to salute the flag or bear arms should be treated with tolerance and understanding. : : : : : : : :

22. In the final analysis parents generally turn out to be right about things. : : : : : : : :

23. The facts on crime and sex immorality suggest that we will have to crack down harder on some people if we are going to save our moral standards. : : : : : : : :

24. Divorce or annulment is practically never justified.	: : : : : : :
25. One way to reduce the expression of prejudice is through more forceful legislation.	: : : : : : :
26. Honesty, hard work and trust in God do not guarantee material rewards.	: : : : : : :
27. Army life is a good influence on most men.	: : : : : : :
28. An insult to our honour should always be punished.	: : : : : : :

For Office Use Only:

CODE [ ] GROUP [ ] SCORE [ ]

Name .... Age .... Sex ....

Instruction:

On the following page there is a 10 x 10 grid. Down the side are listed a number of people who are likely to be important in your life. Before going any further, please fill in their names where it is appropriate.

Across the top are a number of personality dimensions: I would like you to rate each person on each personality dimension, using the number scale at the top.

To take an example, if you think that you yourself are a very interesting person, mark +3 in the top left hand box, where "interesting", "dull" and "yourself" coincide. If you think that A.N. Other - the person you dislike - is moderately dull, but not exceptionally dull, mark -2 in the next box below.

N.B: 1. Please do not change the name of one of the people half way through the judgements, i.e. if you decide that 'A' will do as "friend of opposite sex", do not decide that you really prefer 'B' half way through the personality dimension.

2. Work through the grid systematically and fill in every box. Work fairly quickly, but with care. Most people finish this test in little over 10 minutes.

If you have any questions, feel free to ask the supervisor any time you want.

P.C.T.

Name: Age: Sex:

Below are given a series of words or phrases, which are intended to begin a paragraph. Follow on from each word or phrase and complete a paragraph on that subject in not more than 60 words. There are no correct answers in this test: try to sum up your own ideas or feelings on the subject.

There is no time limit: take as long as you want.

FOR OFFICE USE ONLY

CODE: TOTAL SCORE:

1. "When I am in doubt..........

2. "Confusion..........

3. "Rules..........

4. "When I am criticised..........

Course on Human Thinking

This term you have tackled the subject matter on "problem-solving" by a form of project work. Project work is rarely used in education for this type of information, although it is, of course, more used where practical work is concerned.

It would be helpful if you could indicate whether this particular project work, as compared with the typical lecture course,

	6	5	4	3	2	1		
increases	____	____	____	____	____	____	decreases	interest
	____	____	____	____	____	____		memory
	____	____	____	____	____	____		comprehension
	____	____	____	____	____	____		motivation to work
	____	____	____	____	____	____		ability to use the information outside the course
	____	____	____	____	____	____		inter-relatedness of different parts of the course

Course on Human Thinking

To help me adapt my teaching to your needs, please give your estimate, of this half of the course, on a six-point scale, by putting a tick in the appropriate column.

		6	5	4	3	2	1	
Vocational relevance	good							bad
Interest	good							bad
Clarity of exposition	good							bad
Organisation of information	good							bad
Visual presentation	good							bad
Delivery - audibility	good							bad
- speed	good							bad
Rapport with class	good							bad
Amount learned	A lot							nothing
Problem solving	Improved							unchanged
Subject	Easy							difficult

Do you have any suggestions for future courses?

Pseudonym ....

Class ....

Thank you for your co-operation.

T. J. Lowe
DEPARTMENT OF EDUCATION

Course on Human Thinking

T-F 2

All the questions on this and the following pages can be answered from the information presented in the second half of the course. Some of the questions require memory, and others require some reasoning. The questions are in the form of statements, with which you are asked to agree, disagree, or signify ignorance. If you do not know the answer, please do not guess.

			Agree	Don't Know	Disagree
1.	3	When attempting to design a model of practical use to problem solvers, the flow diagram is of more use than the "stage model".	✓		
2.	4	Analysis of information consists mainly of separating into parts, and judging relevance and veridicality.	✓		
3.	5	The "imposing" of a structure or organisation on information is a considerable aid to memory.	✓		
4.	6	A teacher of engineering is more likely to resort to 'conceptual good figures' than a teacher of economics.			✓
5.	3	The best way to confirm a hypothesis is to go out and look for corroborative evidence.			✓
6.	4	The person who is able to see a problem in a number of different ways is more likely to produce a novel solution.	✓		
7.	5	In order to maximise the efficiency of information search, it is essential to specify exactly how the individual problem solver should go about reading articles, taking notes, etc.			✓

		Agree	Don't Know	Disagree
8. 7	The techniques of stimulating creativity all work by stimulating association in a random way.			✓
9. 5	The general stages which will be followed in problem solving vary with the type of problem.			✓
10. 4	The formulating of problems in real-life situations presumes a minimum level of critical ability.	✓		
11. 3	If the evidence hangs together in a consistent manner, you can be sure that it is valid.			✓
12. 6	A problem, in which the solution involved using a filing cabinet as a step ladder, would be solved more quickly by :			
	a) a clerk			✓
	b) a bricklayer	✓		
13. 8	The ideal <u>creative</u> problem solver would have the following characteristics :			
	a) flexible in thinking	✓		
	b) unhibited	✓		
	c) down-to-earth			✓
	d) well-motivated	✓		
	e) able to organise information			✓
	f) methodical			✓
	g) detached	✓		
	h) emotional			✓
14. 8	The ideal <u>over-all</u> problem solver would have the following characteristics :			
	a) emotional			✓
	b) detached	✓		
	c) methodical	✓		
	d) unconventional in dress			✓
	e) impulsive			✓
	f) analytic	✓		

			Agree	Don't Know	Disagree
15.	3	The algorithm is most useful in those problems which are ambiguous and badly defined.			✓
16.	4	Newspaper articles are a good place to look for 'conceptual good figures' because they are written in such haste.	✓		
17.	5	A state of healthy confusion may be conducive to creative problem solving.	✓		
18.	7	de Bono's account of the four ways of being right suggests that thinking can never be sure of being right without empirical support.	✓		
19.	3	The particular solution which the problem solver arrived at will depend on his own characteristic understanding of the problem.	✓		
20.	3	On the whole, information which conforms to 'good figure' will tend to be judged as valid.	✓		
21.	4	People who are trained to solve problems creatively feel less inhibited in their thinking as a result.			✓
22.	4	The importance of working to a schedule lies in the fact that it makes one concentrate on higher things, and consequently prevents one continually thinking about the trivialities of life.			✓
23.	5	The research on 'pecking order' showed that the dominance relations between birds did not fit into a 'good figure' largely because they do not have language.			✓
24.	6	It would be possible to write an algorithm on how to play chess.	✓		

			Agree	Don't Know	Disagree
25.	3	The most indepensable pre-cpndition for creative problem-solving is the ability to follow a single idea through to the end.			✓
26.	3	A major hindrance to the effective evaluation of evidence is the constant use of 'emotion-laden' words.	✓		
27.	4	It is important to put information into some sort of organisation or structure even though it makes thinking slower.			✓
28.	4	One of the main advantages of the algorithm is that it is simple to write one that is general across different problems.			✓
29.	6	The ancient idea that the universe consists of air, earth, fire and water owes more to 'good figure' than to observation.	✓		
30.	5	The attempt to stimulate one's own creative problem solving depends on a close examination of the creative thinking processes as they occur.			✓
31.	5	The main problem, when initially collecting information, is one of selection.	✓		
32.	6	The following argument is not valid: "I see no reason for taking any notice of your arguments for closing the branch-line: after all, your taxi company would benefit."	✓		
33.	5	It is not worth producing an algorithm unless that problem is likely to occur frequently.	✓		

Course on Human Thinking

T-F-1

Instructions. Below are a series of statements relating to the course that you have just taken. Indicate with a tick in the appropriate column whether of not you agree with them, or you do not know. Use the 'do not know' column where appropriate, as incorrect answers will be penalised.

		Agree	Do not Know	Disagree
3	1. The evidence seems to be that normal perception and imagery are two entirely different processes			✓
4	2. The mind is organised in a hierarchical fashion because every act or thought pre-supposes a number of other acts or thoughts.			✓
5	3. The evidence on perception suggests that in some cases we cannot rely on the evidence of our own eyes.	✓		
2	4. One can measure the duration of the visual icon by indicating which one of four rows of letters a person must report immediately <u>before</u> he is shown the display.			✓
3	5. A key characteristic of the human mind is the desire to organise all its beliefs into a logically consistent system.	✓		
4	6. Subjects who had to search for one example from four target letters took no longer than other subjects who were searching for only one target letter.	✓		
5	7. The requirement that a creative problem solution should be novel means that people with worse memories in fact make better problem solvers.			✓
6	8. The speed of chase scenes on the cinema is exaggerated by the way the eye perceives them.	✓		
2	9. Creative work flourishes in an atmosphere of stimulating mutual criticism.			✓
Y	10. The evidence that perception, dreaming and imagery all involve the perceptual system relies on the analysis of eye movements.	✓		
3	11. A person is less likely to see an emotionally-charged and noxious element of a situation.	✓		

		Agree	Do Not Know	Disagree
3	12. In most problem situations, all the elements of thinking are conscious.			✓
4	13. The best way to rote memorise something is to read it over and then copy it down.			✓
5	14. The experienced driver is more skilled precisely because he can carry on a conversation whilst driving.	✓		
6	15. A good tennis player would initially have an advantage over a non tennis player if they were both novices at table tennis.			✓
7	16. A perceptual experiment will only indicate something about the nature of perception when the person can be induced to make a mistake.			✓
8	17. The following qualities have been maximised in the evolution of perception:			
	a. speed............................	✓		
	b. accuracy of detail................			✓
	c. general impressions...............	✓		
	d. objectivity.......................			✓
	e. emphasis on personally relevant....	✓		
	f. storage of detail for later use....			✓
	g. minute analysis...................			✓
	h. isolation of unified objects.......	✓		
3	18. The creative person is the person who is more able to take psychological chances.	✓		
4	19. The ability to maintain a constant perception of size despite variations in retinal image depends on the perception of cues signifying the distance of the perceived object.	✓		
4	20. An experiment demonstrated that children who had less money around the house 'saw' coins as smaller.			✓
5	21. The principles of memory and perception suggest that the contents of dreams a re at least partially random.	✓		
6	22. It would seem likely that people who are placed in an environment where there is no perceptual variation at all will very quickly start having 'dreams, even though still wide awake.	✓		
2	23. A man who has bought a new car will search for information to justify his prejudices.	✓		

		Agree	Do not Know	Disagree
3	24. When a situation is expected to contain a particular aspect, that aspect is less likely to be noticed,			✓
4	25. The time-scale over which the perception of a stimulus occurs tells us little about the nature of perceptual processes.			✓
5	26. A knowledge of the importance of logical consistency in thinking suggests that as the cognitive component of attitudes is taught at University, the affective and behavioural components are automatically inferred.			✓
2	27. A consideration of how people judge size suggests that many sightings of the Loch Ness Monster may have been sightings of ducks, logs, etc..	✓		
8	28. The idea of the TOTE as the basic unit of mind suggests the following basic characteristics:			
	a. hierarchical organisation..........	✓		
	b. feedback is important..............	✓		
	c. many mental operations can be pursued simultaneously................	✓		
	d. the mind is good at taking an average over different impressions and ideas..			✓
	e. thinking is movement and action which has been internalised..........			✓
	f. it is difficult to do two things at once (e.g. overtake and talk)......			✓
3	28. In general, those things that are more likely to be seen are less likely to be remembered.			✓
4	29. The ability to produce a creative solution to a problem depends partly on the ability to see any situation in a number of different ways.	✓		
5	30. The template-matching model of pattern recognition can very easily be adapted to a novel perceptual environment.			✓
6	31. A person trying to estimate the distance of a tree would be wildly misled if doing it from a moving train.			✓
3	32. It is necessary to postulate some sort of pre-processing of the visual stimulus which must take place before the different features are recognised.	✓		
4	33. Information with more logical structure is easier to remember.	✓		

		Agree	Do Not Know	Disagree
5	34. There is likely to be considerable loss of information, as the information in theiconic memory is transferred to the verbal memory.	✓		
5	35. The perception of the colour of an object can depend on the shape of the object.	✓		
6	36. In reading a book, it would be helpful to memory to attempt to predict the contents of each chapter before actually reading it.	✓		

	+3 +2 +1 / -1 -2 -3	interesting / dull	independant / dependant	considerate / inconsiderate	responsible / irresponsible	cheerful / ill-humoured	interested in others / self-absorbed	calm / excitable	decisive / indecisive	adjusted / maladjusted	Outgoing / shy (+3 +2 +1 / -1 -2 -3)
Yourself											
Person you dislike											
Mother											
Person you'd like to help											
Father											
Friend of same sex											
Friend of opposite sex (or spouse)											
Person with whom you feel most uncomfortable											
Immediate Superior											
Person difficult to understand											

SPEED OF COLOR DISCRIMINATION TEST

The following items consist of samples or patches of four different colors—red, blue, green, and orange. For example:

*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****

You are to print under each color the first letter of the color's name. Print R under each patch of *red,* B under each patch of *blue,* G under each patch of *green,* and O under each patch of *orange.* Here is how a set of items should look when completed.

*****	*****	*****	*****	*****	*****
B	G	O	B	G	O
*****	*****	*****	*****	*****	*****
R	O	G	R	B	R

THE ITEMS MUST BE COMPLETED IN ORDER beginning at the top of the page and working each row from left to right. Do not omit any items.

This test is highly speeded, so work as quickly as you can without making errors. There will be four separately timed parts. Wait for the signal before turning the page. Remember, work as fast and as accurately as possible.

DO NOT TURN THIS PAGE UNTIL YOU ARE ASKED TO DO SO.

Developed under NIMH Contract M-4186, S. Messick, Principal Investigator.

*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****

STOP HERE.
DO NOT TURN THIS PAGE UNTIL YOU ARE ASKED TO DO SO.

*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****

STOP HERE.
DO NOT TURN THIS PAGE UNTIL YOU ARE ASKED TO DO SO.

*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****

STOP HERE.
DO NOT TURN THIS PAGE UNTIL YOU ARE ASKED TO DO SO.

*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****
*****	*****	*****	*****	*****	*****

STOP HERE.
DO NOT TURN THIS PAGE UNTIL YOU ARE ASKED TO DO SO.

The following items consist of the *names* of four colors printed in different colored inks. For example, the name "orange" may be printed in either blue, red, green, or orange ink. Here are some sample items:

orange red blue orange green orange blue green

green green blue green orange green red blue

You are to print under each word the first letter of the *color in which the word is printed.* Print R under a word printed in Red ink, B under a word printed in Blue ink, G under a word printed in Green ink, and O under a word printed in Orange ink. Ignore the meaning of the words themselves and indicate only the color of the ink used. Here is how a set of items should look when completed.

red	blue	red	blue	orange	red	red	orange	blue
O	G	B	O	R	O	G	B	R

blue	green	red	red	blue	blue	red	blue	orange
G	O	G	O	R	O	B	R	B

THE ITEMS MUST BE COMPLETED IN ORDER beginning at the top of the page and working each row from left to right. Do not omit any items.

This test is highly speeded, so work as quickly as you can without making errors. There will be four separately timed parts. Wait for the signal before turning the page. Remember, work as fast and as accurately as possible.

DO NOT TURN THIS PAGE UNTIL YOU ARE ASKED TO DO SO.

orange red green blue orange red blue green orange green blue

orange green orange green red orange red red orange red orange

green orange blue red blue green blue orange green orange green

orange red blue red red orange red blue green orange red

blue green red orange red blue red blue green blue green

green blue red blue red orange red orange blue red orange

blue green red green blue green orange blue orange red green

red red blue red green orange green green green blue orange

green blue blue green red blue red orange orange blue green

green red blue red orange orange red red orange green red

blue blue blue orange green orange red orange green orange green

blue orange green orange blue green red red orange orange red

blue blue green red blue red orange green orange green green

blue red blue green red blue orange blue red orange blue

red orange red green blue orange green orange blue red red

red blue green red orange blue green orange red orange blue

STOP HERE.
DO NOT TURN THIS PAGE UNTIL YOU ARE ASKED TO DO SO.

orange red blue orange green orange blue green orange green blue

green green blue green orange green red blue green red green

red blue red blue orange red red orange blue red green

blue green red red blue blue red blue orange orange orange

red orange red green green blue green red orange blue red

orange blue green blue red green red orange blue red green

blue orange red green blue orange orange red orange blue orange

blue red green blue orange green orange blue green red blue

orange red green orange blue red red blue orange blue orange

red green blue green orange orange red orange green blue red

red blue blue orange blue red orange blue green red orange

red blue green blue green orange red orange red orange blue

red green red green blue red red blue blue red green

blue orange green green blue orange green red blue green red

blue red blue orange green orange red orange green orange green

blue green orange blue red blue orange red orange green blue

STOP HERE.
DO NOT TURN THIS PAGE UNTIL YOU ARE ASKED TO DO SO.

blue green orange blue green blue red blue green orange red

green green orange blue blue green orange red green orange red

blue red orange orange blue green orange green red orange red

orange red blue orange blue blue red green orange green red

green orange blue green blue orange green blue orange red red

blue blue green blue orange green red orange red green blue

blue green red blue green red red orange green red red

orange green red orange orange green blue orange blue red orange

blue orange blue red blue green red green red blue orange

green orange blue green red orange red blue orange red blue

orange blue red green blue red blue green green orange green

orange red blue red green blue blue red orange orange red

red green blue orange green orange blue green orange green orange

red orange green orange blue blue blue red green orange red

red orange orange red blue red green green blue orange orange

blue orange red green blue green orange orange red orange green

**STOP HERE.
DO NOT TURN THIS PAGE
UNTIL YOU ARE ASKED TO
DO SO.**

orange red blue red orange red orange red green blue green

orange blue orange blue orange red green green red red green

orange blue orange green blue green orange red green red blue

orange red orange blue green red green blue red green orange

green blue green orange red blue red green red green blue

orange blue orange green blue orange blue green blue red orange

green red red orange blue blue orange blue green red green

red blue red orange green red green red blue green blue

green blue red red blue green red orange blue green orange

red orange blue red blue orange green blue green orange green

orange red orange green orange blue red blue red green blue

red green orange blue green green orange blue green red blue

blue red red blue green red green red blue orange red

orange red orange green blue green blue red orange red green

blue orange red blue orange blue blue red red blue green

orange red orange orange green blue green red orange blue orange

STOP HERE.
WAIT FOR FURTHER
INSTRUCTIONS.

L. S. Questionnaire

Below are some questions relating to attitudes to life, vocation and University, and to other factors which may influence your approach to your studies. Please answer the questions as sincerely and honestly as possible. If you object to any particular question, please refuse to answer it rather than put down a false answer.

The completed questionnaires should be returned to

T. J. Lowe,
Research Student,
Dept. of Education.

I would like to state categorically that all personal information obtained in this investigation will remain strictly CONFIDENTIAL.

Code Score. Item.
......

Name: [redacted] .

1. Name, in order of preference, three occupations you would like to pursue after having finished your studies

1.1....................

1.2...................

1.3...................

2. If all goes as you hope at University, which of these do you think you will take up?...................

3. How certain are you that this is your vocation?

Uncertain Certain

:___:___:___:___:___:___:___:___:___:___:

4. Suppose, through some unforseen circumstance, you were to fail your degree, or get a lower grade than you had hoped, how difficult would it be to continue with that career choice?

Impossible No effect

:___:___:___:___:___:___:___:___:___:___:

5. How relevant is your present course of studies in the Dept. of Education to your vocation? To what extent will you use the knowledge and skills you are now learning in your preferred future occupation?

Irrelevant Highly relevant

:___:___:___:___:___:___:___:___:___:___:

......

Code Score. Item

......

6. How essential do you consider the following items to your future happiness ?

6.1.Marriage: :___:___:___:___:___:___:___:___:___:___:

6.2.Wealth: :___:___:___:___:___:___:___:___:___:___:

6.3.Fame: :___:___:___:___:___:___:___:___:___:___:

6.4.Popularity: :___:___:___:___:___:___:___:___:___:___:

6.5.Service to individuals: :___:___:___:___:___:___:___:___:___:___:

6.6.Service to mankind: :___:___:___:___:___:___:___:___:___:___:

Irrelevant Essential

7. Suppose, through unforseen circumstances, you were to fail your degree, or get a lower grade than hoped for, how do you think this would affect your future life-style?

7.1.Marriage: :___:___:___:___:___:___:___:___:___:___:

7.2.Wealth: :___:___:___:___:___:___:___:___:___:___:

7.3.Fame: :___:___:___:___:___:___:___:___:___:___:

7.4.Popularity: :___:___:___:___:___:___:___:___:___:___:

7.5.Service to individuals: :___:___:___:___:___:___:___:___:___:___:

7.6.Service to mankind: :___:___:___:___:___:___:___:___:___:___:

A little A lot

8. To what extent are you satisfied with the amount of work you are doing? Is your effort adequate to get you the level of degree you are aiming for?

Not at all Completely

:___:___:___:___:___:___:___:___:___:___:

9. Was your present course at Aston your first choice? Or would you have preferred to attend another University if you had able to ? (Please circle)

First/second/third/fourth/fifth/sixth

10. Do you at present have a problem of any kind (emotional, financial, social, health, organisational) which is causing you anxiety or worry ?

Yes / No

11. Have you had this, or another, problem at other times in the past two years?

Yes / No

......

Code Score. Item

......

12. How much anxiety does this present problem cause you ? *

Slight Serious Extreme

:___:___:___:___:___:___:___:___:___:___:

13. How much does it affect your college work? *

No effect Hinders work Prevents work

:___:___:___:___:___:___:___:___:___:___:

* Omit this question if you are not suffering from any worry at all.

Sociogram Section

In this section, I would like some information on social relationships, in order to build up a picture of the patterns of friendship within the class, and then to relate them to study habits, interests and academic performance criteria. I would be grateful if you would co-operate by answering the questions below. Realising that this is a sensitive area, I would like to remind you that all the personal information obtained in this research is completely confidential.

All these questions relate to the people in your class who study Education.

14. Name, in order, the three people you feel most close to or friendly to

14.1......................

14.2......................

14.3......................

15. Name, in order, the three people you feel least close to

15.1......................

15.2......................

15.3......................

16. Are you satisfied with your social contacts and friends at University? :-

16.1 Is there sufficient opportunity to make friends?

Yes / No

16.2. Are your fellow students friendly enough?

Yes / No

......

..3 Are they too friendly - even interfering?

Yes / No

Educational Environment Perceptions

Scheduled Interview

Interviewer's Manual

Introduction.

The aim of this interview procedure is to assess certain aspects of the subject's perceptions of the educational situation, with a view to relating them to his academic performance. The logic behind this is that a student's level of academic performance will be partly determined by the way he perceives and categorises his educational environment, and this categorisation process is implicit in his preferences for different activities, and his interests in different aspects of his work.

The interview proceeds through a four-phase process, which is repeated on different subject matter :

1. The List. The subject is asked to list all aspects of a category of his behaviour (for instance, all study activities). He writes the list on index cards, one item to each card.
2. The Sort. The subject is asked to sort the previous cards according to a criterion (for instance, how much he enjoys the different study activities). The interviewer notes down the scores on the score-sheet.
3. Abstraction of Qualities. The subject takes the most highly-rated items, and attempts to express essential ways in which they are similar: for instance, what quality it is, which applies to some or all of these study activities, which makes them more enjoyable than other study activities. This is repeated with the items that have the lowest ratings.
4. The Re-sort. Other items, which have been elicited in a different part of the interview are sorted against the qualities elicited above.

The interviewer's function in these four phases is as follows:

1. The List. The interviewer must explain precisely what is required - what it is that the subject must list. He must answer any questions that the subject might have, and correct the subject if he appears to be listing items which are not in the required category. If the subject produces no items, or only a few, the interviewer must prompt the subject with questions, or, in the last resort, supply the subject with the appropriate check-list.
2. The Sort. The interviewer's task in this phase is restricted to explaining what is required, and noting the scores on the score-sheet.
3. Abstraction of Qualities. This phase is particularly difficult for the subjects, and hence may engender slight resistance. The interviewer's task is to answer any doubts the subject might have, and to give encouragement, and prompting questions. It is important that the interviewer does not impose his own abstraction of qualities: he must, however, monitor the subject's judgements. Thus the question

is "In what essential ways are these items similar?", and the interviewer must use his own judgement to decide the meaning of 'essential' in that particular context.

4. Re-sort. Here, items are re-sorted against the qualities abstracted. Many items will be intrinsically unsortable against the qualities. There must therefore be a preliminary sort into categories of 'Relevant' and 'Irrelevant' (i.e. the quality has no relevance to that item). Here again, the interviewer's task is to maintain the subject's motivation, and prompt him where necessary. Finally, the interviewer must note down the results of the sort on the score-sheet.

Throughout the interview, the interviewer should bear the following points in mind:

1. Good rapport is essential. The interviewer should explain the aim of the investigation, and enlist the subject's co-operation. Furthermore, as the ultimate research criterion is academic performance, it is essential to emphasise that all results of the interviews are confidential, i.e. they are restricted to the interviewer (yourself), and the researcher (T. J. Lowe, Dept. of Education), and are witheld from members of staff, and anybody who might be concerned with assessment.

2. Stick to the schedule. The procedure involves a number of lists, a number of sorts, and cross-combinations of these. There are therefore various piles of cards, which have to be used at different times in the interview. The cards are supplied in envelopes and folders, and are identified by code-letters. To avoid confusion, replace all cards in their envelopes as soon as they are finished with, and follow the schedule as closely as possible.

The schedule itself is divided into two parallel parts - SAY and DO. Instructions to the subject, explanations, and prompt questions are written out in full in the SAY section, although it would be better if they were paraphrased rather than read out verbatim.

The Schedule

SAY | D

(1. INTRODUCTION)

This interview is taking place as part of a piece of research in the Dept. of Education. The research is investigating how a student's patterns of activities, and perceptions of his subject affect his level of performance. The idea behind this is that attitudes towards the educational situation, and perceptions of the educational situation, to some extent determine performance. If it proves possible to show that certain types of perception - certain interpretations of what University is about - are associated with students who do well, then we may be able to give help to students who are misinterpreting their subject and surroundings.

(2. CONFIDENTIALITY)

In this interview, and in another questionnaire which you will be given, there are some questions which you may feel wary of answering. For instance, one question asks if you are doing enough work. If, by chance, you are one of these people with a guilty conscience about how much work they are doing, this may be a sensitive area. I would like to give a categorical assurance that all the personal information obtained in this investigation is absolutely confidential. As soon as all the information is coded, the identity of individual respondents will be destroyed. There is no way in which any information you give can get back to members of staff who may be involved in assessing you.

If it happens that you do violently object to a particular question, please refuse to answer it. Please do not put down a false answer, as it will be taken seriously, and may jeopardise the results of the investigation.

(3.CO-OPERATION)

We hope you will feel able to co-operate in this study. What we are interested in is you as an individual, your patterns of interest and perception, and why you are succeeding or not fulfilling your potential.

(ACTIVITIES SECTION)

(1. LIST)

Firstly, I would like to explore your patterns of study activities. I would like a list of the study activities you spend a significant part of your time on. This includes any activity you perform as part of your studies, or which contributes to your studies. Please write down one activity on each card.

Take out cards from the envelope marked SA, and give them to the subject.

Prompts: Try to think your way through a day and write down each study activity. What did you do today, for instance?

SAY	DO
From the time you got up, what study activities have you been involved in? What about yesterday? The day before? What about in the evenings? What about at weekends? (If necessary give the subject the checklist, Checklist S).	
(2. SORT)	Take the cards from the subject, shuffle them, and hand them back.
Now I would like you to sort these cards into categories according to how much you enjoy the activities written on them. Imagine there is a ten-rung ladder across the table. Put the activity or activities which you enjoy most on your extreme right, and then fill up the rungs down to the tenth pile on the left, in which are the activities you least like.	
	When the subject has finished, take the pile on your extreme left,, and confirm that they are the activities he most enjoys. Take Scoresheet S, and find the cell where row 10, and column SA, meet. Enter in that cell the codenumbers on the backs of the cards in that pile. Replace that pile in its original position, and repeat with all the other piles, filling up the cells in Column SA.
	Now remove piles 4,5,6,7, and replace them in envelope SA. Take piles 8,9,10 (most enjoyable activities), shuffle them, and give them to the subject.
You have said that these activities are the most enjoyable of the study activities. Now I want to try to find out why. What is it about them that makes them enjoyable to you? What quality do all of them share? In what essential ways are they similar? Prompts: If you cant think of something common to all of them, is there something common to two or three of them?	
	Take out cards from envelope SQ, and place them near to the subject. When he has thought of a quality common to the activities, ask him to write it on a card. Use up to <u>six</u> of the twelve SQ cards. Remove piles 8,9,10, and replace them in envelope SA. Take piles 1,2,3, shuffle, and give them to the subject.
Now I would like you to do the same thing with these cards. These are the activities you dislike, or like least. What is it about them that makes them less enjoyable?	

SAY	DO

Use the remaining SQ cards. When the subject can abstract no more qualities , replace the SA cards and the SQ cards in their envelopes.

(The sequence up to now is repeated for leisure activities)

Remove the cards from the envelope marked . LA and give them to the subject.

Now could we do the same sort of thing for your leisure activities. First, I would like a list of the leisure activities which you spend a significant part of your time pursuing. This includes things like going to the pub, going to the cinema, watching TV, reading, pursuing various sports and hobbies. It does not include what might be called routine things, like cooking, travelling, shopping and washing; unless, that is, you have a special and non-routine interest in such activities. Also, in case there is any doubt, we are not inquiring into your sex life, or lack of it.

As before, please write down one activity on each of these cards.

Prompts: Think your way through a normal day. What leisure activities are you likely to pursue today?
What about yesterday?
The day before?
At weekends?
In the holidays?
What would you do on an unexpected day off?
(If necessary, show the subject Checklist L)

When the subject can list no more leisure activities, shuffle the cards and hand them back to him.

As before, could you no sort these activities according to how much enjoyment you get from them. The ones you most enjoy on the extreme right, the ones you least enjoy on your extreme left: ten categories in all.

When the subject has finished, list the codenumbers of the cards in each category on Scoresheet L, Column LA.
Now remove piles 4,5,6,7, and replace them in envelope LA. Take piles 8,9,10, shuffle them, and hand them back to the subject

What is it about these activities that makes them enjoyable? What quality or quakities do they have in common? In what essential ways are they similar?

Take out the cards from envelope LQ, and get the subject to write down his answers, one to each. Do not use more than six of the twelve cards.

SAY DO

Remove piles 8,9,10, and place them in envelope LA.
Take piles 1,2,3, shuffle, and give them to the subject.

Now, could we do the same with these cards? These are the activities you dislike, or like least. What is it about them that makes them less enjoyable?

Use the remaining LQ cards. When the subject can abstract no more qualities, remove the remaining LA cards and replace them in the envelope. Leave the LQ cards on the table.

(RE-SORT)

Take out the SA cards, shuffle them and hand them to the subject.

Here we have the list of study activities which you gave me. I want you to sort them into categories again, but this time in terms of the qualities of pleasurable leisure activity. The idea behind this is that we want to get some measure of the cohesiveness of your mental functioning: the extent to which your work activities and leisure activities are psychologically separated. The extent to which they fulfill the same, or different psychological functions.
It may be that some of the qualities of leisure activities simply do not apply to study activities. So before you start sorting them into ten piles, it would be best for you to go through the study activities, and divide them into one pile where the quality is applicable, and another pile where it isn't.

Take card LQ1.

Let's begin with this quality - it is "........" (read out).
First we have to check which activities can be judged on this quality - which activities can be said to possess more or less of this quality, and which activities the quality simply does not apply to. Could you do that now please? Look through the activities and sort them into two piles, according to whether or not the quality "......." applies.

Take the pile of cards judged as 'irrelevant' and note down their codenumbers on Scoresheet S, at the bottom of Column LQ1. Move the 'irrelevant' pile aside, SAY | DO

You judged that the quality does apply to these activities. Now I would like you to sort them into ten piles, as before, according to how much they possess "........". Activities with the most "........", put on the extreme right and then work down to the left.

When the subject has finished the sort, note down the results in the rest of column LQ1 (Scoresheet S). Make sure that you are putting the 'high' category in Row 10.

(Repeat this procedure i.e. from the arrow on the previous page, for the remaining LQ cards, making sure you take them in the order 2,3,4,etc., and write the results in the appropriate column of Scoresheet S.)

Remove the SA cards and the LQ cards, and replace them in their envelopes. Take out the LA cards, shuffle and hand to the subject. Take out the SQ cards.

(Repeat the above procedure, from the arrow on the previous page, using the Leisure Activity cards, and sorting them against the Leisure Qualities. Note scores in the appropriate columns on Scoresheet L)

Remove the LQ cards and replace them in their envelopes.

Finally, with these leisure activities, please sort them into ten categories according to how much you think they contribute towards your general development. By this I mean how much they improve your skills and personality, your development into what you would like to be.

Note down the results on Scoresheet L. Remove the LA cards and replace them in their envelope. Take out the SA cards, shuffle, and hand them to the subject.

Now sort these study activities into the ten piles according to how much they improve your ability to succeed in the examinations: how much you think they contribute to your actual examination grades.

Note down the results on Scoresheet S. Replace all cards in their envelope, and all envelopes in the Activities folder.

SAY | DO

(SUBJECT OF STUDY SECTION)
(INTRODUCTION)

Now I would like to explore your ideas about the subject matter of education. As with the last part of the interview, there are no right or wrong answers here. A mass of concepts and ideas, like the subject of Education, or any other subject for that matter, can be perceived in an almost infinite variety of ways. The research, of which this interview is a part, is trying to measure some of these ways, and relate them to how students perform in their examinations. But firstly, I have a couple of questions to ask.

(QUESTIONS)
(Enter the answers on Scoresheet SS)

1. You said your subjects of study in the Joint Honours are....................?

2. Which of these are you most interested in?

3. Which of these are you least interested in?

4. How do you divide up your study time between these subjects? Have a guess at the percentage of your time you spend on ....(2).
How about on ....(3).
Therefore you spend ......% on ....

Let me repeat those figures. (Repeat). That is correct is it?

(LIST)

Now I want you to think about the subject matter of Education. I would like you to imagine that you have just been given the job of planning out a course on Education. It will obviously have to be sub-divided into a number of parts - anything between three and six parts. I would like to know how you would divide up education into course-parts.
By the way, I am not asking you to be deliberately revolutionary. If you think that the course is divided up correctly at present, then don't try to invent new divisions. But if you think the present divisions do not reflect the underlying subject matter, then perhps you could tell me how it might be better organised. Write down each course-part on one of these cards.

Take out some spare (un-numbered) cards and give them to the subject. Use up to six of them.

When the subject has finished, lay out the cards in front of him.

SAY DO

Now take another look at these course-parts. I would like you to try to plan out each one in more detail. Take each one in turn, and try to say what topics would be included in it. Let's take this one first (pick up any one). Supposing that you had to give a course of lectures on ...., how would you divide it up into different topics, each one of which might take up a lecture? Write down each topic on one of these cards.

(SORT)

Take out T cards. When the subject has finished, take the T cards, shuffle them, and hand them back.

I would like you to sort these cards, as you did with the other cards earlier in the interview, as before, into ten piles according to how interesting you personally find each of these topics. Put the ones you find most interesting on the right.

Note down the results of the sort on Scoresheet SS, Column TP. Shuffle the cards and hand them back to the subject.

Now could you sort them again? This time in terms of how important a part of the syllabus they are. I am referring now to the syllabus of the Dept. of Education Course, as reflected in the examinations. Put the most important topics on your right.

Note down the results in Column TS. Re-shuffle the cards and hand back.

Finally, can you sort them again according to how much study time you spend on each. Put the topics which you have spent most time on on the right, and the topics you have spent least time on on the left.

Note down the results on the Scoresheet, Column TT. Replace the cards in the envelope and folder. Hand out any supplementary questionnaire, where applicable. Thank the subject for his co-operation.

Study Activities Checklist

- Lectures
- Lecture-discussions
- Tutorials (1 or 2 students)
- Seminars (more than 2)
- Project work
- Interviewing
- Practical work
- Laboratory work
- Demonstrations
- Films
- 'Visits'
- T.V.
- Writing essays
- Writing theses
- Preparing talks/lecturettes
- Reading journals (interest)
- Reading books (interest)
- Reading journals (specific information)
- Reading books (specific information)
- Classifying notes.
- Reading notes
- Rote learning
- Revision
- Informal discussion
- Seeking information from 'experts'.
- Writing essay plans
- Copying notes from others

Leisure Activities Checklist

- Sports
- Hobbies
- Reading novels
- Reading magazines
- Reading newspapers
- Reading non-fiction (not your subject of study)
- Cinema
- Concerts
- Plays
- T.V.
- Art galleries
- Museums
- Discussions
- Debates
- Playing cards
- Other games (chess, draughts, Scrabble, etc..)
- Travel (other than necessary)
- Restaurants
- Pubs
- Dances
- Nightclubs
- Union affairs
- Specialist clubs
- Evening classes (hobbies)

Scoresheet S

Name:______________________ Age:________ Sex: Male/Female

Code: ...... Subjects of Study:______________________

____

____

		SP	LQ1	LQ2	LQ3	LQ4	LQ5	LQ6	Exam. Benefit
Most	10								
	9								
	8								
	7								
	6								
	5								
	4								
	3								
	2								
Least	1								
Irrelevant									

Scoresheet L

Name: ______________________________

Code: ......

	LP	SQ1	SQ2	SQ3	SQ4	SQ5	SQ6	General Benefit
Most 10								
9								
8								
7								
6								
5								
4								
3								
2								
Least 1								
Irrelevant								

Scoresheet SS

Name:..............................

Code: ......

1. Subjects of Study: ....
   ....
   ....

2. Most interested in ....
3. Least interested in ....

4. Percentage of time spent on:

   2: ....%

   3: ....%

   Other: ....%

		TP	TS	TT
Most	10			
	9			
	8			
	7			
	6			
	5			
	4			
	3			
	2			
Least	1			