Fluency and Flexibility of Thinking

A study of open-ended thinking
with pupils of high ability in Science.

A dissertation presented in partial
fulfilment of the requirements for the
degree of Doctor of Philosophy
in Education at Massey University

by

Donald M. McAlpine

ABSTRACT

This study explored some aspects of fluency and flexibility of thinking with a sample of 13 to 15 year old pupils of high ability in science drawn from England, U.S.A. and New Zealand. Stimulus material within the general field of science was presented for response in the open-ended idiom. This procedure established a closer alignment between the nature of the task and the interests and abilities of the subjects than hitherto obtained for such pupils.

Tests included three measures of open-ended thinking employing science stimuli, an intelligence test, an attitude scale, a personal preference questionnaire, and teacher and peer rating scales.

Scoring procedures and interscorer reliabilities for the open-ended measures were determined, and scores from all tests converted to normalized T-scores. Then a series of correlational studies was undertaken which examined relationships between measures of fluency, flexibility, intelligence and personality. Two centroid factor analyses - one in the cognitive, the other in the personality domain - were carried out in an effort to seek further evidence of the internal structure of matrices.

Results from total sample data (by country) were employed for the majority of statistical analyses, although selected samples incorporating high-low scorers ($\pm$1SD) and other selected groups based on sex, age, and classroom differences were utilised on other occasions.

Within the limits of sample selection some inter- and intra-cultural differences in cognitive and personality

behaviours were observed, although the major result was one of variation on a theme rather than of striking contrasts. Scores on either fluency or flexibility of thinking were seen to be less correlated across tests (e.g. generally between .30 to .45 for fluency, and between .25 to .35 for flexibility) than with each other within tests (e.g. generally between .65 to .75). This tendency was supported by factor analysis which suggested a composite fluency-flexibility factor anchored to each test. Although some measure of stable cognitive style was seen to be operating, it was overshadowed in the present study by a tendency for pupils to be task-specific even within the science domain.

Correlations between intelligence and fluency and flexibility of thinking were generally low and positive.

Differences on some attitude factors suggested that within this sample of pupils of superior ability in science high scorers in flexibility (and to some extent in fluency) of thinking tended to approve more of such traits as running risks, doing dangerous experiments, teasing people and investigating the unusual, and approve less of being obedient, accepting expert advice, and getting everything correct.

While the study was not primarily concerned with the wider issues of divergent or creative thinking, nor with the academic bias of science/arts students, they were briefly discussed when the educational implications of the study were considered.

ACKNOWLEDGEMENT

Appreciation is due to Professor Liam Hudson, University of Edinburgh, who offered guidance during the formative stages of the study at the Research Unit on Intellectual Development, King's College, Cambridge, England.

Thanks are also due to the following professors for the opportunities of further study and discussion:

Jacob Getzels, Professor Education and Psychology, University of Chicago;
Donald Johnson, Professor of Psychology, Michigan State University, East Lansing;
Nathan Kogan, Senior Research Psychologist, Educational Testing Service, Princeton;
Donald MacKinnon, Director, I.P.A.R., University of California, Berkeley;
Jack Merwin, Professor of Educational Psychology, University of Minnesota, Minneapolis;
Sidney Parnes, Director, Creative Education Foundation, State University of New York, Buffalo;
Donald Taylor, Professor Psychology, Yale University, New Haven.

Appreciation for the more regular counsel of dissertation supervision is due to

Dr David Barney, Senior Lecturer in Education, University of Auckland;
Dr Harrison Bray, Senior Lecturer in Education, Massey University;

and particularly to Clement Hill, Professor and Head of Education Department, Massey University.

Finally, thanks are due to the staff and pupils of schools involved in the study, to the Nuffield Foundation (Science Teaching Project), London, and to the following for financial assistance:

The British Council, London,
The Carnegie Corporation of New York,
The Humanities and Social Sciences Research Fund, Massey University.

Donald M. McAlpine
December, 1971.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

This investigation represents some initial explorations into the dimensions of fluency and flexibility of thinking with pupils of high ability in science. Central to the study is the presentation of stimuli within the general science domain. The idea of a closer alignment between the nature of the task and the interests of the subjects as a basis for studying fluency and flexibility in greater depth arose from a suggestion by Hudson (1966).

Within the framework of test stimuli and concept of divergent and convergent thinking employed by Hudson, arts specialists were seen to be mostly divergers, and physical scientists, convergers. However, after closer examination of some of the rather ordinary responses of boys known to be clever at science, Hudson comments that one can only conclude that the test failed to catch their fancy (p.47). And, again, of boys clever at science, he observes that we can only hope to measure their capabilities by setting problems within their sphere of interest (p.49). Supportive evidence from word association studies also suggests that response availability and heterogeneity - fluency and flexibility - increase as the subject favours the stimulus (Cramer, 1968).

The present study takes cognizance of these points and has attempted to employ stimuli in the open-ended idiom that have a degree of compatibility and interest for pupils of science. The measurement emphasis is one of encouraging fluency and flexibility of thinking, of generating as many different responses to a particular problem as possible rather than one of arriving at a single 'right' answer. An open-ended mode of presenting science stimuli was consequently designed for four task situations. A sample comprising pupils of high ability in science was deliberately selective, giving a degree of parallelism between task on the one hand and population interest and competence on the other.+

+An incidental consequence of such an approach was the enjoyment that science pupils displayed in performing the tests.

Furthermore, several of the schools in the sample were following programmes from the Nuffield Science Foundation (England and New Zealand) and from the National Science Foundation (U.S.A.). Such programmes emphasize enquiry, experimentation and flexibility of thinking. Therefore, the curriculum background of many of these pupils together with the fact that test stimuli are pitched in the area of their abilities and interests offers every encouragement for the sample to meet the task requirements of fluency and flexibility. However, a wide range of scores can still be expected within this framework.[+]

The initial section of the study, which examines problems of categorization associated with scoring responses to open-ended tests, reflects an S-R emphasis within the general framework of word association theory and linguistics. The remainder mainly reflects a more general cognitive view which considers facets of intelligence, cognitive style

---

[+] The degree to which these procedures might alter the comparative divergent/convergent bias of arts and science students is not the concern of the present study, although the nature of the stimulus material must surely be a condition that influences such results.

and personality.

The study is not premised on any strict sequence of null hypotheses, nor on any predetermined hierarchy of statistical procedures.+ Statistical methodology has nevertheless formed an integral part of the investigation, having been centred on an initial general sequential plan open to modification as the study progressed.

Behind this general area of research, lies a silent substructure of parental influences including child rearing practices, styles of discipline in the home, balance of praise over punishment and success over failure, humour, and encouragement to pursue special interests; of teacher influences including use of open-end or dead-end questioning, encouragement to explore, the balance of problem finding to problem solving; and finally of social and cultural norms which exist to be followed, questioned or broken. These factors undoubtedly interact to produce dramatic differences in the range and style of thinking of which fluency and flexibility are two important dimensions. Ultimately they contribute

+Meehl (1967) has commented that research workers who employ this approach resemble "a potent-but-sterile intellectual rake, who leaves in his merry path a long train of ravished maidens but no viable scientific offspring" (p.114).

to the cyclic tension between 'normal' and 'revolutionary' science (Kuhn, 1962) in which the cognitive style of the scientist plays a significant role.

The present study, however, is concerned with a relatively narrow shaft of this substructure. It is certainly not a study of creativity, although some aspects of research under this general heading impinge on the study. It is not a dichotomous study that contrasts divergent/convergent thinking, or arts/science bias, although again some aspects of such research are relevant. Nor is it primarily a cross-cultural study, although the sample is drawn from England, U.S.A. and New Zealand. In no way is it claimed to be 'representative' of each country in any cross-sectional sense. On the contrary it is selective in terms of high ability in science, thereafter being matched as closely as possible in each country. As a result, greater intra-cultural differences may be expected to exist than inter-cultural differences. The general emphasis of the study is therefore cognitive. Finally, any conclusions that might be drawn from the investigation are not generalizable to wider samples - particularly to pupils of less ability or different interests.

AIMS, METHOD

The major purpose of the study was an investigation into the characteristics of fluency and flexibility of thinking in the responses of pupils of high ability in science. Stimuli generally related to science were presented in the open-ended idiom.

Some measure of the average level of generation of responses in terms of fluency and flexibility, as well as the range of individual differences, was of interest to the study. No comparisons with arts students were made.

More specifically the study aimed at considering the relationship between:

1. Fluency and flexibility of thinking;
2. Fluency, flexibility and measured intelligence;
3. Fluency, flexibility and aspects of personality;
4. Fluency, flexibility and teacher and peer ratings.

The major emphasis of the study was generic, with general tendencies and relationships between variables being studied by way of total sample data.

However, some sections of the study examined comparisons between high and low scorers (±1SD) on certain variables, since interesting relationships are sometimes evident amongst special samples who score towards both ends of a continuum. Other special sample analyses were concerned with age and sex differences.

Statistical procedures included the normalizing of all scores (T scores), correlational studies, and a centroid factor analysis of (i) cognitive variables, and (ii) attitudes; non-parametric techniques such as chi-square and Fisher's exact probability tests were employed in some of the high-low comparisons. These procedures are detailed later (p.77).

The following tests were designed for the present study. Three were open-ended tests employing science stimuli : Test 1, Science vocabulary; Test 2, Definition of science concepts; Test 3, Uses for chemical substances and physical objects; Test 4, was an attitude scale measuring such attitudes as risk-taking, respect for authority, and self-concepts.

Teacher and peer rating scales were also administered. Intelligence was assessed with the AH5 Test (Part I). All tests were administered by the investigator to each school sample on the one day. The development of the tests is detailed under the section 'Stimulus Material - Instrumentation' (p.22).

Since no scoring procedures had been established for the open-ended tests the setting up of criteria for scoring became a central concern of the initial section of the study. Satisfactory levels of objectivity were eventually indicated from the interscorer reliability study. Scoring procedures are discussed in more detail under 'Scoring Responses' (p.49). A fuller description of the sample now follows.

SAMPLE

General: A sample of 506 pupils of above average ability in science over the two year range 13 years 4 months to 15 years 3 months was drawn from 16 schools and 23 classrooms - nine from England, seven from U.S.A., and seven from New Zealand.

The distribution of the major sample according to geographic location, size and sex can be seen from Table 1.+

TABLE 1.

Distribution of Sample by Geographic Location, Size and Sex

Country	Location	School No.	Class No.	Sample Size	Sex M	Sex F
	Cambridge	01	01,02,	33	33	-
	London	02	03	26	26	-
ENGLAND	London	03	04,06,07	57	57	-
	London	04	09	17	17	-
	London	05	10,11	49	49	-
		Σ= 5	Σ= 9	Σ=182	182	-
	Buffalo,N.Y.	06	12	22	9	13
	Chicago	07	13,14,15,	32	32	-
USA	Minneapolis	08	16	28	20	8
	Minneapolis	09	17,18	58	27	31
		Σ= 4	Σ= 7	Σ=140	88	52
	Auckland	10	19	31	14	17
	Auckland	11	20	27	27	-
	Auckland	12	21	20	9	11
NZ	Wellington	13	22	26	26	-
	Wellington	14	23	21	21	-
	Wellington	15	24	24	12	12
	Wellington	16	25	35	-	35
		Σ= 7	Σ= 7	Σ=184	109	75
		Σ=16		Σ=506		

+The English sample was tested first, followed by the U.S.A. sample and the New Zealand sample. The opportunity to test in coeducational schools in U.S.A. and New Zealand offered the further possibility of considering sex differences within a cross-cultural framework.

Science Ability

All pupils in this study have been designated by their school as well above average ability in science, coming as they have from either the top stream in science or from the top general academic stream where no 'setting' is employed. In England and New Zealand pupils from several schools were involved in the Nuffield Foundation Science Teaching Project, while in the United States the four schools chosen were participating in a progressive and challenging science programme.

In the initial stages of determining the sample several science achievement tests standardized in England and U.S.A. were considered. However, test content proved too specific and sensitive to curriculum emphases to be suitable as a valid cross-cultural criterion of science attainment. Obviously each pupil has been subjected to many formal and informal evaluations based on a wide range of criteria by the time he is 13 or 14 years. These multiple criteria, sensitive to each pupil in his own school and culture, appear more appropriate as a basis of selection for this sample than a series of externally standardized tests that are far less 'fair' cross-culturally. Thus, the selective procedures of the educative process, saturated as they are with all their complexities were allowed to operate in each culture to determine at this point in time those pupils deemed to have above average to exceptional ability in science.

Intellectual Ability

The general intellectual level of these pupils may be inferred to be above average in terms of measured intelligence. Nevertheless some pupils who are in the top stream for science may be of only average intelligence. In order to assess the intellectual level of these pupils as measured by the conventional intelligence test, the AH5 test was administered to the sample.+ The results of this testing confirmed assumptions of generally high intellectual ability across the sample and will be presented in more detail later (pp. 79-100).

Age

The distribution of the sample over the two year age range appears in Table 2 arranged for the three countries by three eight-month intervals:-

---

+AH5 Group Test of High-Grade Intelligence
A.W. Heim. N.F.E.R.

TABLE 2

Distribution of Sample by Age and Country

Age Intervals Yrs/Mths		Yrs/Mths	ENG M.	ENG F.	USA M.	USA F.	NZ M.	NZ F.	Total
13.4	-	13.11	49	-	13	16	24	2	104
14.0	-	14.7	107	-	47	19	40	30	243
14.8	-	15.3	26	-	28	17	45	43	159
		Σ =	182	-	88	52	109	75	506

Socio-economic and Cross-cultural Characteristics

Three-quarters of the 16 schools were situated in attractive suburbs while the four central city schools drew pupils from high socio-economic backgrounds. These characteristics were common to the samples from each culture. However, the English sample was drawn from schools which reflect greater social and educational stratification and selection. By contrast the New Zealand and U.S.A. samples were drawn from schools which portrayed a more egalitarian approach to educational organization.

The sample makes no attempt to reflect any purported cultural traits, nor to be in any way 'representative' of each culture. Rather, it is selective, being matched in each country by characteristics of well above average ability in science, 13 to 15 year age range, well above average general intellectual ability, and generally high socio-economic status.+

Summary and theoretical position

Since the present investigation is exploratory in nature no single theoretical position structures the study as a whole, but rather cognizance is taken of several different positions which monitor the interpretation of data and suggest directions for further study. These positions include (i) stimulus response theories - particularly associationist theory as interpreted by Mednick (1962), and the concept of aligning the stimulus material with the interests and abilities of the subjects to facilitate higher degrees of fluency and heterogeneity of associates (Cramer, 1968), and (ii) cognitive theories - particularly those that relate fluency and flexibility of thinking with aspects of personality associated with risk-taking, self concept, conformity and respect for authority.

In view of the above the more orthodox position which attempts to gather together at the outset all the relevant research as the basis of a single comprehensive

---

+It is interesting to note the similarity in the U.S.A. sample - in terms of size, age, and socio-economic status - to that of Wallach and Kogan (1965).

statement of theoretical position (Travers,[+] 1964, pp.31-50) was set aside in favour of a less pre-determined position which still takes account of relevant research but which mediates theory and data and is progressively self-regulatory in nature. Such an approach is employed by several workers in the present field of research (Cropley, 1968; Hudson, 1966, 1968; Mackworth, 1965). More flexible theoretical positions such as the above are well authenticated in literature on empirical research (e.g. Sax[++], 1968, pp. 12, 14-16, 27, 71-72).

Further description of the theoretical background which undergirds the study, and which formed the basis of the first pilot study employing scientific concepts as a means of assessing aspects of cognitive style (p.27), will be described in the following chapter.

[+] Travers, Robert M.W. An Introduction to Educational Research, Macmillan Co. New York 1964.

[++] Sax, Gilbert. Empirical Foundations of Educational Research. Prentice-Hall. 1968.

CHAPTER TWO

THEORETICAL BACKGROUND AND INSTRUMENTATION

The general field of psycholinguistics (Deese, 1970) and semantic theory (Katz, 1967; McNeill, 1966; Mednick, 1962) undergirds the initial section of the study on fluency and flexibility of thinking, since 'results' must first be derived from techniques of word analysis.

Relevant to such a study is the point of view expressed by Vygotsky (1962) that the "development of verbal thought is made possible by the use of word meaning as the analytic unit" (p.121) and that "the word is a microcosm of human consciousness" (p.153).

Attempts to explain the origin of consciousness and its expression through modes of divergent thinking vary from psychoanalytic interpretations (Trosman, 1969) to specific neurological ones - e.g. cortical control and inhibition in creative thinking processes (Morozov, 1968), and the role of the corpus callosum as a mediating influence on ideational fluency (Bogen and Bogen, 1969).

As far as the interpretation of divergent thinking responses are concerned, two broadly contrasting theories have been described by Cropley (1970): (i) S-R theories, and (ii) Cognitive process theories.

(i) The S-R approach has been developed by Mednick through the analysis of word associations, whereby divergent people tend to link stimuli with highly unlikely responses, while most people associate them with those commonly paired in the past. He further points out that S-R theorists have placed considerable emphasis on the so-called 'mediating processes', proposing different theories for 'explaining' the intervening variables between the S-R bond (e.g. Osgood, 1953, 1957).

Berlyne (1965) also modified the S-R approach to include an examination of intervening variables in the chain of symbolic thinking processes, and to emphasise the place of the organism in giving direction and structure to thinking. His theory of habit strength hierarchies has been further elaborated (McAlpine, 1970) as a restructuring process in creative thinking. Berlyne's integrative neoassociationist approach can be seen to some extent as bridging the gap between S-R and cognitive theorists. This general position, reported by Shouksmith (1970), is held by several contemporary theorists who would maintain that "the present best guess still favours a connectionist theory albeit modified by symbolic mediating responses" (pp. 33-34).

(ii) The cognitive process theorists, the other broad group described by Cropley (1970), emphasize the priority of the individual himself in organizing information - coding, categorising etc. Cognitive theorists, he continues, are more concerned with broader issues such as category width and cognitive style, and with related personality characteristics such as risk-taking and authoritarianism.

The present investigation moves from an initial analytic examination of stimulus-response situation which involves defining categories, establishing taxonomies and determining scoring procedures, to a more general cognitive point of view which examines the relationship between certain cognitive and personality variables. It is appropriate here to describe briefly some of the more relevant research related to S-R connectionist and associationist theory as it pertains to this initial phase of the investigation.

In general there has been a considerable revival of interest during the last decade in word association. Fox (1970) claims that "more psychologists have asked more subjects to free associate to verbal stimuli in the past decade than over the previous half century combined" (p.31). Similarly Cramer (1968) holds that "this renewal of interest in word association in large part reflects the general increase in investigations of verbal learning and verbal behaviour and, perhaps, an increasing interest in the study of cognitive processes". To some extent it also reflects an increasing emphasis on the open-ended mode of asking questions. Cramer (1968) continues by pointing out that "for some investigators, the identification of natural language associative habits is but a first step in the study of other cognitive processes" (p.1), and that in this type of study the interest is in providing calibrated experimental materials for assisting the understanding of broader cognitive processes. It is this emphasis, rather than a detailed study of word association per se, that has been taken in the present investigation.

Employing a similar emphasis, Mednick (1962) explained how associative hierarchies are built up from the responses of a stimulus word. Hierarchies which have a steep slope reflect conventional responses, those with a flat slope unconventional or remote associates. Divergent thinkers tend to generate more associates (fluency), more unique associates (originality), and more heterogeneous associates (flexibility). Thus the word associations of the divergent thinker are characterized by less stereotypy and commonality. As a result of this theory Mednick devised the Remote Associates Test (RAT) which has formed the basis of a considerable amount of research in creative thinking and verbal learning behaviour: the relationship between creativity and the need for associative novelty (Harris and Hall, 1970); the role of creativity and intelligence in conceptualising (Jacobson et al., 1968); the relationship between intelligence and verbal ability and ratings of research graduate students on creativity (Ginsburg and Whittemore, 1968); and differences in associative responses of honours and non-honours students (Abney, 1970).

Further studies which have specific relevance for this investigation include: a discussion of principles involved in coding sentence structure (Steinberg, 1970); elaborative activities and selective processes that determine the ordering of responses (Di Vesta and Bernstein, 1969); associative strength as the result of category membership (Henry and Voss, 1970); associative groupings and free recall (Matthews and Hoggart, 1970); contrast and oppositional tendencies in associations (Pollio, Deitchman and Richards, 1969); delayed instructions (latency) and resultant influence on level of originality on word associations (Masters et al., 1970; Masters and Anderson, 1970); the role of memory in the production of divergent thinking responses (Pollert, Feldhusen et al., 1969); and, the function of memory and acoustically similar responses on recall (Bruce and Crowley, 1970).

Age has also proved to have been a critical variable in many studies. In this respect Cramer (1968) reports that "cultural-chronological studies of word association indicate that the responses of both children and adults have become more stereotyped over the years. This increased response consistency is most marked in the associative behavior of children" (p.224). She continues, "It is not just that today's children are more consistent in the responses they give; they also give a different set of responses than did the children of earlier times" (p.224). This changing nature of responses to science words like space, gravity, solar system, has also been noted by Mason (1969) and Schmidt (1968). Cramer continues, "In addition to this increased stereotypy, there is also evidence that there are more idiosyncratic responses among today's children (and possibly among adults, also). This finding suggests a polarization of individuals, in terms of associative behavior, such that the flexible middle of earlier times has been replaced by the extremes of response stereotypy (high-commonality subjects) and response idiosyncrasy (low-commonality subjects)" (p.224).

Some studies (Entwisle, 1966) have employed factor analysis to assist scoring procedures and to arrive at meaningful clusters of associates.

In this respect Neman and Dixon (1970) comment that factor analysis "will probably provide a more refined structural view of the stimulus and response components of word association" (p.507). Similarly Rotberg (1965) extended Deese's procedure for classifying groups of words similar in associative meaning by using factor analysis. Comprehensive categories of meaning were determined based on the assumption that the common free associates to a set of related stimuli are mediating responses and are part of the associative structure of the set.

As with factor analysis, Stone et al. (1966) outlined a computer approach to scoring in an effort to bring more precise and comprehensive techniques to bear on the problem of word analysis.

The present investigation, however, does not employ some of the sophisticated techniques of linguistic analysis described above, nor factor analysis as a scoring procedure, although it is used later in the study as a technique for describing the internal structure of matrices relating to intelligence, fluency and flexibility. Nevertheless, a satisfactory level of objectivity appears to have been achieved in scoring which perhaps reflects a greater degree of precision than obtained by Torrance (1966) who maintained, "it is not necessary to have special training in scoring these tests", and suggested, "If the examiner does not yet have an understanding of the concepts of fluency, flexibility and originality, he should do some supplementary reading of the rationale of the Torrance tests of creative thinking" (p.11).

STIMULUS MATERIAL:- INSTRUMENTATION

Subsequent discussion is devoted to a consideration of the following:

A. Three open-ended tests of thinking employing science stimuli, viz:

Test 1, Science Vocabulary

Test 2, Definition of Scientific Concepts

Test 3, Uses for Chemical Substances and Physical Objects

B. Test 4, Attitude Scale and Personal Preference Questionnaire

C. Test 5, AH5 Test of Intelligence (Part I)

D. Teacher and peer rating scale information

E. A description of the pilot study with the above tests concludes this section on instrumentation.

A. Three open-ended tests of thinking employing science stimuli

General Instructions

Instructions for Tests 1, 2 and 3 aimed at:

1. Introducing the pupil to the general tone and demands of the open-ended situation,
2. Emphasizing fluent ('many') and flexible ('different') thinking responses,
3. Allaying any concern about spelling, or inclusion/exclusion of category membership - if in doubt, pupils were encouraged to record their response. This was particularly stressed with Test 1 where 'words to do with science' was the stimulus,
4. Offering some freedom of stimulus choice within the science stimulus field (Tests 2 and 3),
5. Encouraging pupils to work 'fairly quickly', although at the same time pointing out that it is 'not a race against the clock'.

Introductory comments to pupils appear below:

> "I want you to help me by answering some questions. Some of these are to do with science, some are questions about yourself. These questions are different from many you have at school, in that there are no right or wrong answers to them.
>
> Your answers are confidential and none of your teachers will be seeing your papers.
>
> It is important that you keep your work to yourself and work quietly. You should not have to ask any questions once you start writing. Finally, you should work fairly quickly, although it is not a race against the clock.
>
> You may use ball point (biro) pen, and you will need something to rule with."

A relaxed atmosphere was sought within the framework of the demands of the task, but there was no attempt to create the 'game-like' atmosphere of Wallach and Kogan's (1965) study.+ All instructions were read aloud while pupils paced the reading from their own script. Pupil responses were recorded on unlined paper (cf Torrance 1966).

Test 1: Science Vocabulary

This test, designed last in the series of open-ended tests, was formulated to assist rapport and introduce the subject to the general style of the two main tests. Although associative processes are basic to the three tests as a whole,they appear in more simple structure form in Test 1. Because of its single word response pattern it also presents some of the scoring problems in a rather more simplified form.

+More recent research by Kogan and Morgan, (1969) reports that "No clear-cut superiority for test-like or game-like contexts in reference to creativity level was observed" (p.125). However, further work by Nicholls (1971) and Vernon (1971) suggest that this is a fruitful area for further research.

Instructions
Test 1

Write down as many words to do with science as you can think of.

Think of as many different words as you can.

List your words in a column.

Work fairly quickly.

A sample of responses from two pupils for Test 1 appears on the following page.

Sample of Responses, *two Pupils, Test 1

Pupil 579	Pupil 365
Chemistry	biology
biology	astronomy
phisiology	species
physics	geology
zoology	ion
flux	atom
density	molecule
barn	palientology
neutron	genus
roentgen	archeology
curie	botany
centimetre	hypothesis
alkali	sociology
acid	kingdom
base	marsupial **
reactor	
neutron	
deuterium	
plasma	
ion	
electron	
- meson	
muon	
deutron	
nitrogen	
specific heat	
tensile	
steam	
water	
ice	
mercury	
atomic	
weight	
mass	
light	
reflection	
camera	
photon	
quanta	
relitvity	
theory	
hypothesis	
planet	
galaxy	
iron	
nickel	

*Pupil responses represent exact copies from original scripts

**Several scientists who have examined the responses to Test 1 have expressed surprise at the wide range of scientific vocabulary for 13 and 14 year olds.

Test 2: Definition of Scientific Concepts

Description

A preliminary study in New Zealand with 143 school boys and girls of high academic ability (age 11 to 12 years) had indicated that the definition of scientific words (concepts) held promise as a method for assessing not only 'knowledge', in the traditional sense, but also such aspects of intellectual style as fluency and flexibility of thinking.

It appeared from initial experiments, that stimulus words which were fairly pervasive and perennial - such as TIME, ENERGY and SPACE appeared to have the greater scope. Words like WAVE also seemed profitable since they could be seen in many different contexts - light wave, micro-wave, heat wave, goodbye wave, and not-in-use wave - as in 'wave that regulation' if the pupil showed a lack of spelling ability or a sense of cognitive sport.

Words such as TIME, ENERGY and SPACE also have the advantage of being able to be used more fairly across classrooms and cultures than more specific words or concepts which may be more sensitive to syllabus content. It is not claimed that words like TIME, ENERGY, SPACE and WAVE do not feature in syllabi, but rather that they are more generic than PHOTOSYNTHESIS, VECTOR, CATALYST or QUINOL ($C_6H_4(OH)_2$). They also have the advantage of being relevant over a wide age range.

From an initial experimental list of 86 words the following 23 were selected as having some promise.

SOUND	CHAIN-REACTION
BACTERIA	AQUATIC
HEAT	MAGNETISM
PLANET	WATER
VIRUS	GAS
GRAVITY	PLASTICS
VACUUM	PRIMATE
WAVE	LIGHT
SPACE	FORCE
PROTEIN	ENERGY
TIME	REPRODUCTION
COMMUNITY	

Finally the following four words were chosen as the stimuli for Test 2:- SPACE, TIME and ENERGY, together with a choice of one from WAVE, BACTERIA or WATER. Pupils' responses for SPACE and TIME have been analysed in the present study.

Instructions: Test 2

Although this test is concerned with the definition of scientific concepts, a test instruction which bluntly demands "Define ENERGY" appeared inappropriate to the aim of establishing a relaxed atmosphere and the generation of many different ideas. The purpose was to establish an atmosphere that combined the demands of knowledge of scientific terms with the encouragement to explore as widely as possible. The following instructions were consequently devised and read aloud to pupils who followed them from their own paper:

> <u>Instructions</u>
>
> <u>Test 2</u>
>
> "I am going to write some words on the blackboard and I want you to tell me all you know about them, I will write one word at a time and then you write as <u>many different ideas</u> as you can about it.
>
> Write down everything that comes into your head. You do not have to write complete sentences, just short phrases will do i.e. a sort of 'telegram style'.
>
> You can list the points one under the other if you like. Write the word down each time and rule off when you have finished writing all you know about it.
>
> Work fairly quickly."

Each stimulus word was printed on the blackboard and presentation time controlled with a stop watch.

It is interesting to compare Wallach and Kogan's (1965) test instruction at this point e.g. Test 1 (p.29)

> "In this game I am going to tell you something and it will be your job to name as many things as you can think of that are like what I tell you. For example, I might say 'things that hurt'. Now you name all the things you can think of that hurt". (The experimenter then lets the child try.) "Yes, those are fine. Some other kinds of things might be falling down, slapping, fire, bruises, or a knife".............

It will be noticed that Wallach and Kogan offer the pupils examples of responses. It was decided in the present study that to give specific examples of responses was to hand pupils a model - just the opposite of what is required for measuring flexibility.

Sample of Responses, two pupils, Test 2

SPACE (Pupil 143)
"Vacuum, around us, contains gases, bounded by solids or liquids, contains 5 letters, space between everything not joined, around everything, can only be seen because there is nothing else to see, space even in solids - molecules moving, lack of friction, rockets travel into space, spaceship, space race (America Russia etc.), lack of space,"

TIME (Pupil 143)
"Time machine, travelling in time, watches, clocks, hours, minutes, seconds, time wasting, time to kill, no time, time is measured, all the time in the world, contains 4 letters, magazine, time and tide, future past, fortune telling - able to see into future, time is essential - time flies, appears to go v slowly, time to get up."

SPACE (Pupil 115)
"Emptiness, void, moons, starts, rockets, vacuum, planets, no atmosphere, loneliness, no end, infintesimal, asteroid, travel through, astronomy, packing into,"

TIME (Pupil 115)
"arbitrary, basic setting of time for whole world, 4th dimension, essential to concepts of work, International date line, No true beginning or end, relentless, measured by clocks, set by astronomical features e.g. pos. of moon sun in relation to earth."

Test 3: Uses for Chemical Substances and Physical Objects

Description

In keeping with the general rationale of the study, science pupils were asked to respond to the 'uses' test with science stimuli. As in Test 2 stimulus material was kept as general as possible. Furthermore, pupils were offered a choice amongst four stimulus words, viz:

(1) OXYGEN or ACID

and (2) WHEEL or MAGNET

The element of choice has certain advantages in inducing a more positive attitude to the stimulus word and hence in the generation of ideas, but has disadvantages in setting up two sets of scoring norms and of producing different sets of N for statistical analyses.

Instructions

Test 3

"I want you to think of as many different uses as you can for each of the words I write on the blackboard.

Write the word down each time and then think of as many different uses for it as you can.

Write down everything you can think of.

Write your ideas down in a column.

Work fairly quickly."

Samples of Responses, two pupils, Test 3

OXYGEN (Pupil 145)

"Breathing
Forming oxides
Using in a blow lamp
In hydrogen peroxide as a bleach
For making water
For burning
For pumping under pressure
In a fluorescent light"

WHEEL (Pupil 145)

"Turning around
As a grindstone
For playing hoola-hoop
For playing quoits
To support me swimming
To hold up a car
To steer a car
To keep an engine going as a flywheel
To open and shut air locks
To lock a locomotive boiler
To steer a ship
To hold up candles as a chandeleer
To make a round about."

ACID (Pupil 347)

"Making hydrogen
neutralization
making salts
burning through metal
getting rid of an enemy
garbage disposal
suicide device
indigestion promoter
persuasion
creating new art forms
experimentation"

MAGNET (Pupil 347)

"Particle accelerator
a play thing
picked up spilled nails
experimentation
cheating at bowling
fixing roulette
picking up metal objects
making electricity
making a compass"

This concludes a description of tests of the stimulus-associative response kind. Test 4 anticipates relationships of the general cognitive-personality type.

B. Test 4 Attitude Scale and Personal Preference Questionnaire

Description

Items in this test reflect some of the findings of other research workers concerning the personality traits of creative scientists (Barron, 1969; Getzels and Csikszentmihalyi, 1967; Hudson, 1963, 1966, 1966a, 1967, 1968; Kinsbourne, 1968; MacKinnon, 1962, 1966; Pearce, 1968; Shapiro, 1968; Taylor, 1963, 1963a; Taylor and Barron, 1963; Walberg and Welch, 1967). Test 4 has two sections.

Section 1 Attitude Scale

The basic form of this section follows the 'Personal Qualities Questionnaire' designed by Hudson (1966, p.167). Seven of his original 30 items+ appear in the present

---

+Mixing well socially, being neat and tidy, having a low opinion of yourself, being highly imaginative, having set opinions, accepting expert advice, and being a good team member.

36 item scale. Some of the new items designed by the investigator are more specifically anchored in the science domain e.g. 'Doing a dangerous experiment', but most are concerned with basic attitudes that cut across subject matter areas.

The six scales can be referred to tentatively as:

(1) Respect for Authority

(2) High Conformity

(3) High Risk

(4) Serious Minded Worker

(5) Freedom of Emotional Expression

and (6) High Self Concept (ego strength)

While the test was designed with six scales in mind it was fully realised that items were in no way discrete to each scale. Since items are responded to on a rubric which moves from 'highly approve' to 'highly disapprove' the same item can be interpreted as belonging to two (or more) different scales at the same time e.g. highly disapproving of 'getting into trouble' can be seen as Respect for Authority (Scale 1) while highly approving of 'getting into trouble' can be seen as High Risk (Scale 2).

The initial stage of design represents an empirical approach to measurement whereby items thought to differentiate high/low scorers on fluency and flexibility are included experimentally (Cattell, 1967; Edwards, 1970). One or two items such as 'smoking' were used without being initially assigned to a specific scale in the belief that such items might prove efficient in differentiating high and low scorers in the two cognitive domains. The tentative grouping of items appears as follows:

Scale 1 Respect for Authority

Item No.	Item
1	being obedient
4	sticking to the truth
13	getting into trouble (as a negative) (Scale 2 also)
15	showing respect for teachers
18	being neat and tidy (Scale 3 also)
26	keeping quiet in class
30	being very polite (Scale 3 also)
32	questioning the truth of a text book (as a negative)
33	doing one's duty (Scale 3 also)

(Nine items)

Scale 2 High Risk

Item No.	Item
5	being highly imaginative
6	being cautious (negative)
9	running risks
13	getting into trouble (Scale 1 also)
14	doing new things (Scale 3 also)
16	teasing people
19	investigating the unusual (Scale 5 also)
20	taking the lead (Scale 6 also)
22	doing a dangerous experiment
23	smoking
25	making jokes (Scale 5 also)
27	arguing with people
28	showing off
31	daydreaming (Scales 4,5 also)
34	exaggerating (Scale 5 also)

(Fifteen items)

Scale 3 High Conformity

Item No.	Item
8	having set opinions
14	doing new things (negative, Scale 2 also)
18	being neat and tidy (Scale 1 also)
24	accepting expert advice
30	being very polite (Scale 1 also)
33	doing one's duty (Scale 1 also)
36	being a good team member

(Seven items)

Scale 4 Serious Minded Worker

Item No.	Item
2	working hard
12	doing homework thoroughly
21	getting everything correct
29	taking things seriously
31	daydreaming (negative, Scale 2, 5 also)

(Five items)

Scale 5 Freedom of Emotional Expression

Item No.	Item
7	being enthusiastic
10	having self control (negative)
19	investigating the unusual (Scale 2 also)
25	making jokes (Scale 2 also)
31	daydreaming (Scales 2, 4 also)
34	exaggerating (Scale 2 also)
35	not hurting other people's feelings

(Seven items)

Scale 6 High Self Concept

Item No.	Item
3	having a low opinion of yourself (negative)
11	mixing well socially
17	shrugging off criticism
20	taking the lead (Scale 2 also)

(Four items)

Item membership in the above scales is tentative, dependent upon factor analysis indicating alternative and more parsimonious ways of clustering items and suggesting scales.

<u>Instructions</u>

<u>Test 4, Attitude Scale</u>

"Which of the following do you approve or disapprove of?

Place a tick under the heading that best represents <u>your</u> feeling.

Answer honestly. Remember that this is confidential, and that none of your teachers will be seeing your papers.

Work fairly quickly."

A copy of Test 4, Section 1, Attitude Scale, appears on the following page.

Sample Copy, Test 4, Section I, Attitude Scale

Which of the following do you approve or disapprove of?
Place a tick under the heading that best represents **your** feelings
Answer honestly. Remember that this is confidential, and that none of your teachers will be seeing your paper
Work fairly quickly

	Strongly Approve	Mildly Approve	Neutral or Not Sure	Mildly Disapprove	Strongly Disapprove
	+ +	+	?	–	– –
being obedient					
working hard					
having a low opinion of yourself					
sticking to the truth					
being highly imaginative					
being cautious					
being enthusiastic					
having set opinions					
running risks					
having self control					
mixing well socially					
doing homework thoroughly					
getting into trouble					
doing new things					
showing respect for teachers					
teasing people					
shrugging off criticism					
being neat and tidy					
investigating the unusual					
taking the lead					
getting everything correct					
doing a dangerous experiment					
smoking					
accepting expert advice					
making jokes					
keeping quiet in class					
arguing with people					
showing off					
taking things seriously					
being very polite					
day dreaming					
questioning the truth of a text book					
doing one's duty					
exaggerating					
not hurting other people's feelings					
being a good team member					

Test 4. Section 2. Personal Preference Questionnaire

Description

A small number of multiple-choice questions concerned with personal choice in the science domain were derived from aspects of the Nuffield Science Foundation Teaching Programme (1967) and from other related research to form the basis of Section 2 of Test 4.

Instructions

Test 4. Personal Preference Questionnaire

Which would you rather do in each of the following?

Place a tick alongside your choice.

Remember this work is confidential.

Work fairly quickly.

Items included in the Personal Preference Questionnaire appear on the following page.

<u>Sample Copy of items in Personal Preference Questionnaire</u>

1. If an experiment failed would you rather
   (a) Find out why it failed? ☐
   or (b) Start a different one? ☐

2. When doing an experiment do you usually like to
   (a) Work alone? ☐
   or (b) Work in a group? ☐

3. Would you prefer to
   (a) Test the accuracy of a measuring instrument? ☐
   or (b) Design a new measuring instrument? ☐

4. Would you prefer to study how animals
   (a) Co-operate? ☐
   or (b) Compete? ☐

5. In building something from plans do you prefer to
   (a) Make up your own plans? ☐
   or (b) Work from ready-made plans? ☐

6. Do you prefer to
   (a) Work out theories? ☐
   or (b) Observe results? ☐

7. Do you prefer to have
   (a) One or two friends? ☐
   or (b) Many friends? ☐

8. If more time were given to science at school would you like to
   (a) Read about science? ☐
   or (b) Try out ideas of your own? ☐
   or (c) Work with the teacher on experiments? ☐

9. Do you think pupils should be encouraged to
   (a) Reach a good even standard of work? ☐
   or (b) Follow their special interests? ☐

C. Test 5, AH5 Test of High-Grade Intelligence (Part I)

Introduction and Description

After examining several group intelligence tests the AH5 (Part I) was selected as being the most appropriate because of its theoretical structure (Mehryar and Shapurian, 1970); its high demands on thinking (Liggett, 1965); its age range (separate norms for 13,14,15,16,17 and 18 year old pupils of high ability); and its suitability for being used cross-culturally (Keats, 1959; Mehryar and Shapurian, 1970). The test has also been used in other studies in England (Hudson, 1966, 1968; Johns, 1970; Povey, 1970) and in Wales (Shouksmith, 1970) where dimensions of cognitive style and divergent thinking have been examined.

D. Teacher and Peer Rating Scale Information

Teacher Evaluation Questionnaire

While there is considerable research on teacher perception of divergent thinking in pupils (Cropley, 1969; Getzels and Jackson, 1962; Gowan, Demos and Torrance, 1967; Hamilton, 1970; Massialas and Zevin, 1967; Shigaki, 1970; Taylor and McKean, 1968; Williams, 1967) little appears known concerning the cues that teachers employ in making such observations.

In the initial pilot run of the investigation science teachers were asked to rate each pupil in their class on such traits as originality and ability in science and to consider the degree to which they enjoyed teaching each pupil, on a five point scale. The later aspect was aimed at assessing the degree to which high scorers in fluency and flexibility, in so far as they can be associated with divergent thinkers, might be considered somewhat troublesome to many science teachers. From the pilot study, however, it became apparent that the teacher's task of rating every pupil in the class was too burdensome. It was therefore decided to ask each teacher to nominate the four most outstanding pupils in each rating category. It will be noted that there is no strict equivalence between test score measures (fluency and flexibility) and the rating categories. No precise degree of parallelism was sought, since, to ask science teachers to nominate the most 'fluent' and 'flexible' pupils in science would have been to use terms with which they might be unfamiliar. On the other hand the degree to which 'originality' and 'highest ability in science' might correlate with test scores in fluency and flexibility could prove interesting.

A sample copy of the Teacher Evaluation Questionnaire follows:

TEACHER EVALUATION

CONFIDENTIAL

SCHOOL................

CLASS.................

Would you please assist this project by filling in the following information:

Ability in Science

Please list the names of four pupils in this class who show the highest ability in science.

1 ______________________________

2 ______________________________

3 ______________________________

4 ______________________________

Originality in Science

Please list the names of four pupils who display some of the most original and unusual ideas in science.

1 ______________________________

2 ______________________________

3 ______________________________ (any additional comments welcome)

4 ______________________________

Behaviour in Classroom

Please list the names of four pupils in this class whom you enjoy teaching **most**.

1 ______________________________

2 ______________________________

3 ______________________________

4 ______________________________

Please list the names of four pupils in this class whom you enjoy teaching **least**.

1 ______________________________

2 ______________________________

3 ______________________________

4 ______________________________

Peer Rating Information

Conflicting reports on the value of peer ratings as a technique for identifying traits associated with divergent and creative pupils occur in the literature (McHenry and Shouksmith, 1970; Kurtzman, 1967; Yamamoto, 1964).

In the present study limited use was made of the technique. Peer ratings were sought concerning (i) those pupils who were perceived as working the hardest at science, and (ii) those who were seen as displaying the most original and unusual ideas in science. Renzulli (1970) has pointed out that teachers and pupils seem to equate original with hard working or high achieving in making evaluations on rating scales. Alerting pupils to both traits separately in the present study may overcome this to some degree and increase reliability and validity. Each of the 506 pupils in the study was asked to list two classmates for each of the two categories.

E. Pilot Study

The tests were administered to an English Grammar School where responses to the following questions were sought:

What specific time limits should be set? Are the instructions easily understood? Do they produce a relaxed atmosphere yet motivate pupils to the task? And finally, do the pupils enjoy doing the test? Pupils in the pilot study were encouraged to write down comments about the test, particularly concerning words or phrases not understood.

Time Limits

Observations from the initial study suggested that when half to three quarters of the class appeared unable (or unwilling) to generate any more responses the upper limit for time and rapport had been reached for group administration. This seemed an appropriate time to stop responses for each stimulus word. Such a policy allowed the majority of pupils to have reached that point on response gradients where it is reported that a higher percentage of low frequency words or ideas are generated (Mednick, 1962; Renzulli, 1970; Wallach and Kogan, 1965). On the other hand it does not push the majority of the class to the point of boredom and unrest which would occur if everyone, or nearly everyone, were to work themselves out (although it would be interesting to administer the test individually on this basis). Unfortunately, direct comparisons between group and individual administrations would be difficult since, for some, the face to face situation of individual administration increases nervousness (anxiety) with a consequent decrease in fluency and flexibility (Cramer, 1968; Krop, Alegre, and Williams, 1969; McCutcheon, 1970; Pankove and Kogan, 1968; Ward, 1969; Wood, 1970). If time limits were to be extended to the upper limits of endurance for group administrations a wider range of scores at the upper end of the scale would be obtained, but at the expense of continuing rapport. It is doubtful whether any consistently high scorers in the written response situation are penalized either by the present time limits or by the statistical procedure of taking $\pm$1SD as cut-off points for designating high and low scorers.

The present time limits allow for a wide range of individual differences in fluency and flexibility to be manifested, and appear quite reasonable for classroom group administrations (Kogan, 1967).

On this basis, the following time limits were determined for the tests:

Test 1, Science Vocabulary

Test time, 3 minutes. With administration time, 4 minutes.

Test 2, Definition of Scientific Concepts

Test time, 12 minutes. With administration time, 14 minutes.

Test 3. Uses for Chemical Substances and Physical Objects

Test time, 6 minutes. With administration time, 8 minutes.

Test 4. Attitude Scale

No time limit was placed on this test, all pupils being allowed to finish. Average completed test time was 10 minutes. With administration time, 12 minutes.

Information Page

Test time, 4 minutes. Administration time, 3 minutes for beginning and ending test.

Test 5. AH5 Test of Intelligence, Part I

Test time, 20 minutes. With administration time, 30 minutes.

Total Test Time, 55 minutes.

Total Administration and Test Time, 76 minutes.

Testing Modifications, Instructions and Rapport

The earlier forms of the tests were modified in several ways. A controlled presentation of each stimulus word was found necessary whereby each word was printed on the blackboard and presentation timed by stopwatch. In Test 2 delineation between the beginning and ending of each set of responses was found necessary, and incorporated into the instructions. In Test 4 more strict control over column width was required.

The method of presenting instructions orally, while pupils paced the reading from their own script, seemed to work satisfactorily. Pupils also appeared to enjoy the tasks set by the tests, a condition which was later seen to hold across administrations in each country.

The open-ended tests in science, the AH5 and the teacher evaluation sheet were all administered by the investigator on the same day to each class (cf. Getzels and Jackson, 1962).

CHAPTER THREE

SCORING RESPONSES

Operational definitions of fluency and flexibility, the formulation of taxonomies to guide scoring, and interscorer reliability results are discussed in this chapter.

(i) Operational Definitions of Fluency and Flexibility

Fluency of thinking was defined in this study as the number of written responses representing a basic unit of thought or idea. Obviously no universally acceptable definition of an idea can be given. For the purpose of scoring fluency in these tests an 'idea' was taken as the smallest unit of coherent meaning. In many circumstances single words (often nouns) such as 'vacuum' represent an idea as much as the sentence 'It is a vacuum'.

The scoring of fluency for Test 1 (Science vocabulary, or 'words to do with science') was straightforward, since in this context fluency was premised on a word count. The fluency score for Test 1 was therefore the most objective measure of the test and had the highest interscorer reliability.

However, responses to Tests 2 and 3 caused some problems. Single words or short phrases could be scored fairly easily but it was necessary to break down complex phrases and sentences into more basic units of expression and thought. It was therefore particularly important to be alert to elaboration within the basic unit of expression e.g. "Space is lonely/dark/and mainly a vacuum/" scored three for fluency. "Time is measured in minutes/hours/days/ weeks/ and months/" scored five. "Time goes quickly/ and slowly" scored two - representing as they do, opposite aspects of psychological time (cf Wallach and Wing, 1969, p.35).

The general aim in scoring fluency was to maximize the number of specific units. This has the effect of increasing the number and range of responses and hence individual differences. It also has the effect of increasing the reliability of scores.

While there will always be a margin of error in interpreting responses, the above principle at least helped to reduce it. The logical extension of the principle of maximizing the specificity of components of each phrase or sentence would occur at the point where every word counted as a unit. Further research could investigate the correlation between the present intermediate approach (words and short phrases) with such an extreme one as a word count. For the moment it should be realised that different scoring criteria are sensitive to individual differences in response style, and that changing the criteria for fluency would change the status of scorers to some degree. This would be true of Tests 2 and 3 in the present investigation, but not of Test 1, since this is in fact based on a word count.

Flexibility of thinking was defined in this study as the number of category changes in the sequence of words and/or ideas. A category may be defined as: a systematic arrangement, a classification, an organization, a codification, a class based on essential or fundamental considerations, a taxonomy.

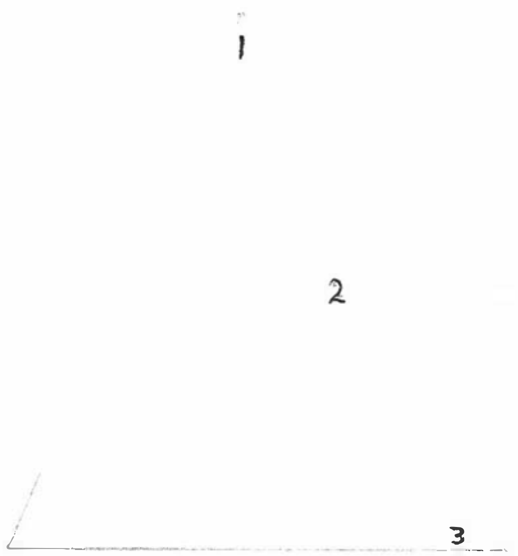
A category change occurs when responses switch from one category membership to another. Thus A,A,A,/B,/C,C,C,C,/D,D/E,/F/G represents six sequential changes of category (seven categories) and hence a flexibility score of six i.e. N - 1 where N = total number of sequential categories). A,/B,/A,/B,/A,/B,/A also represents a flexibility score of six (but with only two categories). The latter pattern, which did not occur in any script, would be a highly contrived and difficult pattern with which to persist. The study was therefore concerned with flexibility of thinking defined as category change and not with the number of categories per se.[+] The problem of what is a 'category' still remains, however. There are 'categories and categories', e.g. 'chemical apparatus' which could represent a single category under one set of definitions could be broken into several categories under another set of definitions, e.g. apparatus made of glass, apparatus used for measuring, apparatus used for heating, etc. Thus 'test tube, pipette, and bunsen burner' could represent three categories (A,/B,/C,/) under the latter definition, and one category under the former. It all depends on the width of categories.

The present policy avoided a highly specific basis for categorization (e.g. each chemical apparatus determining its own category) in preference for an intermediate one. Such an approach allowed the flexibility score to be something different from the fluency score (which would not occur if every response, or nearly every response, was interpreted as a change of category). These

---

[+]Both category change and number of categories were measured in the interscorer reliability studies. More intensive research with smaller numbers could investigate the triple components of fluency, category numbers and category change.

different approaches to category width may be represented diagrammatically in Figure 1.



1 = A few broad categories. Carried to its extreme only one category e.g. in Test 1, all responses could be classified under the single category "words to do with science".

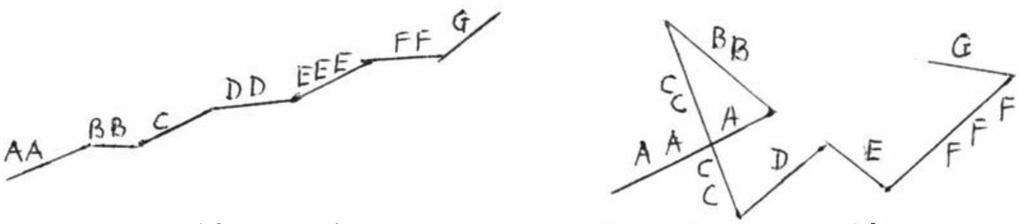
2 = Intermediate number of categories. This offers adequate range of scores.

3 = Highly specific categorization. Every word a category i.e. Fluency = Flexibility.

FIGURE 1

Diagrammatic Representation of Three Category Widths

While some pupils displayed contiguous responses with strikingly different category membership (sharp gradient), others shaded one response into the next in a particularly subtle way (smooth gradient). Such a gradient placed considerable strain on the scoring taxonomy. It was with this latter type of response-gradient that most of the variability in flexibility scores occurred in the interscorer reliability study. To some extent responses reflecting smooth gradients still remain difficult to score, despite a fairly rigorous interscorer reliability study. Smooth and sharp gradients representing different patterns of response-category change are illustrated in Figure 2.



1 smooth gradient   2 sharp gradient

FIGURE 2

<u>Smooth and Sharp Response Gradients Representing Different Patterns of Response-Category Change</u>

Translated into actual responses the above gradients appear as follows:

1. <u>Smooth gradient</u>

A A B B C D D
heat, light/ energy, power/ electricity/ amp, volt/

E E E F F G
electrons, neutrons, ions/ atom, molecule/ water

2. <u>Sharp gradient</u>

A A A B B
Chemistry, Physics, Biology/ photosynthesis, hydrolisis/

C C C C D
Pasteur, Archimedes, Dalton, Watson/ experiment/

E F F F G
microscope/ sodium, sulphur, ammonia/ equivalent weight.

These two kinds of gradients have been related, in one study (Vick and Jackson, 1967), to two types of cognitive style - 'levellers' and 'sharpeners'. Levellers are described as being relatively insensitive to differences between similar stimuli while sharpeners are predisposed to emphasize them. These two styles may be contributing to the two kinds of gradient in the present investigation.

(ii) Taxonomies

Responses for each test stimulus were categorized into a taxonomy which acted as the scoring guide for that particular test.

Category membership in the first place was premised on common or shared characteristics. It was realized that many bases exist for establishing these and that category inclusion or exclusion was a matter of degree. As a result, systems of sets and subsets were employed in the establishment of some categories (e.g. 2A, 2B, 2C). Similarly, 'generic' and 'specific-example' categories were employed in some cases (e.g. "chemical reactions" - generic; "combustion", "oxidation" - specific).

Since no a priori view was taken concerning either the names of categories or the number of categories for each test stimulus, the total population of responses for each test was used as the basis for the formulation of the taxonomy for that test.*

Comparable numbers of categories were sought between parallel or alternative test stimuli within each test. It was necessary for each taxonomy, once formulated, to be resistent to superficial aspects of response, (for example, format, writing style), yet be sensitive to the range of responses premised on fluency and flexibility of thinking. Examples of different format and style appear as follows:

---

*Thus the responses of research scientists might suggest quite different categories from those of the present sample. Yaroshevskii (1969) stresses the need to look at such categorical nets as representative of the logical development of science in a particular sample (group or culture) at a particular time in history, emphasizing that such dimensions of thinking can have no meaning outside such a specific context.

SPACE "consists of stars
is a vacuum
has planets
is being explored
no life as far as we know
goes on and never ends
emptiness
no gravity
infinite
rockets travel into space
America vs Russia
lack of space
a noun"

SPACE "You can fill space with many different kinds of things. It has been used for storing things in. Space on earth is becoming a little more scarce as the population increases and large cities sprawl over the landscape."

OXYGEN "used for oxy-acetalene welding
helping chemical experiments
under water diving
pilots need oxygen masks
for testing sulphur dioxide
as a cleansing agent
makes things burn hotter"

OXYGEN "I would like to develop a sort of oxygen we can breathe which does not support large fires so that when a forest fire occurs, when it reaches a certain temperature it will extinguish itself."

It is not claimed that any of the taxonomies developed here holds a position of any special privilege. Different taxonomies are useful for different purposes. A comparison of several different taxonomies utilizing a smaller sample would prove interesting. However, the aim of the present study was to establish one scoring taxonomy premised on a reasoned interpretation of categories that would stablise the scoring procedures and offer a fairly high degree of consistency and interscorer reliability. If this could be achieved, different scorers should produce similar numbers of category changes and also agree fairly closely concerning the actual <u>position</u> of the category change. Were this not so, two different scorers could arrive at the same number of category changes, but based on quite different taxonomies e.g.

Scorer 1 (Taxonomy 1) AAA/B/CC/D/EEE/F (Five category changes).

Scorer 1 or 2 (Taxonomy 2) AA/BB/CCC/D/EE/F (Five category changes).

A graphic scoring record of category change was made for Test 1 to check this phenomenon. The representation of this procedure together with some evidence of its reliability as derived from a 're-mark' carried out eight weeks later can be seen from Figure 3.

FIGURE 3

Graphic Record of Responses Indicating Position of Category Change and Number of Categories for Original and Re-mark Scoring

The following taxonomies guided the scoring of the total sample and contributed toward a satisfactory level of objectivity and interscorer reliability. Taxonomies for Test 1, Science Vocabulary, and Test 2, Definition of Scientific Concepts, 'SPACE', follow. The remainder appear in Appendix A.

Taxonomy Test 1. Science Vocabulary - Guide Lines for Scoring Responses to Science Vocabulary

1.	Subject Names	chemistry, physics, biology, astronomy, radio physics
2.	Chemical Processes	photosynthesis, electrolysis, distillation, hydrolysis
3.	Names of Scientists	Pasteur, Archimedes, Galileo, Curie, Einstein, Watson
4.	Science Laboratory Equipment	test tube, retort, tripod, bunsen burner, spatula, beaker, filter
5.	Personal Writing and Recording Equipment	science book, lab. book, ruler
6.	Specifically Named Elements, Compounds and Gases	potassium, magnesium, sulphur, phosphorous, nickel, copper, mercury, oxygen, nitrogen, CO_2, Xenon

Specific grouping

7. Observation, experiment, hypothesis, results, discovery, theory

8. Valency, equivalent weight, chemical equivalence, atomic number

9. Specific gravity, density

10. Weight, mass, matter, force, acceleration, gravity

11. Radioactivity, nuclear reactor, radiation, X-rays, geiger counter

12. Electricity, magnetism

13. Heat, light, sound, refraction, reflection

14. Amp, amperage, volt, voltage, voltameter, ammeter

15. Animal, plant, insect

16. Air, water

17. Sex, reproduction

18. Power, energy, work

19. Acid, base, alkali, liquids, solids, compounds, elements

20. Electrons, neutrons, ions, protons, atoms, molecules, nucleus

21. Solids, liquids, gases

22. Wires, batteries, contacts, cell, resistor, capacitor.

Taxonomy Test 2, Definition of Scientific Concepts - Guide Lines for Scoring Responses to 'SPACE'

1A **Bodies in outer space**

e.g. planets (single words)
stars
sun
earth
moon
meteorites
nova
quasers
solar system
galaxies
constellations
universes

1B **Description of universe**

e.g. surrounds heavenly bodies
makes up part of the universe
area in which earth is located
positioning in space
everything in the world
outer space

1C **Structure, composition**

e.g. composed of different elements
dense parts
no weight in it
no gravity
molecules and atoms
energy
all matter occupies it
contains modulating waves
light travels through it (rate, speed)
sound doesn't travel through it

2A **Study of outer space**

e.g. studied by man
research in
scientific study
astronomy
radio telescope

2B **Conquest of, exploration of outer space**

e.g. space travel
disasters, accidents in outer space
voyages
space ships
astronauts
satellites
rockets
U.F.O's
flying saucers
impenetrable
the last boundary
orbits
cost of, expenditure on space exploration

2C <u>Methods of propulsion</u>
e.g. rocket fuels
atomic energy
use of gravity towards

2D <u>Political conquest of</u>
e.g. Russia vs U.S.A.
France and Britain behind
Space race
Cape Kennedy
NASA

2E <u>Specialized communication system</u>
e.g. T.V. Satellites
Telstar
Earlybird
weather stations

2F <u>Life forms in space</u>
e.g. may hold other civilizations
no good living conditions
new worlds
alien men
Mars and vegetation

3A <u>Empty/full</u>
e.g. emptiness
void
nothingness
nowhere
really invisible

4A <u>Size and distance</u>
e.g. immense
vast
huge
large or small
infinite
endless
isolated
dimension
forever
continuous
universal
3 - dimensional
4th dimension
5th dimension
goes in all directions
distance
measured in hundreds of light years

5A <u>Location</u>

e.g.
space is location
a location
an infinite number of points are contained in space
area
no area
planes
volume

5B <u>Space between two things</u>

e.g.
a clearing
a gap
a vacancy
wide open space
a space in which one can move
a parking space
can turn around in this space
something to fill in
room
openness
blank

6A <u>Space needs on earth</u>

e.g.
India needs more space

7A <u>Air, comparison with atmosphere</u>

e.g.
lack of air
little oxygen
gases
little pressure
vacuum
atmosphere
clouds
dust
cosmic dust
cosmic rays
cannot breathe in it

8A <u>Empathy, values, beliefs</u>

e.g.
lonely
cold
is precious
dismal
dangerous
dark
bleak
dreary
quiet

8B <u>Known, unknown</u>

e.g.
unknown mysteries
puzzlement
man dreams to possess it
beyond comprehension
complicated
controversial
untouchable
cannot imagine space
it has fascinated people for a long time
unimaginable

9A <u>Derivation of word</u>

e.g.
a noun
it forms the adjective spacious
the set of all points
from latin spatium

10A <u>Entertainment</u>

e.g.
basis of T.V. programs
science fiction
Startrek

11A <u>God</u>

e.g.
birth of Christ
associated with some religions

12A <u>Space of time</u>

e.g.
a space of time
spacetime

(iii) <u>Interscorer Reliability Results</u>

Most of the principles and problems of scoring previously discussed grew directly out of attempts to establish satisfactory interscorer reliability.

In the first place such reliability was a way of testing the scoring taxonomy and the operational definition of fluency and flexibility. Three scorers were employed in order to gain greater insight into the amount and direction of discrepancy, and to see if any trends between three sets of scorer-correlations would occur, i.e. between scorers 1 and 2, scorers 1 and 3, and scorers 2 and 3. This strategy aimed at yielding more information on interscorer reliability than one employing only two scorers.

In the second place it was crucial to establish satisfactory interscorer reliability at this stage otherwise little faith could be placed in the results and conclusions obtained from scoring the major sample.

Three markers were chosen for their previous experience in content analysis and in science. Three sets of twenty scripts were then selected at random for the first three trial runs. For the final study from which interscorer reliability figures were calculated, sixty scripts were chosen, representing approximately every fifteenth script from the major sample.

The following procedures guided the development of interscorer reliability:

(1) An initial meeting discussed operational definitions of fluency and flexibility and devised tentative taxonomies for scoring sample responses.

(2) A random selection of 20 scripts was then marked by each marker; records were kept of fluency and flexibility scores.

(3) Discussions were held to consider discrepant cases; operational definitions and taxonomies with stricter guide lines for scoring were devised; records were kept of the exact points on a pupil's script where a category change occurred, as well as of the number of categories; each test was marked separately for fluency and then for flexibility, in order to reduce contamination of scores between tests within each individual's responses; results of each test (1,2,3) were kept discrete for recording purposes.

(4) Another random sample of 20 scripts was marked; taxonomy and guide lines were further refined; remaining discrepant scores were discussed. The final random sample of 20 scripts was then marked before the interscorer reliability marking was untertaken.

(5) A selection of 60 scripts representing

approximately every fifteenth script from the major sample was then made for the interscorer reliability study; each test was marked and recorded separately, and raw scores pooled for the three tests for the three markers as follows:

Test 1 = 360 scores, Test 2 = 1080 scores, Test 3 = 720 scores. Total test scores = 2160. Analysis of final scores showed some pockets of discrepancy still remaining, although the overall picture looked promising. Reasons for continuing variability in interscorer reliability were discussed; some suggested reasons appear below:

(a) smooth flexibility gradients in responses of some pupils.

(b) different width of categorizing tendencies between scorers.

(c) different degrees of scientific literacy of scorers.

A sample of results from the interscorer reliability study appears in Table 3. The remaining results appear in Appendix B.

TABLE 3.

Interscorer Reliability Data : Raw Scores for Three Tests by Three Scorers

Column 1 = Scorer 1
Column 2 = Scorer 2
Column 3 = Scorer 3.

Sixty Scripts	Test 1						Test 2						Test 3					
Script No.	Fluency			Flexibility			Fluency			Flexibility			Fluency			Flexibility		
15	13	13	13	6	4	4	5	11	10	4	4	4	7	6	7	6	4	6
							3	7	7	2	3	3						
							3	4	4	2	2	2	7	7	7	5	6	6
30	20	20	20	10	15	10	11	9	9	7	6	7	9	9	9	8	6	8
							7	7	6	3	4	4						
							7	8	7	6	7	6	7	7	7	5	6	6
45	39	38	39	21	21	20	7	9	8	4	5	4	5	6	6	3	3	3
							6	8	7	5	4	6						
							8	10	8	5	5	5	6	6	6	4	5	5
60	19	19	19	13	13	10	7	8	8	6	4	6	5	6	5	4	3	4
							6	7	6	5	4	5						
							8	11	10	4	4	4	5	5	5	4	1	4
75	19	19	19	8	6	5	9	12	10	7	6	7	6	9	7	5	7	6
							7	7	8	6	6	7						
							7	8	7	6	4	6	10	10	10	5	2	5
90	38	37	38	12	17	14	7	8	9	6	6	6	4	4	5	3	5	4
							4	8	6	3	5	3						
							4	11	10	3	4	3	3	5	3	2	3	3
105	29	29	29	17	20	18	10	10	10	8	8	8	3	4	4	2	3	3
							6	9	9	5	5	5						
							5	10	10	4	6	4	4	4	4	3	3	3
120	30	31	31	10	11	10	8	15	14	6	8	6	5	7	7	3	5	4
							7	11	11	5	4	4						
							4	7	7	3	4	3	6	8	7	4	4	4
135	26	26	26	13	14	11	5	5	5	4	2	4	6	6	6	5	5	5
							5	6	7	4	4	4						
							5	6	8	4	5	4	7	7	7	3	6	4
145	33	33	33	17	16	15	11	11	11	9	8	7	8	7	8	7	6	6
							6	10	10	4	4	4						
							6	11	4	4	5	4	13	11	13	12	8	12
150	27	27	27	23	21	18	7	9	11	5	6	6	5	5	5	4	4	4
							4	4	4	3	3	3						
							5	6	5	4	4	3	12	12	12	10	10	10

Means, significance of difference between means, standard deviations, and interscorer reliability coefficients (Pearson's product-moment) for tests separately, and for tests combined, were computed for the above data and appear in Tables 4 and 5.

TABLE 4

Interscorer Reliability Data: Means, and S.D.'s for Three Scorers on Three Tests

TEST 1 Words to do with Science

Thinking Variable	Scorer Variable	Mean	S.D.
	1	29.65	9.4935
FLUENCY	2	29.50	9.3976
	3	29.63	9.4920
	1	15.41	5.4200
FLEXIBILITY	2	16.26	5.7817
	3	14.76	5.3020

TEST 2 (Three Tests) SPACE, TIME, ENERGY

Thinking Variable	Scorer Variable	Mean	S.D.
	1	10.57	5.3166
FLUENCY	2	10.21	4.1230
	3	10.69	4.6297
	1	6.46	3.3779
FLEXIBILITY	2	5.70	2.6708
	3	6.45	3.2703

TEST 3 (Two Tests) ACID or OXYGEN; WHEEL or MAGNET

Thinking Variable	Scorer Variable	Mean	S.D.
	1	8.53	4.2699
FLUENCY	2	8.08	3.6596
	3	8.16	3.9207
	1	5.85	3.2640
FLEXIBILITY	2	5.22	2.7145
	3	5.88	3.1414

t-tests yielded no significant differences between sets of scorer means.

TABLE 5

Interscorer Reliability Data: Correlation Coefficients for Three Scorers on Three Tests

	Scorer 1	Scorer 2	Scorer 3
TEST 1 (Fluency)[+]			
Scorer 1	-----	.9989	.9998
Scorer 2	.9989	-----	.9991
Scorer 3	.9998	.9991	-----
TEST 1 (Flexibility)			
Scorer 1	-----	.8987	.9319
Scorer 2	.8987	-----	.8971
Scorer 3	.9319	.8971	-----
TEST 2 (Fluency)			
Scorer 1	-----	.8518	.9283
Scorer 2	.8518	-----	.9248
Scorer 3	.9283	.9248	-----
TEST 2 (Flexibility)			
Scorer 1	-----	.8488	.9427
Scorer 2	.8488	-----	.8534
Scorer 3	.9427	.8534	-----
TEST 3 (Fluency)			
Scorer 1	-----	.9411	.9667
Scorer 2	.9411	-----	.9583
Scorer 3	.9667	.9583	-----
TEST 3 (Flexibility)			
Scorer 1	-----	.9071	.9547
Scorer 2	.9071	-----	.9208
Scorer 3	.9547	.9208	-----
CORRELATION MATRIX for 3 Scorers on Total Test Data			
Scorer 1	-----	.9645	.9824
Scorer 2	.9645	-----	.9744
Scorer 3	.9824	.9744	-----

[+] word count

No significant difference between scorer combinations occurred although there was a tendency for scorers 1 and 3 to agree most, and 1 and 2 least.

There was little difference between the figures for fluency and flexibility, although coefficients for fluency were on the whole slightly higher.

Finally, there was little difference between the tests, although Test 1 held a slight superiority over Test 3 which in turn held a slight superiority over Test 2. The lowest coefficient (.8488) was for Test 2 on flexibility, between scorer 1 and 2. Correlation coefficients from both total test data and the more stringent individual test data suggest that satisfactory interscorer reliability has been achieved.

Scorers 1 and 3 thereafter scored the responses (64,128) for the total sample of 506 scripts. The results can be seen in Table 6 "Basic Statistics for Samples by Country', page 79.

CHAPTER FOUR

STATISTICAL METHODOLOGY AND RESULTS

Statistical Methodology in the present investigation is related to the aims and purposes of the study, viz, to examine:

(i) the nature of scores in fluency and flexibility of thinking - their distribution, mean, and range within a sample of pupils of high ability in science;

(ii) the relationship between fluency and flexibility of thinking and measured intelligence (total culture samples);

(iii) the relationship between high and low scorers on specific cognitive and personality variables ($\pm$1SD samples by country); and

(iv) the relationship between high and low scorers for observed and expected frequencies according to sex, classroom, teacher and peer variables.

While a general methodological plan was determined before statistical analysis was commenced, there was progressive opportunity to evaluate results, and to consider the appropriateness of each step. All computations were carried out on the Massey University IBM 16/20 with standard programmes.+

Statistical methodology incorporated the following sequential procedures:-

1. Determination of means, standard deviations and range of scores for all cognitive variables.
2. Plotting of frequency distribution curves for the examination of normality of distribution.
3. Translation of all cognitive scores to normalized standard scores (T-scores).
4. Examination of the influence of age as a significant variable within a country sample by use of tests of significance of difference between means (t-tests).
5. Correlational studies between all measures of fluency and flexibility and measured intelligence.

+Computer print-outs and programmes may be obtained on request. General formulae employed in the study appear in Appendix C . Appreciation for advice concerning statistical procedures is acknowledged to
Nola Gordon, Computer Dept., Massey University;
Dr G. Arvidson, Education Dept., Auckland University;
Professor Adcock, Psychology Dept., Victoria University;
Dr H. Beresford, Education Dept. Massey University;
Professor B. Hayman, Mathematics Dept., Massey University.

6. Centroid factor analysis, with orthogonal rotations, of (a) cognitive variables and (b) personality variables involved in the attitude scale, Test 4 Part 1.
7. Analysis of high and low scorers ($\pm$1SD) on specific cognitive and personality variables (t-tests).
8. Use of chi-square and Fisher's exact probability tests in the examination of observed and expected frequencies for high and low scorers (cognitive) according to sex, classroom contributions, teacher and peer ratings.
9. Test-retest reliability together with earlier interscorer reliability studies.

Raw scores were coded separately by:

(a) Country: England, U.S.A. and New Zealand, and

(b) Age: Group I = 13.4 - 13.11 years,
Group II = 14.0 - 14.7 years,
Group III = 14.8 - 15.3 years, and

(c) Test variables (cognitive)

1. Intelligence = AH5
2. Fluency Test 1 = Fl 1
3. Flexibility Test 1 = Fx 1
4. Fluency Test 2A = Fl 2A
5. Flexibility Test 2A = Fx 2A

6. Fluency Test 2B = Fl 2B

7. Flexibility Test 2B = Fx 2B

8. Fluency Test 3A = Fl 3A

9. Flexibility Test 3A = Fx 3A

10. Fluency Test 3B = Fl 3B

11. Flexibility Test 3B = Fx 3B

12. Fluency Test 3C = Fl 3C

13. Flexibility Test 3C = Fx 3C

14. Fluency Test 3D = Fl 3D

15. Flexibility Test 3D = Fx 3D

(d) Test variables (personality)

Test 4, Part 1 - Attitude scale - 36 item correlation matrix for factor analysis

Part 2 - Personal preference questionnaire - 54 sets of probabilities (chi-square).

Statistical analyses at the commencement of the study employed a 3-ages x 3-countries approach with all data, which thereafter became a 3-country approach when it was found that age was not a significant variable within the framework of the present study. The 3-country approach necessitated three separate computations for statistical analyses, e.g. three factor solutions were required for each factor analysis study. These statistical procedures provided the basis for interpreting the results which will now be examined.

RESULTS

Results are organized under the following headings:-

A. COGNITIVE

(i) Description of test variable, sample size, number of responses, measures of central tendency and dispersion, and distribution curves.

(ii) (a) Within-country differences (t-tests), influence of age on scores from cognitive variables;

(b) Between-country differences (t-tests), significance of differences between means on cognitive variables.

(iii) Correlational studies employing cognitive variables.

(iv) High and low scorer comparisons.

(v) Factor analysis - cognitive variables.

B. PERSONALITY

(i) Factor analysis of Attitude Scale, Test 4, Part 1.

(ii) High and low scorer comparisons (fluency, flexibility) on attitude factor scales.

(iii) Chi-square analysis of high and low scorers (fluency, flexibility) on Personal Preference questionnaire, Test 4, Part 2.

C. SEX, CLASSROOM, TEACHER AND PEER RATING DIFFERENCES

Further chi-square analyses of high and low scorer comparisons (fluency, flexibility) with the above variables.

The concepts of reliability and validity will be discussed in the following chapter.

A. COGNITIVE

(i) Description of test variable, sample size, number of responses, measures of central tendency and dispersion, and distribution curves

Since age was not found to be a significant variable within the three age groupings of each sample (see p. 87 for t-tests on age differences), data concerning test variables, sample size, number of responses etc. are presented for each country as a whole in Table 6, 'Basic Statistics for Samples by Country', p. 79.

It will be noted from these results that, with a few exceptions, measures of central tendency and dispersion are similar for each sample. F-tests for homogeneity of variance yielded no significant differences in 30 out of the 45 sets of variances (see Appendix D for F-test results).

TABLE 6

BASIC STATISTICS FOR SAMPLES BY COUNTRY

Test Variable	Country	N	No. of Responses	Mean	Range	S.D.	s^2
	ENG	180	3008	16.7	4–25	3.9188	15.3570
AH5	USA	139	1864	13.4	3–25	3.9412	15.5331
	NZ	184	2661	14.4	6–25	3.6899	13.6154
	ENG	182	5075	27.8	11–50	8.0181	64.2899
Fl 1	USA	140	4015	28.6	9–58	8.7649	76.8235
	NZ	184	6636	36.0	19–62	7.5361	56.7928
	ENG	182	2450	13.4	1–28	4.8028	23.0669
Fx 1	USA	140	1923	13.7	2–28	4.8139	23.1736
	NZ	184	2864	15.5	1–31	5.3342	28.4537
	ENG	181	2109	11.6	4–26	4.3018	18.5055
Fl 2A	USA	140	1643	11.7	5–29	5.1963	27.0015
	NZ	184	2736	14.8	2–33	5.5381	30.6706
	ENG	181	1278	7.0	0–23	3.1112	9.6796
Fx 2A	USA	140	967	6.9	2–19	2.9792	8.8756
	NZ	184	1513	8.2	0–18	3.1638	10.0096
	ENG	182	2191	12.0	2–23	4.5266	20.4901
Fl 2B	USA	140	1521	10.8	1–29	4.9533	24.5352
	NZ	184	2762	15.0	2–39	6.4798	41.9878
	ENG	182	858	4.7	0–12	2.0180	4.0723
Fx 2B	USA	140	616	4.4	0–15	2.2847	5.2199
	NZ	184	1112	6.0	1–18	2.7695	7.6701
	ENG	138	1156	8.3	3–17	2.8211	7.9586
Fl 3A	USA	96	720	7.5	2–19	2.9379	8.6313
	NZ	145	1252	8.6	2–27	3.4637	11.9972
	ENG	138	695	5.0	0–14	2.4920	6.2101
Fx 3A	USA	96	464	4.8	1–13	2.2741	5.1715
	NZ	145	695	4.7	0–13	2.4178	5.8458
	ENG	44	394	8.9	2–20	4.0517	16.4163
Fl 3B	USA	42	274	6.5	2–15	3.2999	10.8893
	NZ	38	328	8.6	4–17	3.6791	13.5358
	ENG	44	259	5.8	1–12	2.5988	6.7538
Fx 3B	USA	42	177	4.2	0–10	2.5040	6.2700
	NZ	38	200	5.2	1–11	2.4349	5.9287
	ENG	138	1453	10.5	2–23	4.2686	18.2209
Fl 3C	USA	116	980	8.4	3–21	3.8069	14.4925
	NZ	151	1701	11.2	2–35	4.6053	21.2088
	ENG	138	832	6.0	0–14	3.1759	10.0863
Fx 3C	USA	116	540	4.6	0–12	2.2921	5.2537
	NZ	151	886	5.8	0–16	3.0346	9.2088
	ENG	44	382	8.6	2–19	4.1862	17.5243
Fl 3D	USA	22	163	7.4	4–11	1.9918	3.9673
	NZ	33	254	7.6	3–27	5.4571	29.7799
	ENG	44	245	5.5	1–13	2.9046	8.4379
Fx 3D	USA	22	120	5.4	3–19	2.0172	4.0691
	NZ	33	156	4.7	1–19	3.5643	12.7042

Consequently, there appears to be a measure of comparability in the sampling as seen through these results. More sensitive tests of significance of differences between means are nevertheless likely to indicate many differences on specific test variables. This, however, does not destroy the general overall picture of homogeneity of variance within the responses of the three samples by country.

Distribution curves (frequency polygons) were plotted for each test variable for each country. Distribution curves for AH5 and Fluency 1 and Flexibility 1 can be seen in Figures 4, 5 and 6. (The remaining distribution curves can be seen in Appendix E, Figures 1-12.)

It can be seen that each distribution curve for the sets of AH5 scores approximates a normal distribution, although there is a tendency for the English sample to be slightly negatively skewed and for the New Zealand sample to be slightly bimodal.

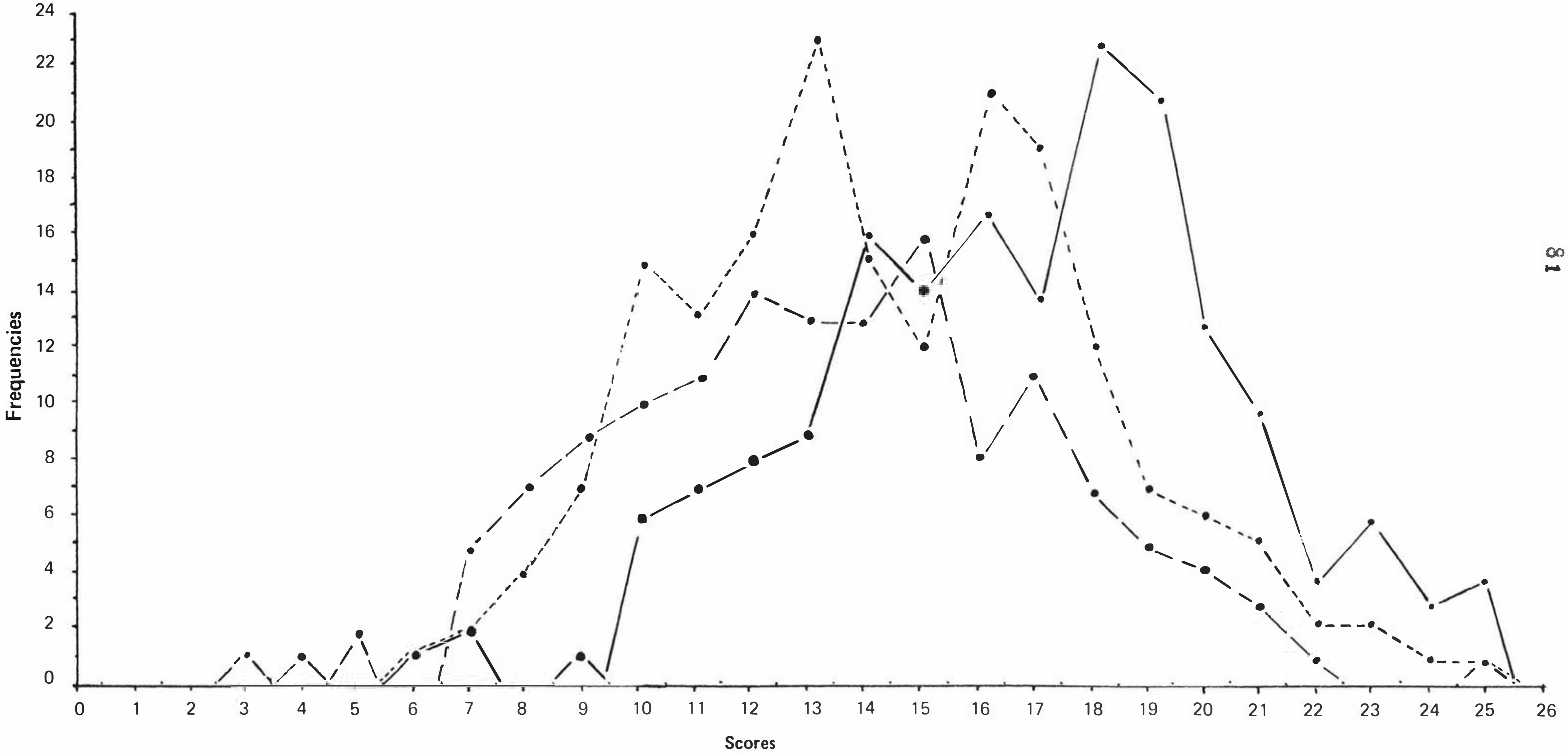


FIGURE 4 Distribution of Raw Scores on AH5 (Part 1) Test of Intelligence for Test Samples, England ——— (N = 180) USA – – – – (N = 139) NZ · · · · · · (N = 184)

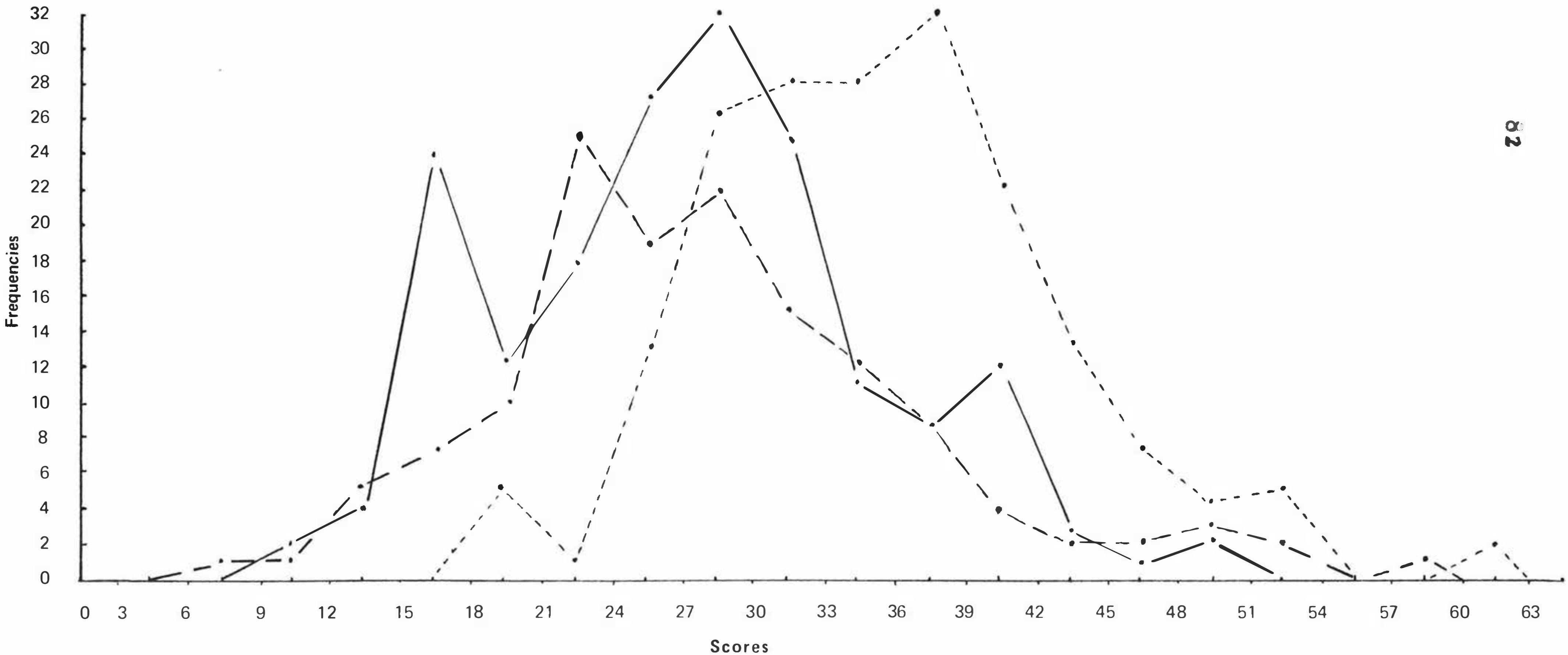


FIGURE 5 Frequency Polygon for the Distribution of Raw Scores, Test 1, Fluency. England ———— (N = 182) USA — — — — (N = 140) NZ . . . . (N = 184)

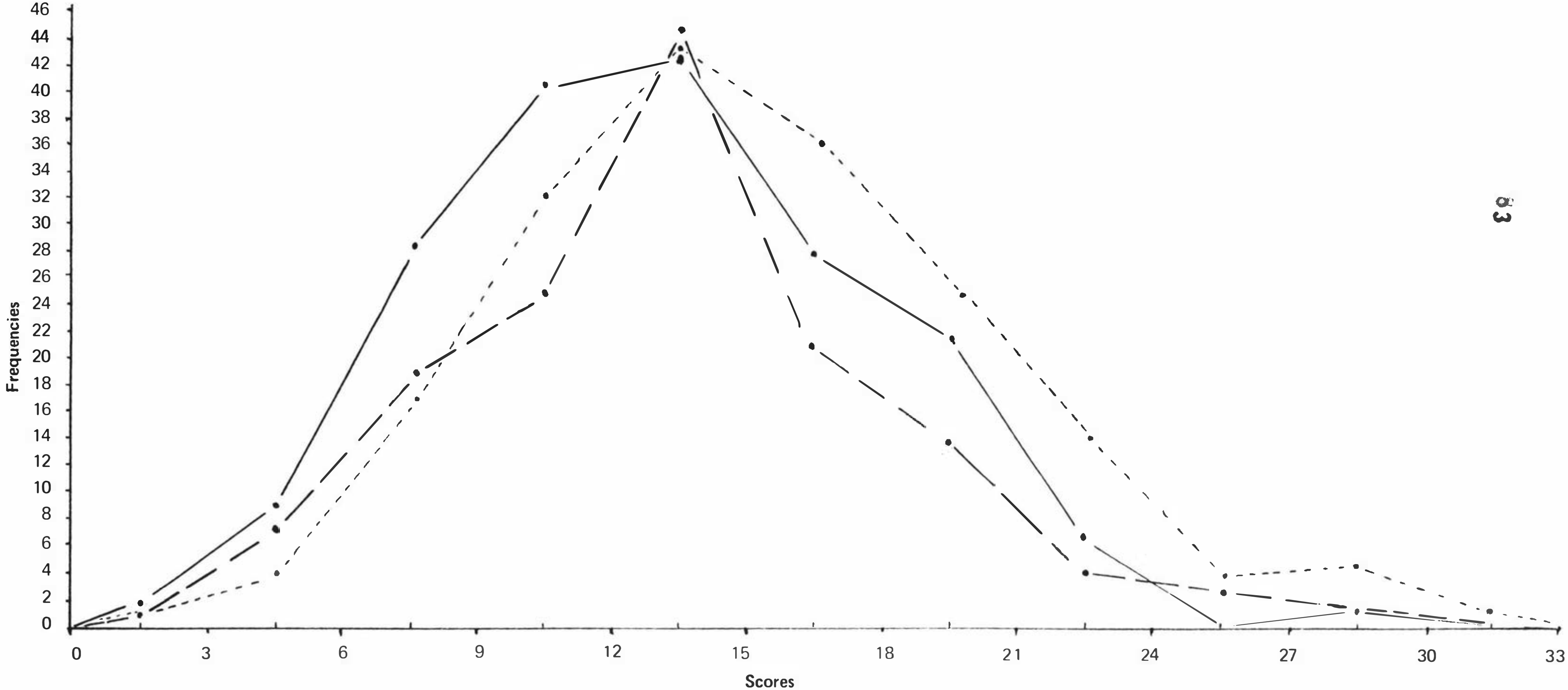


FIGURE 6 Frequency Polygon for the distribution of Raw Scores, Test 1, Flexibility. England ——— (N = 182) USA — — — — (N = 140) NZ . . . . (N = 184)

In the light of sample size, smoothing the curves yields a near normal distribution. It is interesting to note that the frequency polygons show graphical differences in characteristics between the samples. Although the English sample is generally superior on measured intelligence it should not be thought that the New Zealand or U.S.A. samples are 'only average'. In fact the means of both these countries are very high, matching the AH5 level equivalent to the top 25% of a population sample already selected in terms of superior intelligence.

The frequency polygon for Fluency Test 1 (Figure 5) shows differences in mean scores between countries, but a general tendency toward normal distribution can be seen in each case. Similarly for Flexibility Test 1 (Figure 6) approximation toward normal distribution is apparent, although the shapes of the curves for Flexibility Test 1 are different from those obtained for Fluency Test 1 and reflect different standard deviations. The curves of distributions for Test 2A, and 2B, on fluency and flexibility show slight tendencies toward being positively skewed (Appendix E, Figures 1-4).

Since differences in the nature of distributions occur both from test to test and country to country, all scores were converted to T-scores. In translating to T-scores not only are standardized scores employed, but distributions are normalized. This procedure is of fundamental importance for subsequent statistical analyses.

From the 15 sets of cognitive variables recorded as raw scores in Table 6 (Basic Statistics for Samples by Country) six further calculations determined:

(i) the means of Tests 2A and 2B (Fl 2M, Fx 2M)

(ii) the means of Tests 3A/3B, 3C/3D (Fl 3M, Fx 3M) and

(iii) the means of all test fluencies (Fl M) and test flexibilities (Fx M).

This resulted in 21 sets of scores for cognitive test variables which were converted to T-scores.+ Of these 21 sets of scores 20 were for open-ended test variables and one for the AH5. And of the 20 open-ended 18 were recorded for each individual, the other two being alternatives for Test 3.

Finally, all scores were recorded in three separate age groups for each country separately.

+Unless stated otherwise 'scores' in subsequent data presentations refer to T-scores.

(ii) (a) Within-country differences (t-tests) - influence of age on scores from cognitive variables

The influence of age on scores from cognitive variables was sought at an early stage of analysis. If results from cognitive test variables classified by the three age groupings within countries showed few significant difference, the consequent degree of homogeneity on the basis of age would suggest that the groups could be combined to form a total sample for each country. On the other hand if a fairly high incidence of significant differences occurred no such procedure could be adopted. Moreover, it would be necessary for tendencies of non-significance to be apparent in all of the three countries before such a step could be taken.

An analysis of the significance of differences between the means (t-tests) on 153 sets of scores (17 test variables by three age groups by three countries) was undertaken on a within-country basis. A condensed report using 81 sets of scores (nine test variables by three age groups by three countries) derived from the nine major test variables is reported in Table 7, page 87.

TABLE 7

THE EFFECT OF AGE ON SCORES FROM COGNITIVE VARIABLES : WITHIN-COUNTRY DIFFERENCES

An analysis of the significance of difference between means (t–tests) on 81 sets of scores for test variables
(Intelligence : 1; Fluency : 4; Flexibility ; 4) for three age groups and three countries - (England, U.S.A., New Zealand)

Age Groups within Countries	Cognitive Variables Showing Significance	Level of Significance P<	Group with Higher Mean
ENG I – ENG II	None	––	–––
ENG I – ENG III	None	––	–––
ENG II – ENG III	None	––	–––
	Σ=0 (out of 27 sets of significance)		
USA I – USA II	Fl 1	.05	USA I
	Fx 2M	.05	USA I
USA I – USA III	Fl 2M	.05	USA I
	Fx 2M	.01	USA I
	Fl M	.05	USA I
USA II – USA III	None	––	–––
	Σ=5 (out of 27)		
NZ I – NZ II	AH5	.05	NZ II
NZ I – NZ III	AH5	.05	NZ III
NZ II – NZ III	None	––	–––
	Σ=2 (out of 27)		
Total	Σ=7 (out of 81)		

KEY

Ages I = 13.4 – 13.11 yrs; II = 14.0 – 14.7 yrs; III = 14.8 – 15.3 yrs

Sets of Scores for Cognitive Variables

AH5 = Intelligence
Fl 1, Fl 2M, Fl 3M, Fl M = Fluency Test 1, Fluency Test 2 (Means), Fluency Test 3, (Means), Mean of all Fluencies
Fx 1, Fx 2M, Fx 3M, Fx M = Flexibility Test 1, Flexibility Test 2 (Means), Flexibility Test 3 (Means), Mean of all Flexibilities.

Significance of Difference

p ≤ .05 represents .02 to .05; p ≤ .01 represents .01 and beyond; – represents p > .05. This interpretation will be used throughout the treatment of results.

Example of Interpretation:

Age Groups within Countries	Cognitive Variables Showing Significance	Level of Significance	Group with Higher Mean
NZ I – NZ II	AH5	.05	NZ II

Interpretation: With sample NZ I age 13.4 – 13.11 yrs, against NZ II age 14.0 – 14.7 yrs, on AH5 (Intelligence) the null hypothesis can be rejected at the .05 level of confidence. From the significance of the difference of two means (NZ I against NZ II on AH5) the group with the higher mean is NZ II.

Relatively few significant differences appear within sets of scores on cognitive test variables in relation to age. 74 out of 81 sets of differences are not significant when nine sets of test variables are employed.+ 140 out of 153 sets of differences are not significant when 17 sets of test variables are employed.++

Differences within tests

Of the 72 sets of differences for the fluencies and flexibilities 67 are not significant. The five that are significant occur in the U.S.A. sample, the higher mean favouring the younger age group in each case.

Of the nine sets of differences for the AH5, seven are not significant. The two that are significant occur in the New Zealand sample, the higher mean favouring the older age group in each case.

Differences within countries

Of the 27 sets of differences for each country, 27 are not significant for England (49 are not

---

+81 = 9 tests x 3 countries x 3 ages. For each country this means 27 test results (3 AH5 scores, 12 fluencies, 12 flexibilities).

++153 = 17 tests x 3 countries x 3 ages. For each country this means 51 test results (3 AH5 scores, 24 fluencies, 24 flexibilities).

significant on the 51 sets of scores). For New Zealand, of the 27 sets of differences 25 are not significant (49 are not significant on the 51 sets of scores). For U.S.A., of the 27 sets of differences 22 are not significant (42 are not significant on the 51 sets of scores). Of the total 153 sets of differences between the means for all countries on all variables 140 are not significant. Thus age does not appear to be operating as a significant variable on the test scores within countries in the three nine-month age groupings. If it is indeed still a significant variable, its influence on all of the nine subsamples is being masked by other overriding factors (e.g. teaching style), or by compensatory sample selection factors. Both of these are unlikely, except that for AH5 scores a compensatory factor could be operating whereby younger but brighter pupils appear in the same classrooms as slightly older pupils. On a raw score basis (AH5) there is no significant difference between the means of the three age groups in England or U.S.A.

(ii) <u>(b) Between-country differences (t-tests)</u>

Between-country differences on cognitive variables present a somewhat different picture (Table 8, page 90).+

+For further details see Appendix P

TABLE 8

ANALYSIS OF BETWEEN–COUNTRY DIFFERENCES

Analysis of the significance of difference between means (t–tests) on 45 sets of scores for cognitive variables (Intelligence ; 1, Fluency ; 7, Flexibility ; 7) for three countries, England, U.S.A., and New Zealand *

Countries	Cognitive Variables Showing Significance	Level of Significance	Group with Higher Mean
ENG – USA	AH5	P< .01	ENG
ENG – NZ	AH5	.01	ENG
USA – NZ	AH5	.05	NZ
ENG – USA	Fl 1	—	——
ENG – NZ	Fl 1	.01	NZ
USA – NZ	Fl 1	.01	NZ
ENG – USA	Fx 1	—	——
ENG – NZ	Fx 1	.01	NZ
USA – NZ	Fx 1	.01	NZ
ENG – USA	Fl 2A	—	——
ENG – NZ	Fl 2A	.01	NZ
USA – NZ	Fl 2A	.01	NZ
ENG – USA	Fx 2A	—	——
ENG – NZ	Fx 2A	.01	NZ
USA – NZ	Fx 2A	.01	NZ
ENG – USA	Fl 2B	.05	ENG
ENG – NZ	Fl 2B	.01	NZ
USA – NZ	Fl 2B	.01	NZ
ENG – USA	Fx 2B	—	——
ENG – NZ	Fx 2B	.01	NZ
USA – NZ	Fx 2B	.01	NZ
ENG – USA	Fl 3A	.05	ENG
ENG – NZ	Fl 3A	—	——
USA – NZ	Fl 3A	.05	NZ
ENG – USA	Fx 3A	—	——
ENG – NZ	Fx 3A	—	——
USA – NZ	Fx 3A	—	——
ENG – USA	Fl 3B	.01	ENG
ENG – NZ	Fl 3B	—	——
USA – NZ	Fl 3B	.01	NZ
ENG – USA	Fx 3B	.01	ENG
ENG – NZ	Fx 3B	—	——
USA – NZ	Fx 3B	—	——
ENG – USA	Fl 3C	.01	ENG
ENG – NZ	Fl 3C	—	——
USA – NZ	Fl 3C	.01	NZ
ENG – USA	Fx 3C	.01	ENG
ENG – NZ	Fx 3C	—	——
USA – NZ	Fx 3C	.01	NZ
ENG – USA	Fl 3D	—	——
ENG – NZ	Fl 3D	—	——
USA – NZ	Fl 3D	—	——
ENG – USA	Fx 3D	—	——
ENG – NZ	Fx 3D	—	——
USA – NZ	Fx 3D	—	——

* KEY, see page 87

These results indicate that in approximately half of the 45 sets of means the difference is statistically significant. Quite a clear pattern emerges from these differences whereby the English sample records a higher mean score on AH5 than New Zealand or U.S.A., and whereby New Zealand scores higher than U.S.A. On fluency the New Zealand sample records higher mean scores than U.S.A. in six out of the seven tests, and on flexibility in four out of the seven tests (the remaining one and three sets of comparisons being not significant). The New Zealand sample records higher mean scores than the English sample in only two out of the seven tests of fluency, and in three out of the seven tests of flexibility (the remaining five and four sets of comparisons being not significant). The English sample records higher mean scores than the U.S.A. sample in four out of the seven tests of fluency, and in two out of the seven tests of flexibility (the remaining three and five sets of comparisons being not significant). Differing degrees of cross-cultural difference which will reappear in other analyses are therefore discernible in the sample. These differences again highlight the necessity for separate statistical analyses for each country in terms of correlational studies, and factor analysis.

(iii) Correlational Studies Employing Cognitive Variables

One of the major questions concerned the way in which the two dimensions of fluency and flexibility of thinking related to each other.[+]

Although several other studies (Getzels and Jackson, 1962; Guilford, 1959, 1959a, 1964) have investigated aspects of fluency and flexibility of thinking as aspects of divergent or creative thinking, none has looked at the specific relationship between these two dimensions of thinking. Instead, such studies appear to have submerged the relationship between fluency and flexibility in a single test score (sometimes incorporating originality) which is then correlated with other tests in the battery. The present analysis, by contrast, examines the relationship between fluency and flexibility as separate measures.

Information was also sought on the relationship between measured intelligence and (a) fluency, and (b) flexibility. Of course any relationship with intelligence would depend in part on the correlation

---

[+] It is understood that the relationship is within the present framework of scoring criteria, science test stimuli, and sample. Although the sample represents a restricted range on the basis of ability in science the range becomes less restricted in terms of fluency and flexibility. In addition all scores have been normalized (See McNemar, 1964, p.879).

between fluency and flexibility. If the correlation between fluency and flexibility were low it could then be suggested that intelligence may correlate more highly with flexibility than fluency, since it requires the ability to encode and recall over many categories, a skill which could be claimed to require higher level intellectual functioning than that required for narrow categorization and recall. On the other hand sheer ideational fluency with a tolerable degree of flexibility (on the basis of present scoring criteria) may represent a more powerful intellectual predictor.

Evidence which assists to answer some of these questions appears in the following tables of correlation coefficients.

Correlation coefficients for

(a) fluency across tests are presented in Table 9,

(b) flexibility across tests are presented in Table 10,

(c) fluency against flexibility within tests are presented in Table 11, and

(d) measured intelligence (AH5) against (a) fluency and (b) flexibility are presented in Table 12.

TABLE 9

CORRELATION COEFFICIENTS FOR FLUENCY ACROSS TESTS FOR ENGLAND, U.S.A., NEW ZEALAND

Fluency	Test 1	Test 2A	Test 2B	Test 3A/B	Test 3C/D
Fl 1	.	ENG .41 USA .51 NZ .37	.34 .41 .28	.37 .41 .36	.32 .26 .31
Fl 2A		.	.61 .60 .59	.39 .49 .47	.45 .37 .36
Fl 2B			.	.39 .30 .46	.40 .23 .36
Fl 3A/B				.	.56 .65 .61
Fl 3C/D					.

TABLE 10

CORRELATION COEFFICIENTS FOR FLEXIBILITY ACROSS TESTS FOR ENGLAND, U.S.A, NEW ZEALAND

Flexibility	Test 1	Test 2A	Test 2B	Test 3A/B	Test 3C/D
Fx 1	.	ENG .38 USA .35 NZ .32	.20 .10 .15	.24 .42 .17	.34 .22 .15
Fx 2A		.	.58 .55 .52	.35 .49 .41	.40 .18 .32
Fx 2B			.	.39 .33 .36	.29 .16 .26
Fx 3A/B				.	.50 .49 .63
Fx 3C/D					.

TABLE 11

CORRELATION COEFFICIENTS FOR FLUENCY AGAINST FLEXIBILITY WITHIN TESTS FOR ENGLAND U.S.A., NEW ZEALAND

Flexibility \ Fluency	Fl 1	Fl 2A	Fl 2B	Fl 3A/B	Fl 3C/D
Fx 1	ENG .66 USA .69 NZ .51				
Fx 2A		.67 .68 .65			
Fx 2B			.43 .63 .59		
Fx 3A/B				.77 .85 .79	
Fx 3C/D					.77 .67 .71

For sample size see Page 79.

TABLE 12

<u>Correlation Coefficients for Intelligence Against Fluency and Flexibility on Tests for England, U.S.A. New Zealand</u>

Intelligence \ Fl & Fx		Fl 1	Fx 1	Fl 2 (Mean)	Fx 2 (Mean)	Fl 3 (Mean)	Fx 3 (Mean)	Fl M (Overall	Fx M Means)
	ENG	.31	.18	.26	.10	.14	.10	.30	.16
AH5	USA	.24	.16	.19	.19	.12	.16	.24	.22
(Part 1)	NZ	.04	-.02	.00	-.02	.00	.06	.01	.01

For sample size see Page 79.
The level of significance for coefficients in Tables 9 - 12 is, in nearly all cases, between .01 and .05.*

A fairly consistent pattern of results across tests and samples by country can be seen from the correlation coefficients in these tables. Some differences between tests and samples are apparent, and these will be elaborated, but the general picture is one of variation on a theme.

*Employing tables for the values of the correlation coefficient for different levels of significance (Fisher, 1925, <u>Statistical Methods for Research Workers</u> in Popham, (1967)) the probabilities for df. 100 (which is the sample size in most of the present cases) are:
.01 for r .254; .02 for r .230; .05 for r .195.

Fluency across tests

The coefficients revealed a generally low to moderate positive correlation. On the whole, Tests 2 against Tests 3 indicated slightly higher coefficients. The highest coefficients occurred between Tests 2A/B (.61, .60, .59) and Tests 3A-B/C-D (.56, .65, .61) which are equivalent-part correlations or measures of internal consistency.

When viewed cross-culturally the samples showed few differences in correlational behaviour across tests:
Test 1 against Test 2 (mean correlations employing Fisher's z coefficient for averaging coefficients[+])
England .38, U.S.A. .46, N.Z. .33;
Test 1 against Test 3, England .35, U.S.A. .34, N.Z. .34;
Test 2 against Test 3, England .43, U.S.A. .35, N.Z. .42.

Flexibility across tests

As with fluency, generally low positive correlations across tests were observed for flexibility. However, they are slightly lower than for fluency and indicate a wider range. The highest coefficients occurred between tests of equivalent parts 2A/B (.58, .55, .52) and 3A-B/C-D (.50, .49, .63). When viewed cross-culturally the mean correlations (Fisher's z) showed greater

---

[+]Guilford (1950), p.355, 616.

variability than for fluency:

Test 1 against Test 2, England .29, U.S.A. .23, N.Z. .24;

Test 1 against Test 3, England .29, U.S.A. .32, N.Z. .16;

Test 2 against Test 3, England .36, U.S.A. .29, N.Z. .34.

Fluency against flexibility within tests

Moderate to high positive correlation coefficients occurred for fluency against flexibility. Moreover, the pattern was quite consistent across tests and countries. The highest correlations were found with Test 3A-B (.77, .85, .79). Mean coefficients for fluency against flexibility were:

Test 1, England .66, U.S.A. .69, N.Z. .51;

Test 2, England .56, U.S.A. .66, N.Z. .62;

Test 3, England .77, U.S.A. .78, N.Z. .75.

The correlations for the mean of all fluencies against the mean of all flexibilities for each country were: England .79, U.S.A. .76, N.Z. .72.
The question of how much influence intelligence has over the fluency/flexibility relationship should be considered. As a result a first order partial correlation holding AH5 constant was calculated for the mean of all fluencies against the mean of all flexibilities for each country, yielding the following results: England .78, U.S.A. .75, N.Z. .72. These figures indicate that a moderately high correlation

between the two variables of fluency and flexibility was sustained when the influence of the third variable (AH5 intelligence) was partialled out.

Intelligence against (a) fluency and (b) flexibility

A generally low to slight positive correlation occurred between intelligence and (a) fluency and (b) flexibility for the population and content-sample of the present study.[+] Again, there was a fairly high degree of consistency in the pattern of correlations. But some interesting cross-cultural differences were apparent, e.g. a correlation of .30 was obtained between intelligence and the overall mean of all fluencies (Fl M) for the English sample, while for New Zealand it was virtually zero. For the U.S.A. sample it was .24.

Differences between the two dimensions of fluency and flexibility in their relationship with intelligence existed in a few correlations e.g. Test 1: Fluency .31, Flexibility .18, Fl M .30, Fx M .16, (England). Over the sample as a whole, however, differences between fluency and flexibility in their correlations with intelligence were not great. But where differences

---

[+]While it is possible to correct for attenuation and claim greater generality over a wider sample it is not the concern of this study to do so.

did occur, fluency indicated a higher correlation with intelligence than flexibility.

Further analyses of these correlations will be made later by way of factor analysis (p.102).

(iv) High and Low Scorer Comparisons:
Extreme scorers at the $\pm$1SD level for fluency and flexibility on the one or the other or on both dimensions were examined in each sample by country for any significant changes and trends in correlational behaviour. It could well be suggested that different relationships between intelligence and fluency/flexibility, or between fluency and flexibility, might be apparent with extreme scorers. The general tendency, however, was one of reduced correlation (attributable to restricted range) and of increased variability. In the light of these two points and since no clear trends were discernible it was decided not to correct the correlations for attenuation. A more fruitful approach to the issue of extreme scorers may be one of individual analysis whereby, for example, pupils with exceptionally high or low fluency-flexibility ratios are examined further in relation to other cognitive variables. It might well be, however, that personality traits rather than cognitive variables are of more value in illustrating significant

relationships. Some insight into this question was obtained from later analyses of extreme scorers and from the results of the Attitude Scale (Test 4).

<u>Tests of significant differences with high and low scorers: AH5 against fluency and flexibility</u>

Using the same groups of $\pm$1SD extreme scorers, another approach was made to examine in a more precise way the relationship between measured intelligence and fluency and flexibility of thinking.

A modification of the usual plan of employing high and low scorers alone was suggested by MacKinnon (1967) whereby a control group of mean scorers might be used to further illustrate the nature and extent of the difference. Thus a three-way t-test programme was implemented - (i) high against low (ii) high against mean, and (iii) low against mean. Results appear in Table 13.

TABLE 13

Tests of Significant Differences Between the Means of Hi-Lo Scorers for Fluency and Flexibility Against Intelligence (AH5), England, U.S.A., New Zealand

Country	Fl Groups on AH5	T ratio	d.f.	p<	Group With Higher Mean on AH5
ENG	+1SD/-1SD	3.1972	68	.01	+1SD
	+1SD/Mean	2.8298	145	.01	+1SD
	-1SD/Mean	-1.0567	141	---	---
USA	+1SD/-1SD	3.4334	35	.01	+1SD
	+1SD/Mean	2.7271	118	.01	+1SD
	-1SD/Mean	-2.1726	119	.05	Mean
NZ	+1SD/-1SD	-.4175	50	---	----
	+1SD/Mean	-.7853	160	---	----
	-1SD/Mean	-.1964	152	---	----
	Fx Groups on AH5				
ENG	+1SD/-1SD	2.1475	57	.05	+1SD
	+1SD/Mean	1.5830	153	---	----
	-1SD/Mean	-1.1542	144	---	----
USA	+1SD/-1SD	2.2759	36	.05	+1SD
	+1SD/Mean	1.6639	117	---	----
	-1SD/Mean	-1.3922	119	---	----
NZ	+1SD/-1SD	-.1959	56	---	----
	+1SD/Mean	-1.4327	155	---	----
	-1SD/Mean	-1.1453	151	---	----

With fluency, the +1SD group showed a superiority on measured intelligence which was significant in every case in both the English and U.S.A. samples. The trend was also supported in terms of -1SD and mean differences. With flexibility, the superiority for the +1SD group was again significant for the English and U.S.A. samples, but was less clear for -1SD and mean groups. It would seem that fluency is again more powerful than flexibility in indicating a relationship with measured intelligence. For New Zealand, however, no significant differences occurred between extreme scorers in fluency or flexibility and intelligence.

Significant differences with high and low scorers: Age

Using the same groups of ±1SD extreme scorers for fluency and flexibility, the question of whether high (or low) scorers were significantly younger (or older) was asked. Using ±1SD on the mean of all fluencies and on the mean of all flexibilities for each country in turn no significant differences in age were indicated.

(v) Factor Analysis - cognitive variables

Factor analysis was seen as an efficient method of further describing the internal structure of matrices of covariances and correlations relating to intelligence and fluency and flexibility of thinking.

Since relatively few cognitive factors

account for the covariance of test variables in the present study the centroid method was selected as being most appropriate (Arvidson, 1970; Fruchter, 1954; Lawley and Maxwell, 1963; Thomson, 1956). Orthogonal rotations were undertaken, and Tucker's phi used as a guide to the termination of factors. These procedures resulted in a relatively fast convergence to three factors as the best fit for explaining the covariances in the present 7 x 7 matrix. A separate factor analysis was undertaken for each set of correlations for each country, the results of which appear in Tables 14, 15 and 16.

The rotation to three factors yielded an efficient and parsimonious solution for all three countries. The three factors for each country appear below:

<u>ENG</u>	Factor I	Two high loadings Fl 3 (.849), Fx 3 (.850)
	Factor II	Two high loadings Fl 1 (.778), Fx 1 (.709)
	Factor III	One high one moderate Fl 2 (.712), Fx 2 (-.564)
<u>USA</u>	Factor I	Two high loadings Fl 3 (.827), Fx 3 (.870)
	Factor II	Two high loadings Fl 1 (-.786), Fx 1 (-.750)
	Factor III	Two high loadings Fl 2 (.841), Fx 2 (.776)
<u>NZ</u>	Factor I	Two high loadings Fl 2 (.784), Fx 2 (.782)
	Factor II	Two high loadings Fl 1 (-.717), Fx 1 (-.655)
	Factor III	Two high loadings Fl 3 (-.802), Fx 3 (-.849)

The solution in each of the three cases bears a similar pattern although the identity of the first and third factors is reversed for New Zealand.

TABLE 14

FACTOR ANALYSIS RESULTS FOR INTELLIGENCE (1), FLUENCY (3), & FLEXIBILITY (3), - ENGLAND

Unrotated

VAR	H2	F1	F2	F3
AH5	.180	.309	.220	–.188
Fl 1	.708	.709	.447	.074
Fl 2	.644	.690	–.132	–.387
Fl 3	.841	.779	–.419	.243
Fx 1	.588	.631	.399	.172
Fx 2	.530	.663	–.226	–.196
Fx 3	.800	.744	–.326	.374

Three Factors Rotated

VAR	H2	F1	F2	F3
AH5	.179	–.033	.327	–.266
Fl 1	.707	.238	.778	–.213
Fl 2	.643	.284	.233	–.712
Fl 3	.841	.849	.146	–.312
Fx 1	.586	.270	.709	–.103
Fx 2	.529	.426	.170	–.564
Fx 3	.799	.850	.220	–.165

Lower Triang Corr. Matrix

AH5	AH5	Fl 1	Fl 2	Fl 3	Fx 1	Fx 2	Fx 3
Fl 1	.307						
Fl 2	.263	.410					
Fl 3	.140	.389	.537				
Fx 1	.179	.661	.331	.315			
Fx 2	.102	.386	.594	.505	.352		
Fx 3	.104	.397	.396	.821	.360	.472	

TABLE 15

FACTOR ANALYSIS RESULTS FOR INTELLIGENCE (1), FLUENCY (3), & FLEXIBILITY (3), - U.S.A.

Unrotated

VAR	H2	F1	F2	F3
AH5	.111	.288	–.130	.106
Fl 1	.728	.713	–.461	.085
Fl 2	.792	.717	.218	.479
Fl 3	.781	.719	.349	–.377
Fx 1	.632	.613	–.475	–.176
Fx 2	.693	.685	.257	.395
Fx 3	.839	.731	.269	–.482

Three Factors Rotated

VAR	H2	F1	F2	F3
AH5	.111	.034	–.269	.193
Fl 1	.728	.143	–.786	.298
Fl 2	.791	.183	–.221	.841
Fl 3	.780	.827	–.153	.267
Fx 1	.632	.258	–.750	.045
Fx 2	.691	.239	–.174	.776
Fx 3	.839	.870	–.230	.168

Lower Triang Corr. Matrix

AH5	AH5	Fl 1	Fl 2	Fl 3	Fx 1	Fx 2	Fx 3
Fl 1	.240						
Fl 2	.185	.510					
Fl 3	.118	.379	.429				
Fx 1	.160	.687	.239	.282			
Fx 2	.187	.361	.742	.377	.270		
Fx 3	.157	.323	.359	.814	.414	.386	

TABLE 16

FACTOR ANALYSIS RESULTS FOR INTELLIGENCE (1), FLUENCY (3), & FLEXIBILITY (3), - N.Z.

Unrotated

VAR	H2	F1	F2	F3
AH5	.014	.032	–.040	–.110
Fl 1	.584	.607	–.462	.048
Fl 2	.713	.710	.255	.379
Fl 3	.787	.760	.314	–.331
Fx 1	.443	.436	–.491	.106
Fx 2	.724	.717	.185	.418
Fx 3	.793	.728	.232	–.456

Three Factors Rotated

VAR	H2	F1	F2	F3
AH5	.014	–.073	–.039	–.087
Fl 1	.584	.175	–.717	–.196
Fl 2	.712	.784	–.205	–.236
Fl 3	.785	.359	–.111	–.802
Fx 1	.442	.102	–.655	–.040
Fx 2	.723	.782	–.271	–.190
Fx 3	.791	.219	–.150	–.849

Lower Triang Corr. Matrix

AH5	AH5	Fl 1	Fl 2	Fl 3	Fx 1	Fx 2	Fx 3
Fl 1	.042						
Fl 2	.004	.364					
Fl 3	.004	.362	.509				
Fx 1	–.024	.509	.141	.153			
Fx 2	–.025	.337	.723	.421	.279		
Fx 3	.065	.301	.372	.794	.177	.407	

It is clear that no overall intelligence or fluency or flexibility factor can be postulated. This in itself is a significant finding and confirms in a more striking way the lack of any pervasive fluency or flexibility trait across tests mentioned earlier.

The consistently high loadings on fluency-flexibility suggests some kind of composite factor. But further examination shows that the fluency-flexibility composite is anchored to specific tests. Further illumination of this can be seen by determining the percentage of variance contributed by the major fluency-flexibility test-composite for each factor for each country (Table 17).

TABLE 17

Percentage of Variance Contributed by Major Fluency-Flexibility Test-Composite for Each Factor for Each Country

Country	Factor	High Loadings	% Variance	Rank
ENG	I	Fl 3 - Fx 3	38.30	(i)
	II	Fl 1 - Fx 1	30.02	(ii)
	III	Fl 2 - Fx 2	27.35	(iii)
			Σ 95.67	
USA	I	Fl 3 - Fx 3	35.41	(i)
	II	Fl 1 - Fx 1	30.00	(iii)
	III	Fl 2 - Fx 2	32.63	(ii)
			Σ 98.04	
NZ	I	Fl 2 - Fx 2	35.42	(ii)
	II	Fl 1 - Fx 1	25.32	(iii)
	III	Fl 3 - Fx 3	38.90	(i)
			Σ 99.64	

From this procedure it is evident that the fluency-flexibility test-composite contributing most of the variance is the same in each country, i.e. Test 3, (Factor I England, Factor I U.S.A., Factor III New Zealand). Thereafter differences occur by country in the ranking of importance of the contribution of test-composite factors toward the percentage of variance, but these are slight in most cases. What <u>is</u> significant is the same pattern of the fluency-flexibility test-composite factor across the three countries, and the common priority of Test 3 in accounting for the highest percentage of variance.

The question of interpreting these factors in the light of the stimulus material now remains. First of all no general cognitive factor, either as general intelligence (AH5) or as fluency or flexibility of thinking, can be constructed from the results. Secondly, while there is consistent evidence for some kind of composite fluency-flexibility factor, it is so convincingly test-oriented that it would seem necessary to refocus attention on the specific nature of each test. It is not simply a case of some abstract cognitive-style complex (fluency-flexibility) but of a somewhat specific cognitive-style factor that emphasizes set and task. What originally might have been thought to have been a fairly general stimulus-response ground within the domain of science has turned out to be otherwise. Thus scientific vocabulary (Test 1), definition of scientific concepts (Test 2) and uses of chemical substances and physical objects (Test 3) elicit very different competencies despite the overriding request to think as fluently and flexibly as possible on all occasions. And this is so with the three culture-samples. Interpreted another way, what originally might have been thought to have been a fairly stable cognitive style trait of fluency (or flexibility) operating across tests has turned out to be somewhat less pervasive and more task-oriented.

It would thus seem that the nature and demand of each task is powerful in determining the three factors. Further analysis suggests that while the fluency-flexibility composite may be partly cognitive, it may also be partly attitudinal - premised on initial attitudes of like/dislike, interest/uninterest toward each set of test stimuli. Both cognitive and attitudinal aspects appear to be interacting and be anchored in past experiences associated with each test task. Moreover, this explanation seems to hold across the three culture samples. It also seems to carry with it a warning about referring to fluency or flexibility (or divergent thinking or creativity) as if it existed outside a particular field of endeavour such as science or architecture or even outside a specific sphere within the discipline.

B. PERSONALITY

(i) Factor analysis of Attitude Scale, Test 4, Part I

The 36 item attitude scale which comprised Part I of Test 4 was subjected to a factor analysis employing a centroid approach with orthogonally rotated axes. An iterative process incorporating Tucker's phi resulted in a fairly rapid convergence to four factors.

Since aspects of self-perception relating to the Attitude Scale (Test 4) are likely to be sensitive to cultural differences, separate factor analyses for each sample by country were necessary. In the present situation the nature and order of factors could well be very different for each country. The results of the rotated solution to four factors for each of the three countries are presented in Tables 18, 19, and 20.+

+Details of unrotated and rotated solutions for the 36 x 36 matrix obtainable from computer print-out.

TABLE 18

FACTOR ANALYSIS RESULTS FOR ATTITUDE SCALE : ENGLAND

VAR	H2	F1	F2	F3	F4
1	.389	.576	.023	.215	.103
2	.337	.470	.133	–.194	.244
3	.193	.120	.100	.406	–.065
4	.214	.417	.164	.109	–.034
5	.065	–.012	.108	–.083	.215
6	.055	–.014	.079	–.019	.221
7	.240	.117	.457	–.002	.132
8	.048	–.141	.158	–.055	.017
9	.243	–.270	.134	–.194	–.339
10	.347	.306	.502	.021	.021
11	.344	.056	.516	–.231	.144
12	.431	.595	.063	–.197	.184
13	.324	–.526	.044	–.055	–.207
14	.104	.099	.260	–.080	–.143
15	.474	.651	.174	.140	–.007
16	.270	–.406	.274	–.025	–.172
17	.059	–.077	.004	–.120	–.197
18	.286	.473	.225	–.106	–.019
19	.307	.026	.026	–.535	–.141
20	.317	.044	.197	–.511	.124
21	.235	.219	.183	–.345	.186
22	.365	–.425	.201	–.207	–.317
23	.291	–.526	–.025	.091	.079
24	.202	.250	.352	–.029	.118
25	.354	–.324	.475	–.058	.142
26	.293	.536	.073	–.027	–.005
27	.136	–.204	.192	–.117	–.208
28	.121	–.202	.039	–.190	–.205
29	.290	.432	.115	–.295	.058
30	.345	.554	.160	.079	.078
31	.192	–.080	–.147	.185	–.360
32	.100	–.081	.075	–.190	–.228
33	.402	.475	.388	.095	.126
34	.064	.028	–.009	.018	–.251
35	.081	.132	.063	–.036	.241
36	.236	.138	.461	.001	.063

TABLE 19

FACTOR ANALYSIS RESULTS FOR ATTITUDE SCALE : U.S.A.

VAR	H2	F1	F2	F3	F4
1	.440	.502	–.124	.407	.080
2	.353	.173	–.017	.306	.479
3	.098	–.020	.054	–.056	–.303
4	.259	.126	.080	.196	.444
5	.148	.009	–.379	–.060	–.003
6	.208	.334	–.089	.295	.037
7	.271	.076	–.370	.311	.176
8	.122	–.272	–.153	.130	.090
9	.243	–.320	–.290	–.234	.028
10	.520	.279	–.134	.634	.146
11	.382	–.193	–.212	.533	.124
12	.471	.207	.040	.260	.599
13	.401	–.506	.149	–.221	–.271
14	.355	.220	–.453	.317	–.013
15	.492	.444	–.260	.465	.105
16	.146	–.192	–.102	–.310	.054
17	.123	–.287	–.056	–.021	–.193
18	.384	.295	–.275	.433	.183
19	.335	–.017	–.534	.161	.151
20	.308	.011	–.479	.153	.232
21	.234	–.091	–.187	.004	.436
22	.269	–.249	–.409	–.181	.083
23	.205	–.389	.119	–.138	–.141
24	.376	.013	–.059	.520	.318
25	.298	–.394	–.255	.250	.120
26	.267	.491	–.013	.137	.086
27	.192	–.337	–.099	–.228	.129
28	.236	–.071	–.128	–.447	–.119
29	.203	.263	–.263	–.034	.251
30	.460	.449	–.280	.413	.093
31	.390	–.502	–.344	.024	–.136
32	.163	–.014	–.245	–.069	–.313
33	.302	.142	.007	.498	.182
34	.118	–.028	–.288	–.122	–.140
35	.246	.108	.127	.457	.095
36	.392	.033	–.022	.623	–.038

TABLE 20

FACTOR ANALYSIS RESULTS FOR ATTITUDE SCALE : NEW ZEALAND

VAR	H2	F1	F2	F3	F4
1	.397	.256	.317	–.040	.478
2	.245	–.023	.230	.067	.433
3	.041	.047	.130	–.118	.089
4	.201	.106	.152	.138	.384
5	.153	–.376	–.068	–.068	–.047
6	.109	.034	.316	.073	–.054
7	.078	–.179	–.021	–.013	.212
8	.123	.083	–.208	–.268	.025
9	.293	–.467	–.207	–.085	–.157
10	.151	–.011	–.080	.142	.352
11	.074	–.107	.052	–.042	.240
12	.359	–.039	.094	.113	.579
13	.361	–.236	–.258	–.458	–.167
14	.158	–.202	.159	.230	.198
15	.507	.267	.100	.245	.604
16	.463	–.113	.043	–.626	–.236
17	.065	.118	–.033	–.181	–.130
18	.459	.353	.039	.205	.539
19	.136	–.296	.124	.147	.106
20	.287	–.482	.160	.025	.168
21	.124	–.086	.316	–.003	.132
22	.338	–.542	–.144	–.072	–.136
23	.240	–.252	–.252	–.335	.013
24	.296	–.056	.503	.186	.072
25	.227	–.341	.066	–.291	.146
26	.209	.221	–.003	.199	.348
27	.158	–.375	.113	.055	–.044
28	.170	–.069	–.138	–.351	–.151
29	.085	.007	.036	.261	.125
30	.340	.328	.233	.161	.390
31	.265	–.412	–.125	–.271	–.072
32	.164	–.301	.198	.172	.066
33	.381	.132	.478	.206	.305
34	.342	–.002	.041	–.581	–.054
35	.090	–.072	.048	.036	.286
36	.325	–.063	.447	–.100	.334

Factor loadings above .3 were then extracted for each factor for each country and keyed to the appropriate item in the original questionnaire. These results are presented in Table 21.

These results highlight items that constitute an attitudinal factor as well as reveal cross-cultural similarities and differences in self-perception. Some similarities exist on Factor I for the three samples by country:

Item 25 (making jokes), ENG -324, USA -394, NZ -341
Item 30 (being very polite), ENG 554, USA 449, NZ 328

On the whole, the degree of correspondence in Factor I is greater between England and U.S.A. After these initial similarities in Factor I, many differences between the countries in pupil self-perception are apparent for Factors II, III and IV. This does not mean lack of meaningful agreement within a factor, since all significant items are related by definition of the factor analysis solution; e.g. Factor IV; item 4 (sticking to the truth) USA 444 relates with item 21 (getting everything correct) USA 436. Indeed the factor analysis of the Attitude Scale not only contributed towards a degree of parsimony and meaningfulness in the grouping of items but also contributed toward the validity of the instrument as a whole. Further examination of the factors reveal a

TABLE 21

Items with Factor Loadings Above .3 Grouped Under the Appropriate Four Factors for Test 4, Attitude Scale (ENG, USA, NZ)

Item No	Item	Loadings Factor I			Loadings Factor II			Loadings Factor III			Loadings Factor IV		
		ENG	USA	NZ	ENG	USA	NZ	ENG	USA	NZ	ENG	USA	NZ
1.	being obedient	576	502				317		407				478
2.	working hard	470							306			479	433
3.	having a low opinion of yourself							406				-303	
4.	sticking to the truth	417										444	384
5.	being highly imaginative			-376		-379							
6.	being cautious		334				316						
7.	being enthusiastic				457	-370			311				
8.	having set opinions												
9.	running risks		-302	-467							-339		
10.	having self control	306			502				634				352
11.	mixing well socially				516				533				
12.	doing homework thoroughly	595										599	579
13.	getting into trouble	-526	-506							-458			
14.	doing new things					-453			317				
15.	showing respect for teachers	651	444						465				604
16.	teasing people	-406							-310	-626			
17.	shrugging off criticism												
18.	being neat and tidy	473		353					433				539
19.	investigating the unusual					-534		-535					
20.	taking the lead			-482		-479		-511					
21.	getting everything correct						316	-345				436	
22.	doing a dangerous experiment	-425		-542		-409					-317		
23.	smoking	-526	-389							-335			
24.	accepting expert advice				352		503		520			318	
25.	making jokes	-324	-394	-341	475								
26.	keeping quiet in class	536	491										348
27.	arguing with people		-337	-375									
28.	showing off								-447	-351			
29.	taking things seriously	432											
30.	being very polite	554	449	328					413				390
31.	day dreaming		-502	-412		-344					-360		
32.	questioning the truth of a text book			-301								-313	
33.	doing one's duty	475			388		478		498				305
34.	exaggerating									-581			
35.	not hurting other people's feelings								457				
36.	being a good team member				461		447		623				334

that Factor I is bipolar for all three countries, but again this is consistently meaningful; e.g. item 30 (being very polite) NZ +328 can be seen against item 32 (questioning the truth of a text book) NZ -301. Thereafter, Factor III (ENG and USA) and Factor IV (USA) are bipolar. The remaining factor solutions contain either all negative or all positive loadings. Attention to sign values is therefore important in attributing psychological meaning to each factor.

Items with the highest loadings were then extracted and ranked (highest to lowest) for each factor for each country. These were related back to their original phenomenological grouping under the six original scales Tables 22, 23, and 24.

TABLE 22

FACTOR ANALYSIS AND PHENOMENOLOGICAL APPROACHES ON ATTITUDE SCALES : ENGLAND

FACTOR I

Rank	Loading	Item	Item No.	Original Scale (Phenomenological)
1	651	showing respect for teachers	15	Respect for authority
2	595	doing homework thoroughly	12	Serious minded worker
3	576	being obedient	1	Respect for authority
4	554	being very polite	30	Respect for authority, High conformity
5	536	keeping quiet in class	26	Respect for authority
6=	–526	getting into trouble	13	Respect for authority (negative), High risk
6=	–526	smoking	23	High risk
8	475	doing one's duty	33	Respect for authority, High conformity
9	473	being neat and tidy	18	Respect for authority, High conformity
10	470	working hard	2	Serious minded worker
11	432	taking things seriously	29	Serious minded worker
12	–425	doing a dangerous experiment	22	High risk
13	417	sticking to the truth	4	Respect for authority
14	–406	teasing people	16	High risk
15	–324	making jokes	25	Freedom of emotional expression High risk
16	306	having self control	10	Freedom of emotional expression (negative item originally)

FACTOR II

Rank	Loading	Item	Item No.	Original Scale (Phenomenological)
1	516	mixing well socially	11	High self concept
2	502	having self control	10	Freedom of emotional expression (negative originally)
3	475	making jokes	25	High risk, Freedom of emotional Expression
4	461	being a good team member	36	High conformity
5	457	being enthusiastic	7	Freedom of emotional expression
6	388	doing one's duty	33	High conformity, Respect for authority
7	352	accepting expert advice	24	High conformity

FACTOR III

Rank	Loading	Item	Item No.	Original Scale (Phenomenological)
1	–535	investigating the unusual	19	High risk, Freedom of emotional Expression
2	–511	taking the lead	20	High risk, High self concept
3	406	having a low opinion of yourself	3	High self concept (negative originally)
4	–345	getting everything correct	21	Serious minded worker

FACTOR IV

Rank	Loading	Item	Item No.	Original Scale (Phenomenological)
1	–360	day dreaming	31	High risk, Freedom of emotional Expression
2	–339	running risks	9	High risk
3	–317	doing a dangerous experiment	22	High risk

TABLE 23

FACTOR ANALYSIS AND PHENOMENOLOGICAL APPROACHES ON ATTITUDE SCALES : U.S.A.

FACTOR I

Rank	Loading	Item	Item No.	Original Scale (Phenomenological)
1	–506	getting into trouble	13	Respect for authority, High risk
2=	502	being obedient	1	Respect for authority
2=	–502	day dreaming	31	High risk, Freedom of emotional expression
4	491	keeping quiet in class	26	Respect for authority
5	449	being very polite	30	Respect for authority, High conformity
6	444	showing respect for teachers	15	Respect for authority
7	–394	making jokes	25	Freedom of emotional expression, High risk
8	–389	smoking	23	High risk
9	–337	arguing with people	27	High risk
10	334	being cautious	6	High risk (negative originally)
11	–302	running risks	9	High risk

FACTOR II

Rank	Loading	Item	Item No.	Original Scale (Phenomenological)
1	–534	investigating the unusual	19	High risk, Freedom of emotional expression
2	–479	taking the lead	20	High risk, High self concept
3	–453	doing new things	14	High risk, High conformity (Negative originally)
4	–409	doing a dangerous experiment	22	High risk
5	–379	being highly imaginative	5	High risk
6	–370	being enthusiastic	7	Freedom of emotional expression
7	–344	day dreaming	31	High risk, Freedom of emotional expression.

FACTOR III

Rank	Loading	Item	Item No.	Original Scale (Phenomenological)
1	634	having self control	10	Freedom of emotional expression
2	623	being a good team member	36	High conformity
3	533	mixing well socially	11	High self concept
4	520	accepting expert advice	24	High conformity
5	498	doing one's duty	33	Respect for authority, High conformity
6	465	showing respect for teachers	15	Respect for authority
7	457	not hurting other people's feelings	35	Freedom of emotional expression
8	–447	showing off	28	High risk
9	433	being neat and tidy	18	High risk, Respect for authority
10	413	being very polite	30	High conformity, Respect for authority
11	407	being obedient	1	Respect for authority
12	317	doing new things	14	High conformity
13	311	being enthusiastic	7	Freedom of emotional expression
14	–310	teasing people	16	High risk
15	306	working hard	2	Serious minded worker

FACTOR IV

Rank	Loading	Item	Item No.	Original Scale (Phenomenological)
1	599	doing homework thoroughly	12	Serious minded worker
2	479	working hard	2	Serious minded worker
3	444	sticking to the truth	4	Respect for authority
4	436	getting everything correct	21	Serious minded worker
5	318	accepting expert advice	24	High conformity
6	–313	questioning the truth of a text book	32	Respect for authority
7	–303	having a low opinion of yourself	3	High self concept (negative originally)

TABLE 24

FACTOR ANALYSIS AND PHENOMENOLOGICAL APPROACHES ON ATTITUDE SCALES : NEW ZEALAND

FACTOR I

Rank	Loading	Item	Item No.	Original Scale (Phenomenological)
1	–542	doing a dangerous experiment	22	High risk
2	–482	taking the lead	20	High risk
3	–467	running risks	9	High risk
4	–412	day dreaming	31	High risk
5	–376	being highly imaginative	5	High risk
6	–375	arguing with people	27	High risk
7	353	being neat and tidy	18	Respect for authority, High conformity
8	–341	making jokes	25	High risk
9	328	being very polite	30	Respect for authority, High conformity
10	–301	questioning the truth of a text book	32	Respect for authority

FACTOR II

Rank	Loading	Item	Item No.	Original Scale (Phenomenological)
1	503	accepting expert advice	24	High conformity
2	478	doing one's duty	33	High conformity, Respect for authority
3	447	being a good team member	36	High conformity
4	317	being obedient	1	Respect for authority
5=	316	being cautious	6	High risk (negative originally)
5=	316	getting everything correct	21	Serious minded worker

FACTOR III

Rank	Loading	Item	Item No.	Original Scale (Phenomenological)
1	–626	teasing people	16	High risk
2	–581	exaggerating	34	High risk, Freedom of emotional expression
3	–458	getting into trouble	13	Respect for authority (negative originally) High risk
4	–351	showing off	28	High risk
5	–335	smoking	23	High risk

FACTOR IV

Rank	Loading	Item	Item No.	Original Scale (Phenomenological)
1	604	showing respect for teachers	15	Respect for authority
2	579	doing homework thoroughly	12	Serious minded worker
3	539	being neat and tidy	18	High conformity, Respect for authority
4	478	being obedient	1	Respect for authority
5	433	working hard	2	Serious minded worker
6	390	being very polite	30	Respect for authority
7	384	sticking to the truth	4	Respect for authority
8	352	having self control	10	Freedom of emotional expression
9	348	keeping quiet in class	26	Respect for authority
10	334	being a good team member	36	High conformity
11	305	doing one's duty	33	Respect for authority, High conformity

Comparison between the earlier phenomenological approach and the later factor analysis approach in grouping items is interesting. For the most part there is agreement, although some unexpected groupings of original scale items exist within some factors (notwithstanding sign changes in loadings which are quite consistent).

ENGLAND, Factor I is mostly loaded with items from the original Respect for Authority and High Risk scales. Factor II is mainly composed of items from the original High Conformity, High Self Concept and Freedom of Emotional Expression scales. Factors III and IV comprise items from the original High Risk scale.

U.S.A., Factor I comprises items almost equally drawn from the original High Risk and Respect for Authority scales, and squares with England Factor I. Factor II comprises items almost exclusively from the original High Risk scale. Factor III represents a mixture from the following original scales: Freedom of Emotional Expression, High Self Concept, Respect for Authority, High Conformity and Serious Minded Worker. Factor IV comprises items from Serious Minded Worker, High Self Concept, Respect for Authority, and High Conformity.

NEW ZEALAND, Factor I is almost completely composed of items from the original High Risk scale and again resembles England and U.S.A. Factor I. Factor II draws items from the original High Conformity, High Risk and Serious Minded Worker scales. Factor III involves the original High Risk, Freedom of Emotional Expression, and Respect for Authority scales. Factor IV comprises a majority of items from the original Respect for Authority scale together with some from High Conformity.

On the whole a more meaningful relationship is seen to exist between the phenomenological and factor analysis approaches than might have been predicted. This may be partly attributable to the process of having three professional workers in the field independently rate each item for inclusion in each original scale. However, factor analysis adopted a more parsimonious approach and a more across-the-scale approach based on actual pupil perception.

The naming of factors attempts a harmony between mathematical and psychological meaning. In the present situation items were ranked on the basis of their factor loadings in order to assist in the naming process - those items with highest loadings being given special consideration. As the result of this procedure the following factor labels can be suggested:

England	Factor I	Respect - Obedience
	" II	Self Control
	" III	Self Concept
	" IV	Risk Taking
U.S.A.	Factor I	Respect - Obedience
	" II	Risk Taking
	" III	Self Control
	" IV	Serious Mindedness
New Zealand	Factor I	Risk Taking
	" II	Conformity
	" III	Self Control
	" IV	Respect - Obedience

It will be noticed from Table 25 that the number of items per Factor Scale varies considerably and that for Factor III England and New Zealand and for Factor IV England, there are too few items for any reliable individual, and probably small group, analysis.

TABLE 25

Distribution of Number of Items Per Factor Scale for Attitude Scale (Test 4)

Country \ No. of Items	Factor I	Factor II	Factor III	Factor IV
ENG	16	7	4	3
USA	11	7	15	7
NZ	10	6	5	11

(ii) High and low scorer comparisons on Attitude Factors

The degree to which high and low scorers on fluency and flexibility of thinking might be significantly different in their responses to items comprising attitude factor scales was investigated. In order to examine this question 24 sets of t-tests between high and low scores on fluency and flexibility, separately, and for each of the four factor scales for each country were calculated (Table 26).

Differences according to country, factor scale, fluency, and flexibility were observed as follows:
Country differences: U.S.A. yielded four sets of statistically significant differences, England two, and New Zealand one.
Factor scale differences: Factor II revealed three sets of significant differences; Factor III, two; Factor I, two; and Factor IV, none.
Fluency/flexibility differences: Flexibility accounted for five of the seven sets of significant differences, and would therefore appear more sensitive than fluency in revealing personality differences in this study.

TABLE 26

HI–LO FLEXIBILITY AND FLUENCY SCORERS ON ATTITUDE FACTORS

(t--Tests Between High (+1 S.D.) and Low (–1 S.D.)
Scorers on Fluency and Flexibility (separately) on each of the Four Attitude Factors)

Factor	Degrees Freedom	Means	Variance	No.Obs.	t	P <
ENG. Fluency						
1	1134	2.354 2.268	1.570 1.641	592 544	1.146	—
2	495	1.579 1.600	.799 .634	259 238	–.283	—
3	282	3.770 3.941	1.231 .893	148 136	–1.386	—
4	211	2.819 2.735	1.318 1.547	111 102	.513	—
ENG. Flexibility						
1	942	2.318 2.122	1.654 1.382	544 400	2.390	.01
2	411	1.533 1.485	.610 .478	238 175	.644	—
3	234	4.060 3.772	.956 1.205	100 136	2.075	.05
4	175	2.843 2.573	1.504 1.417	102 75	1.455	—
U.S.A. Fluency						
1	416	2.516 2.421	1.656 1.602	209 209	.764	—
2	264	4.240 3.676	.754 1.226	133 133	4.603	.01
3	568	1.842 1.592	1.024 .697	285 285	3.199	.01
4	264	2.037 2.255	1.705 1.498	133 133	–1.399	—
U.S.A. Flexibility						
1	427	2.722 2.400	1.789 1.694	209 220	2.524	.01
2	271	4.165 3.864	.814 1.188	133 140	2.469	.01
3	583	1.670 1.676	.796 .872	285 300	–.085	—
4	271	2.157 2.107	1.832 1.652	133 140	.316	—
N.Z. Fluency						
1	518	3.296 3.236	1.881 1.735	300 220	.502	—
2	310	1.966 1.840	1.054 1.073	180 132	1.061	—
3	258	2.373 2.136	1.580 1.408	150 110	1.531	—
4	570	1.990 1.851	1.190 .903	330 242	1.593	—
N.Z. Flexibility						
1	578	3.438 3.251	1.775 1.795	310 270	1.677	—
2	346	2.048 1.802	1.261 .677	186 162	2.294	.05
3	288	2.400 2.177	1.543 1.272	155 135	1.580	—
4	636	2.076 1.936	1.307 .854	341 297	1.684	—

The interpretation of high and low scorers toward items comprising each factor scale was obtained from the direction of significance of the two means (1-5 coding corresponding to "strongly approve" - "strongly disapprove" for positively loaded items and 5-1 for negatively loaded items). The results appear below:

England (Fx, $p < .01$) Factor I (Respect - Obedience)

- - High scorers in flexibility tended to be neutral toward, while low scorers tended to strongly approve of...

  "being obedient"
  "working hard"
  "sticking to the truth"
  "having self control"
  "doing homework thoroughly"
  "showing respect for teachers"
  "being neat and tidy"
  "keeping quiet in class"
  "taking things seriously"
  "being very polite"
  "doing one's duty"

  But high scorers tended to approve more of .....

  "getting into trouble"
  "teasing people"
  "doing a dangerous experiment"
  "smoking"
  "making jokes"+

England (Fx, $p < .05$) Factor III (Self Concept)

- High scorers in flexibility tended to disapprove of, while low scorers tended to be neutral about ...

  "having a low opinion of yourself"

  but high scorers tended to approve more of ....

  "investigating the unusual"
  "taking the lead"
  "getting everything correct"

+These last five items have negative loadings (Factor I is bipolar) and hence require reversal of initial interpretational rubric above. Such is also the case with England, Factor III; U.S.A., Factors I and III.

U.S.A. (F1, $p < .01$) Factor II (Risk Taking)

- High scorers in fluency tended to approve of, while low scorers tended to be neutral about...

"being highly imaginative"
"being enthusiastic"
"doing new things"
"investigating the unusual"
"taking the lead"
"doing a dangerous experiment"
"day dreaming"

U.S.A. (F1, $p < .01$) Factor III (Self Control)

- High scorers in fluency tended only mildly to approve of, while low scorers tended to strongly approve of ....

"being obedient"
"working hard"
"being enthusiastic"
"having self control"
"mixing well socially"
"doing new things"
"showing respect for teachers"
"being neat and tidy"
"accepting expert advice"
"being very polite"
"doing one's duty"
"not hurting other people's feelings"
"being a good team member"

but high scorers tended to approve more of ....

"teasing people"

"showing off"

U.S.A. (Fx, $p < .01$) Factor I (Respect - Obedience)

- High scorers in flexibility tended to be neutral toward, while low scorers tended to mildly approve of ...

"being obedient"
"being cautious"
"showing respect for teachers"
"keeping quiet in class"
"being very polite"

but high scorers tended to approve more of...

"running risks"
"getting into trouble"
"smoking"
"making jokes"
"arguing with people"
"day dreaming"

U.S.A. (Fx, $p < .01$) Factor II (Risk Taking)

- High scorers in flexibility tended to approve of while low scorers tended to be neutral about...

"being highly imaginative"
"being enthusiastic"
"doing new things"
"investigating the unusual"
"taking the lead"
"doing a dangerous experiment"
"day dreaming"

New Zealand (Fx, $p < .05$) Factor II (Conformity)

- High scorers in flexibility tended to only mildly approve of...

"being obedient"
"being cautious"
"getting everything correct"
"accepting expert advice"
"doing one's duty"

Conclusions from this section of Test 4 suggest a consistent pattern in the direction and trend of the seven sets of significant differences. High scorers on both fluency and flexibility are seen as more approving of running risks and being imaginative, but less approving of showing respect for authority, being a serious minded worker, and conforming, than low scorers. It is as if their increased fluency (i.e. generation of many ideas) and flexibility (switch of category) bring them into conflict with the status quo.

Nevertheless the degree and incidence of these findings for each factor scale and each country is uneven, and some caution is necessary. First of all, the actual amount of directional change, although significant, is small. It is certainly not a bipolar change - from strongly approve to strongly disapprove. It is only the direction of the attitude trend that significantly differentiates high and low fluency and flexibility scoring groups, and it would be incorrect to see high and low scorers on fluency and flexibility as being diametrically opposed on the factor scales.

The incidence of significant differences is uneven for the three countries. Taken overall they suggest a trend, but for either New Zealand or England separately, the results must be viewed tentatively. In both these countries the number of factors showing significant differences between high and low scorers on fluency and flexibility is small and within two of these factors the number of items is also small. For the U.S.A. sample, however, the evidence is more substantial.

Further study employing more specifically item-keyed factor scores with a wider sample of pupils would seem advisable before any generalizable conclusions could be made.

(iii) Chi-square analysis of high and low scorers (fluency and flexibility) on Personal Preference Questionnaire, Test 4, Part II

This short multiple choice questionnaire which formed the final section of Test 4 was of a personal preference type; e.g. If an experiment failed would you rather

(a) Find out why it failed? ☐

or (b) Start a different one? ☐

Results were analysed by employing chi-square and Fisher's exact probability tests to the responses of high and low scorers on fluency and flexibility separately for each item. In only two out of the 54 sets of probabilities were the differences beyond chance at the .01 or .05 level. It would therefore seem that unlike the factorised Part I (Attitude Scale), Part II proves a poor differentiator of extreme scorers in the fluency and flexibility domain.

C. SEX, CLASSROOM, TEACHER AND PEER RATING DIFFERENCES

Further chi-square analysis of high and low scorers

A series of non-parametric analyses employing chi-square and Fisher's exact probability tests were undertaken to examine expected frequency contributions in the above areas.

Sex Differences

The question of whether significantly more boys or girls might fall into either of the high or low scoring categories of fluency or flexibility of thinking is of some concern for science, and of some importance for certain statistical analyses in the present investigation. Chi-square analysis of high and low scorers on fluency and flexibility in U.S.A. and New Zealand yielded no significant differences in expected frequencies by sex (Appendix G). It should be stressed that this conclusion is valid only for the sample of pupils and test content of the present study. One explanation for this conclusion - which is contrary to the expectation of results drawn from a wider heterogeneous sample - is the very fact that these pupils represent the highest ability level in science that each school could offer. This means that a selective process has already occurred in the schools, whereby the girls in the sample are greatly superior to the average run of girls in science - more so than in the case of boys. This is borne out by the smaller number of girls in the sample to begin with. Another explanation (which may run pari passu with the former) is that, particularly in fluency and flexibility of thinking, girls may have a slight superiority over boys - despite any culturally-determined sex disadvantage in working with such science stimulus words as acid, oxygen, magnet, space, etc.

Classroom and school differences

How far different classroom and school samples contribute toward high and low scorers in fluency and flexibility is of some interest in terms of classroom climate, teaching style and other variables. However, the present study can only examine the extent to which they exist and not the causes for them.

Through the use of chi-square and the more stringent Fisher's exact probability test (for cell numbers less than five) the null hypothesis that the contribution to high (or low) scoring groups from each classroom and each school occurs by chance, was investigated. (Table 27)[+]

[+]See Appendix H for compilation of observed and expected frequencies by classroom and school.

TABLE 27

Classroom and School Contributions to Hi-Lo Scorers on Fluency and Flexibility, England, U.S.A., New Zealand. (Chi-square and Fisher's Exact Probability Tests)

Country	Classroom	Chi-square	Significance Level $p<$	School	Chi-square	Significance Level $p<$
ENG	01	Fl 4.9864	.05 (Fisher's exact test .17)	01	Fl 7.8576	.01
		Fx .8514	—		Fx .688	—
	02	Fl 1.8913	—			
		Fx .1704	—			
	03	Fl 3.1963	.05	02	Fl 3.1963	.05
		Fx 1.2437	—		Fx 1.2437	—
	04	Fl 2.0410	—	03	Fl 2.4466	—
		Fx 3.0773	.05		Fx .9117	—
	06	Fl 2.1154	—			
		Fx 1.7285	—			
	07	Fl .4865	—			
		Fx 1.4726	—			
	09	Fl 3.5751	.05 (Fisher's exact test .09)	04	Fl 3.5751	.05 (Fisher's exact test .09)
		Fx 1.6704	—		Fx 1.6704	—
	10	Fl 2.5816	—	05	Fl 1.4866	—
		Fx .3864	—		Fx .7284	—
	11	Fl .0360	—			
		Fx .3270	—			

TABBE 27 (continued)

Country	Classroom	Chi-square	Significance Level p<	School	Chi-square	Significance Level p<
USA	12	Fl .1666	—	06	Fl .1666	—
		Fx 2.1349	—		Fx 2.1349	—
	13	Fl .5230	—	07	Fl .0428	—
		Fx .0378	—		Fx .9102	—
	14	Fl 8.3692	.01)(Fisher's exact) test .20)			
		Fx 7.3565	.01)			
	15	Fl 1.2624	—			
		Fx 1.3532	—			
	16	Fl .4683	—	08	Fl .4683	—
		Fx .5072	—		Fx .5072	—
	17	Fl .0216	—	09	Fl .0404	—
		Fx .4301	—		Fx 1.1926	—
	18	Fl .0428	—			
		Fx .8418	—			

TABLE 27 (continued)

Country	Classroom	Chi-square	Significance Level p<	School	Chi-square	Significance Level p<
NZ	19	Fl 1.6391	—	10		
		Fx .2535	—			
	20	Fl .0303	—	11		
		Fx .9811	—			
	21	Fl 1.4707	—	12		
		Fx .6879	—			
	22	Fl 1.8794	—	13	Same as for classroom	
		Fx .7944	—			
	23	Fl .4029	—	14		
		Fx .0833	—			
	24	Fl .1866	—	15		
		Fx .9183	—			
	25	Fl .0284	—	16		
		Fx .1344	—			

*Chi-square calculations on Olivetti Programma 101 programmed with Yates correction for continuity factor.

Fisher's Exact probability calculations on I.B.M. 16/20 Computer.

The results indicated that for England the null hypothesis can be rejected beyond the .05 level in four out of nine classroom cases, although Fisher's exact probability lowers the strength in two cases. In three out of four cases the rejection was premised on fluency. Two classrooms contributed significantly more pupils to the high scoring fluency category than on the basis of expected frequency; one classroom contributed significantly more pupils to the low scoring fluency category, and one classroom contributed significantly fewer pupils to the high scoring flexibility category.

When school differences are examined for England three out of the five school samples indicated significant differences in their contribution to high and low scores on fluency and flexibility, although Fisher's exact probability test lowered the strength slightly in one case. In each of the three cases the significance was attributable to fluency.

For U.S.A. the null hypothesis can be rejected at the .01 level in only one out of seven classroom cases (but in both fluency and flexibility), although again Fisher's exact probability lowers the strength of the rejection. When school differences are examined for the U.S.A. sample no significant differences occur in their contribution to high and low scores on fluency and flexibility.

For New Zealand the null hypothesis could not be rejected in any of the seven classrooms (schools). Thus the tendency of the New Zealand and American classrooms in this sample to contribute scorers somewhat evenly to high and low scoring fluency and flexibility groups, is to be contrasted with the unevenness of the English classroom contribution in the sample.

Where classroom contribution is significantly different within a school it raises questions, particularly about teaching style (curriculum content and pupil ability can be looked upon as fairly constant in the present within-school samples). Where differences occur between schools but within the same sample by country, a fairly wide range of variables including curriculum content, specific attitudinal and cognitive variables, reward systems and teaching styles, would be open for consideration. Some of these issues will be examined further in Chapter six.

Teacher and peer rating categories

Each senior science teacher nominated four pupils from his class whom he considered displayed: the most original ideas in science (O); and, the highest ability in science (A). In addition he was asked to nominate: four pupils whom he enjoyed teaching most (B+); and four pupils whom he enjoyed teaching least (B-). Peer rating nominations were obtained for: pupils who displayed the most original ideas in science (O); and, pupils who worked the hardest in science (W).

Observed frequencies for each of the above categories were obtained for high-low fluency and flexibility groups (Appendix I). Observed and expected frequencies were then determined for chi-square calculations (Appendix J). Probability studies were then undertaken to test the null hypothesis that the number of responses in each teacher and peer rating category (O,A,B+,B-,O,W) relate to high-low scorers in fluency and flexibility on the basis of chance. The results appear in Table 28.

TABLE 28

Teacher and Peer Rating Probabilities (Chi-square) on the Basis of Hi-Lo Fluency and Flexibility.

England, U.S.A., New Zealand

(O = most original A = highest ability B+ = enjoy teaching most B- = enjoy teaching least
W = hardest working)

Country	Hi-Lo Fl. Px	Teacher Rating Category O	p	A	p	B+	p	B-	p	Peer Rating Category O	p	W	p
ENG	Hi-Lo Fl.	2.6750	-	.7919	-	.9100	-	.2937	-	17.0760	.01	2.0215	-
	Hi-Lo Fx.	1.4388	-	.5129	-	.0134	-	.0446	-	19.3304	.01	13.9779	.01
USA	Hi-Lo Fl.	1.7163	-	5.2707	.05	.0020	-	.3421	-	21.1367	.01	107.0469	.01
	Hi-Lo Fx.	.4464	-	1.8115	-	.0029	-	.1107	-	25.6825	.01	15.3629	.01
NZ	Hi-Lo Fl.	1.0230	-	1.9013	-	.1643	-	.9402	-	.5046	-	9.7259	.01
	Hi-Lo Fx.	.0559	-	.0435	-	.3488	-	.6091	-	2.3537	-	24.0014	.01

Chi-Square and Probabilities

Olivetti Programma 101 employing Yates correction for continuity factor.

Teacher rating categories

For England no teacher rating category reached a level of statistical significance at or beyond the .05 level when comparing high-low scores in fluency and flexibility. The category which attained the highest level of significance was 'originality' which reached the .10 level for fluency and .20 for flexibility - teachers assigning more pupils in the 'originality' category to the high than to the low scoring test groups.+

For U.S.A. only one teacher rating category reached a level of statistical significance at or beyond the .05 level. This occurred for the 'highest ability' category where significantly more pupils in the high scoring fluency group were nominated by teachers as being of the highest ability. Similarly the trend for originality nominations is in the same direction at the .20 level.

For New Zealand no teacher rating category reached a level of statistical significance. However, the two categories with the highest level were 'highest ability' (.10) and 'originality' (.30) favouring the higher scoring fluency group in each case.

+See Popham (1967) pages 53-54 for descriptive use of p levels up to .30.

Peer rating categories

Peer ratings for 'originality' and 'hardest working' significantly differentiated high and low scoring groups in fluency and flexibility.

For England peer rating nominations reached significance ($< .01$) in the originality category for both fluency and flexibility, thus rejecting the null hypothesis. Peers assigned significantly more originality nominations to pupils who scored high in both fluency and flexibility than to low scoring pupils. (A similar but not so statistically significant trend was observed for teachers). It is perhaps surprising that high and low scorers on fluency and flexibility should be differentiated by peers (and to some extent by teachers) on originality ratings. This raises the question as to how teachers and peers interpret rating categories such as 'original'. Peer rating nominations for 'hardest working' in science also reached significance for flexibility ($< .01$) - fewer observed frequencies being found in the low scoring category than expected by chance.

For U.S.A. peer rating nominations again reached significance ($< .01$) for both originality and hardest working categories favouring high scoring fluency and flexibility groups in each case.

For New Zealand peer rating nominations for 'hardest worker' reached significance ($< .01$) favouring the higher scoring fluency and flexibility groups. Surprisingly, however, the previous trend in the English and American samples for peer ratings in originality to reach significance did not occur for New Zealand, although the tendency was apparent (.10) for the high scoring flexibility groups.

Individual analysis of high peer rated pupils

A more individual analysis of first and second highest peer rated pupils for 'originality' and 'hardest worker' in science was undertaken for each class.[+] The sample for which this information was available comprises 53 pupils from 18 classrooms. Teacher ratings and mean fluency (Fl M) and flexibility (Fx M) scores were then determined for each pupil (Table 29). This analysis highlighted a greater degree of accordance between teacher and peer ratings than might have been expected from the chi-square analysis, although similar trends between teacher and peer

---

[+]Because of the incidence of some composite classes a constant extraction of two pupils was not always appropriate. Furthermore the same pupil sometimes achieved a 'first' rating in one category and a 'first' or 'second' rating in the other category.

TABLE 29

Analysis of Selected Sample of High Peer Rated Pupils

(O = most original, A = highest ability,
B+ = like teaching most, B- = like teaching least,
W = hardest working)

Country	Class	Script	Peer Rating O.	Peer Rating W.	Teacher Rating	Fl M (T Scores)	Fx M (T Scores)
ENG	01	24	1st	2nd	O,A	52	61
	01	22		1st	A	59	61
	02	35	2nd		-	63	60
	03	44		1st	O,A,B+	55	56
	03	58	1st		O,B+	64	52
	03	51		2nd	A	64	72
	03	42	2nd		O,A,B-	66	67
	04	101		1st	A	40	42
	04	120		2nd	A	49	43
	06	68	1st	2nd	O,A	45	57
	06	80	2nd		O,B+	57	67
	07	122		1st	A	46	47
	07	121		2nd	-	42	46
	07	115	1st		O,A,B+	57	62
	07	118	1st		O	70	76
	09	141	2nd	1st	O,A,B+	53	48
	09	145		2nd	A	53	57
	09	128	1st		-	49	54
	10,11	Partial information only					
	N = 7 classes comprising 131 pupils. Present selected sample = 18						
USA	12	287	1st	1st	-	38	36
	12	300		2nd	O,A	51	47
	12	291	2nd		O,A	67	65
	13,14, 15	Partial information only					
	16	354		1st	A	59	47
	16	350	1st	2nd	O,B-	47	57
	16	363	2nd		O,B+	45	52
	17	380	1st	1st	A	64	60
	17	373		2nd	-	71	54
	17	379	2nd		O,A	55	52
	18	405		1st	O,A	65	52
	18	406		2nd	-	55	59
	18	420	1st=		O,A,B+	92	98
	18	410	1st=		A	66	60
	N = 4 classes comprising 108 pupils. Present selected sample = 13						

TABLE 29 (continued)

Country	Class	Script	Peer Rating O.	Peer Rating W.	Teacher Rating	Fl M (T Scores)	Fx M (T Scores)
NZ	19	442		1st	O,A,B+	67	74
	19	456	1st	2nd	O,A,B+	51	54
	19	446	2nd		O,A	45	44
	20	494	1st	1st	O,A,B+	54	57
	20	492	2nd		A	46	44
	20	475		2nd	O	40	49
	21	507		2nd	O,A	64	51
	21	513	1st	1st	O	55	50
	21	524	2nd		A	76	75
	22	566		2nd	O,A	57	56
	22	549	1st	1st	O,A	32	29
	22	572		2nd	-	61	58
	23	579	1st	1st	O,A,B+	81	86
	23	595		2nd	A	56	38
	23	588	2nd		O,A,B+	63	51
	24	638	1st	1st	O,A	58	57
	24	639	2nd		O	42	48
	24	632		2nd	O,B+	36	35
	25	646		1st	A	44	33
	25	649		2nd	A,B+	47	44
	25	660		1st=	-	39	40
	25	669		1st=	B-	55	68

N = 7 classes comprising 184 pupils.
Present selected sample = 22

ratings on originality had been already noted. Furthermore the present analysis is different in confining attention to the two highest rated pupils only - a much sharper focus.

Not only is there a strikingly high degree of similarity between teacher and peer assessment for these pupils but their expected frequencies for assignment to high scoring (+1SD) groups (F1 M and Fx M) is higher than expected from chance for the English and U.S.A. samples, but not for the New Zealand.

The ten highest peer nominations for originality in the English sample were matched by eight teacher ratings for originality. This is a remarkably high degree of parallelism when it is realized that these eight out of ten matched pairs are drawn from a teacher rating sample of 131 pupils and not the 18 of this special sample. Much the same is seen between the matching of the eleven highest peer nominations for hardest worker in science and the ten teacher ratings for most able in science. Apparently the different wording ('hardest worker', and 'most able') counts for little in altering the similarity of perception for the two groups. Again, teacher equivalence is based on the larger pool of 131 pupils. When test scores

(Fl M, Fx M) for peer nominations in England are considered, five cases (Fl M)[1] and eight cases (Fx M)[2] fall above +1SD where the expected frequency in each case is three (1/6 of 18).

The seven highest peer nominations for originality in the U.S.A. sample were matched by five teacher rating nominations when the latter were drawn from a pool of 108 pupils. The eight highest peer nominations for hardest worker in science were matched by four teacher nominations. Both of these results represent quite a high level of agreement. When test scores are considered for U.S.A. six cases (Fl M)[3] and four cases (Fx M)[1] fall above +1SD where the expected frequency in each case is two (1/6 of 13).

The eleven highest peer nominations for originality in the New Zealand samples were matched by nine teacher rating nominations when the latter were drawn from a pool of 184 pupils. This represents a very high degree of equivalence. The sixteen highest peer nominations for hardest worker in science were matched by eleven teacher nominations which again is high. When test scores are considered for New Zealand six cases (Fl M)[1] and four cases (Fx M)[1] fall above +1SD

---

1 = Not significant
2 = Significant at the .01 level
3 = Significant between the .01 and .02 level

where the expected frequency in each case is between three and four (1/6 of 22).

This section of the study, which has sharpened the focus to the one or two most highly rated pupils in each classroom, has shown a high incidence of agreement between teacher and pupil nominations. In many cases these highly rated pupils have also scored in the +1SD group on the present tests. This is not to deny that some discrepancies exist - e.g. Pupil 435, 'A' (Highest ability) teacher rating, zero peer rating for hardest worker in science, although third highest peer rating for originality; Pupil 652, 'O' (originality) teacher rating, zero peer rating. Such discrepant cases would certainly be interesting to study further. More intensive analysis might be profitable, incorporating fewer classrooms, closer parallel of language for rating categories, and more precise statistical analyses (e.g. biserial correlations).

CHAPTER FIVE

RELIABILITY AND VALIDITY

Reliability

Two aspects of reliability have already been mentioned: (i) interscorer reliability (pp. 70-71) and (ii) equivalent-parts reliability as determined from correlations between tests 2A/B and tests 3A-B/C-D (pp. 96-97). Coefficients of equivalence or internal consistency were determined for these tests. Employing the Spearman-Brown prophecy formula for the estimation of reliability, the coefficients for test 2A/B were:

England	.76 fluency	.73 flexibility
U.S.A.	.75 fluency	.71 flexibility
New Zealand	.74 fluency	.70 flexibility

and for tests 3A-B/C-D

England	.71 fluency	.67 flexibility
U.S.A.	.79 fluency	.66 flexibility
New Zealand	.76 fluency	.77 flexibility

These results compare favourably with those obtained by Wallach and Kogan (1965) who reported coefficients ranging from .51 to .93 with the same method.

The main emphasis in the present section on reliability, however, will be given to test-retest reliability (coefficients of stability), including an analysis of the stability of extreme scorers, and of sex differences. A related issue to retest reliability will also be examined. This involves the degree of variability involved in the re-marking of the original responses of the retest sample after an identical time interval. Such a procedure also tests the reliability of scoring procedures.

Three classroom samples comprising 74 pupils (40 males, 34 females) were drawn from three New Zealand schools, one of which was coeducational, one single sex male, and one single sex female. The time interval between retest situations was eight weeks.

Means, standard deviations and t-tests between means were calculated for eight major sets of fluency and flexibility scores for the test-retest situation. No significant differences occurred on any test variable, thus giving one indication of a measure of stability of scores (see Appendix K). Nor was there any trend for an increase in scores which might reflect a practice effect.

Coefficients of stability were then calculated for the eight test variables, and these appear in Table 30.

TABLE 30

Coefficients of Stability for 74 New Zealand Pupils on Eight Test Variables with a Retest Interval of Eight Weeks

Test Variable	Coefficient of Stability
Fl 1	.68
Fx 1	.67
Fl 2M	.57
Fx 2M	.62
Fl 3M	.63
Fx 3M	.46
Fl M	.72
Fx M	.63

These are moderately high coefficients and typical of those reported in the literature. Yamamoto (1962) employing a test-retest interval of ten weeks reported coefficients of stability of .79 for fluency and .64 for flexibility on Torrance Tests. As would be expected, the mean of all fluencies (Fl M) and the mean of all flexibilities (Fx M) yield the highest reliability coefficients. Fluency tends to be slightly more reliable than flexibility, and Test 1 slightly more reliable than the other tests.

An analysis of extreme scores ($\pm$1SD) indicated some interesting shifts toward greater stability. These results can be seen in Table 31.

TABLE 31

Coefficients of Stability for Extreme Scorers (New Zealand, N = 28 for $\pm$1SD) on Eight Test Variables with a Retest Interval of Eight Weeks.

Test Variable	Coefficient of Stability
Fl 1	.91
Fx 1	.84
Fl 2M	.59
Fx 2M	.67
Fl 3M	.80
Fx 3M	.82
Fl M	.88
Fx M	.88

Test 1 again appeared the most reliable single measure, with less difference between fluency and flexibility in terms of reliability than was the case with the total reliability sample. The above results are encouraging and suggest high retest reliability for extreme scorers.

Sex differentiation in retest reliability showed no significant differences. Means, standard deviations and t-tests between means were calculated for both the original test and the retest situations. In neither case was there sufficient evidence to suggest significant sex differences. In the original test situation only one significant difference (.05 level) was observed over the eight test variables, while again only one at the .05 level was observed in the retest situation. (See Appendix L). Coefficients of stability for each sex separately did not differ significantly from each other nor from the total reliability sample.

A related issue concerning the amount of variance resulting from the re-marking process itself was also examined. In order to assess this the responses of the total reliability sample (N = 106, including those sick or absent on retest) were re-marked after the same eight week time interval. Coefficients of correlation between scores from original and re-mark occasions on all test variables were consistently high (.95 to .99) indicating a high degree of reliability in the scoring procedures. Thus it would seem that most of the variance unaccounted for on retest can be attributed to variation of trait behaviour.

In general these results reflect moderately high levels of reliability for measures of fluency and flexibility of thinking, with higher stability for extreme scorers. Since the results are also obtained from fairly closely related measures (fluency and flexibility) and within a fairly restricted range of ability, they can be considered satisfactory. Replication over a wider range of ability (as is customary in the sampling for coefficients reported in many test manuals) would be likely to yield higher coefficients than those reported above.

Validity

In terms of psychological measurement, validity is usually considered by way of the question: Does the test measure what it purports to measure? On this count tests of so-called creativity and so-called intelligence can be readily criticised. Since both creativity and intelligence represent complex concepts, the issue of validity is involved and the responsibility of establishing validity somewhat onerous. If, for example, giving as many different uses of a brick is claimed to be a measure of creativity, or repeating five digits backwards a measure of intelligence, then there is a responsibility on the part of the test designer to substantiate the claim.

The concept of divergent thinking, being less inclusive than creativity, reflects a narrower universe of meaning on which validity is required to be established. Extra claims associated with creativity need not be met. The responsibility of establishing validity still remains, however, particularly if several dimensional scores are combined into a single score and labelled 'divergent thinking'.

The position in the present study represents an even narrower universe, viz, fluency and flexibility of thinking. There is no claim that the tests are measuring creative or divergent thinking, or that conclusions derived from scores are generalisable outside the present instruments and samples. Therefore, in terms of validity it is not necessary for the measures of fluency and flexibility to meet any total criterion-complex associated with creativity or divergent thinking. One of the dangers of tests related to creativity has been the assigning of titles that encompass a much wider universe of meaning (and hence claim) than warranted. Sometimes investigators have begun constructing a test confined to a particular area, but once the study has developed, have been tempted to change the descriptive title to a more encompassing one. A study which represents a good example of the 'name-changing game' (McNemar, 1964) and which displays a sense of anticlimax in terms of validity, is that of Wallach and Kogan (1965). After designing tests scored for fluency and uniqueness of associates the authors immediately refer to them as 'creativity instruments'. The study is described as a "methodologically rigorous investigation of that little understood, variously defined psychological ability called creativity" (and claimed on the dust cover to be concerned with an "empirical validation" of a distinction between intelligence and creativity), yet the authors make no mention of the validity of their 'instruments of creativity' either in the text or in the index. If such investigators score for fluency and uniqueness of response and then call their results 'measures of creativity' the responsibility of establishing validity is a heavy one. Wallach and Wing (1969) have taken pains to discuss the issue in their later commentary.

Some investigators have questioned measures of creativity and divergent thinking when derived from timed or from 'test-like' situations (Ginsburg and Whittemore, 1968; Nicholls, 1971), but the position seems to be very open, with counter claims reported by Guilford, 1959, 1963, Kogan and Morgan, 1969, and more recently by Vernon, 1971.

However, no claim is being made for scores on the present tests to be measures of creativity, although they are two dimensions that are commonly regarded as part of divergent thinking.

In the light of the above comments the concept of validity and its relationship to the present series of tests requires careful consideration. In terms of what it purports to measure, 'fluency' of response has rational meaning as the result of the total number of words or ideas produced within a specific time limit. The term 'fluency' has been traditionally associated with such a characteristic of verbal behaviour in psychology throughout this century. In the light of what it purports to measure 'flexibility' has rational meaning in terms of change (or non-change) in the associative flow of words or ideas. Categories form the basis of scoring taxonomies for determining such change. As mentioned earlier (p.57) the taxonomies in the present study represent reasoned groupings of words and concepts based on responses from the sample. The taxonomies have also been examined by university staff in science. High coefficients of interscorer reliability indicate that the taxonomies can be interpreted consistently among different scorers, a finding which further supports evidence concerning validity. Nevertheless, flexibility of thinking remains a more complex measure than fluency.

It is unnecessary, in terms of validity, to prove that a test measures a particular trait or behaviour and nothing else. It is doubtful if any psychological test can achieve this. Thus, to say that the ability to repeat digits backwards measures intelligence and nothing else, or memory and nothing else, or that the present tests measure fluency or flexibility and nothing else, is not mandatory. However, it is reasonable to expect that tests, including the present ones, represent a relatively high proportion of the measure claimed of them. The independence of fluency and flexibility from intelligence indicates separation from this domain. The permanency of the fluency/flexibility relationship after the use of partial r is also confirmatory of their more fundamental nature.

The area of validity poses serious problems for psychometricians. Discussion of the topic customarily centres on face validity, content validity, construct validity, concurrent validity and predictive validity. Face validity is concerned with the plausibility of the test (often derived from the test title), and motivational concomitants associated with rapport. In this respect the present tests reflect a parallelism between task demands (i.e. to give as <u>many different</u> responses as possible) and nature of scoring (i.e. word or idea count, frequency; and category change, flexibility). Rapport was readily established through the tests.

<u>Content validity</u> is usually referred to as the match between test objectives and instructional objectives. As such it has little relevance to the present tests. If content validity, however, can be interpreted more broadly in terms of the sample of a specific universe of behaviour, the current tests may be judged to have a measure of adequacy. This can be claimed on the basis of three tests which, including subtests, yield seven sets of primary scores for each dimension of fluency and flexibility. Of course wider sampling, up to a point, would increase validity, but the balance of test length, width of objectives, reliability and validity remains continuously delicate.

<u>Construct validity</u> is generally referred to as the extent to which a test illustrates something about a meaningful characteristic of the individual. Evidence of a test's specific construct validity is usually derived from (i) its correlation with other accepted measures of the same construct;

and (ii) its correlation with other characteristics of the individual e.g. attitudes and interests. Fluency and flexibility of thinking have generally been established as meaningful aspects of divergent thinking by such investigators as Guilford (1964) and Hudson (1966, 1968), and to a lesser extent as meaningful aspects of creativity by Torrance (1966) and Getzels and Jackson (1962). Since there still remains some doubt as to the nature of this 'meaningfulness', the extent of its construct validity remains somewhat open - as of course is the case with any other construct.

Taken by themselves, however, fluency and flexibility of thinking are evidenced as meaningful characteristics of an individual's behaviour in many specific contexts by way of oral and written language and performance generally. In terms of the specific construct validity of the present tests no account can be given for (i) above, since the investigator is unaware of other tests employing science stimuli in the open-ended idiom against which a correlation coefficient could be determined. Such another test would also be required to command considerable respect in the literature.[+]

Some evidence of the specific construct validity of the tests in terms of (ii) above can nevertheless be seen from high/low scorers on fluency and flexibility of thinking and other cognitive and affective variables. Differentiation in terms of high scorers exhibiting some superiority in measured intelligence and some attraction toward certain attitudinal factors associated with high risk, non-conformity, disrespect for authority and other factors, lend support for a theory of meaningful relationships between the cognitive and affective domains as well as relating to other research in the wider field of scientific creativity (pp.179-180)

Factorial validity can be looked upon as another aspect of construct validity. In the present study factor analysis of both cognitive and attitudinal variables helped to throw further light

---

[+] The procedure of correlating tests against 'respected' or 'time-honoured' tests - in the way that some intelligence tests are correlated against the Stanford Binet - is of doubtful value if there is fundamental disagreement over the concepts underlying the time-honoured test. In this respect the concept of validity is more truthfully seen as somewhat circular and relative, despite assumptions of linearity, authoritarianism and absolutism in the writings of many research workers and text book authors. Furthermore, so-called coefficients of validity often mask basic reasoning about the concept itself.

on the nature of intercorrelations of the tests and factors common to them (pp. 103-107). The pervasiveness of basic factor conclusions across cultures has been reported by Dacey, Madaus and Allen (1969) as support for construct validity. Such pervasiveness was substantially indicated in present results (p. However, factorial validity tends to be inward looking, and by itself, insufficient (McNemar, 1964; Vernon, 1964).

Concurrent validity is generally referred to as the test's ability to distinguish between two or more groups of individuals whose status at the time of testing is different. In the present circumstances, for example, high/low scorers on fluency and flexibility should be seen to perform differently according to the thinking styles rewarded in classroom science attainment tests, and to be perceived by teachers and peers as having outstanding characteristics in behaviours related to fluency and flexibility. However, the absence of a reliable and valid measure of divergence/convergence bias within the repertoire of teacher-made classroom tests meant that no precise comparison with high/low scorers on the present tests could be made.

Some evidence of teacher and peer observation of characteristics associated with high and low scorers on fluency and flexibility of thinking was nevertheless determined. In terms of teacher ratings only mildly supportive evidence was obtained. This indicated that teachers tended to associate high scorers on fluency and flexibility with 'originality' category, and to a slightly lesser extent with the 'highest ability in science' category. But the level of significance was more often between .10 and .30 than between .01 and .05. In terms of peer ratings, however, the differentiation between high and low scorers was more significant (often beyond the .01 level). Again, high scorers were associated with the 'originality' and 'hardest worker' category. These conclusions were more characteristic of the English and U.S.A. samples than of the New Zealand sample. It is most likely that the desire to reward, and the ability to perceive thinking styles associated with fluency and flexibility, are very uneven amongst teachers within the overall sample.

Predictive validity is generally associated with forecasting the future status of individuals on the basis of test scores or some other measure. One of the difficulties in establishing predictive validity is the establishment of a suitable and reliable criterion, or criteria, against which to correlate the dependent variable. Predictive validity again illustrates the somewhat circular and arbitrary nature of the concept of validity. Nevertheless, over time, predictive validity should be attempted. Such an investigation is contemplated for the present tests and is referred to under 'Further Research', (p. 221). Evidence of the long-term predictive capabilities of dimensions of divergent thinking comes from such studies as Cropley and Field (1969), Field and Poole (1970), and Hamilton (1970). In the study by Cropley and Field, scores on divergent thinking tests predicted terminal grades at university more successfully than either entering grades or I.Q.

In conclusion, the reliability and validity of the present tests while not comparable with some more-conventional measures of educational attainment, are nevertheless in line with measures of cognitive style and attitude. Bearing in mind the somewhat restricted range of ability from which the reliability coefficients have been determined, quite satisfactory levels have been achieved. When current criteria for evaluating the validity of tests have been considered, satisfactory evidence for the measures has been indicated. In the light of the above results and comments, it may be judged that the reliability and validity of the instruments and procedures used are as satisfactory as might be achieved in this area at the present exploratory stage, and compare more than favourably with most other work in the field.

CHAPTER SIX

DISCUSSION

This section elaborates in a more general way some of the specific details of results discussed in chapter four, and relates them to contemporary research. The discussion is organized under the headings: (A) Cognition, (B) Personality, and (C) Educational Implications.

A. COGNITION

Three tests of open-ended thinking in science were designed for the present study emphasizing fluency and flexibility of thinking. Following Guilford's (1964) fluency factor terms, Test 1 (Science Vocabulary) stresses 'word fluency', Test 2 (Definition of Scientific Concepts) 'expressional fluency', and Test 3 (Uses of Chemical Substances and Physical Objects) 'ideational fluency'.+

+Word fluency: the production of many word associates premised on a specific letter, word, or concept.

Expressional fluency: the production of many short phrases or sentences associated with a specific word or concept.

Ideational fluency: the production of many words or phrases associated with classes of objects or uses of objects.

Increasing importance is being attached to fluency as a component of divergent thinking in contemporary research. Renzulli (1970) claims that "fluency as measured by a simple frequency count of relevant responses may well be the single most important contributor to the variance of creativity" (p.4), and continues, "Wallach (in press) has pointed out that when the effects of intelligence are partialled out of measures in the divergent thinking domain, a 'cognitive disposition remains that seems most adequately characterized as ideational fluency' .... Assessment tasks that might tap this factor would appear to be a highly valuable line of research endeavour" (p.5).

Related to fluency in each of the tests in the present study is an analysis of flexibility of thinking. Flexibility is premised on category change in the associational flow of ideas, and is closely related to breaking a mental set (Jackson and Messick, 1967). In terms of Guilford's (1964) classification, spontaneous flexibility - a restructuring process in problem-solving, was emphasized in the tests. More recently flexibility of thinking has been stressed as an important 'modality of divergent thinking' and as a valuable predictor of certain personality traits (Carlier, 1970).

In selecting pupils of high ability in science to respond in the open-ended idiom to stimuli in the science domain, the present study differed from others which have presented more general stimuli to pupils whose interests and abilities range more widely across the arts and sciences. Evidence (Cramer, 1968) suggests that increased fluency and flexibility of response occur as the stimulus material approaches the sphere of the subject's interests and abilities. Thus it would seem that individuals have different levels of fluency and flexibility and that what is being reported in the current study is a specific cognitive level toward the upper end of the productive continuum. However, differences in the appeal of stimulus test-material, and in individual differences to produce associates, reduce scores from any general hypothetical 'high'. These same influences contribute toward a range of scores in the dimensions of thinking under study. Indeed the range of scores in fluency and flexibility is one of the most striking characteristics of responses in the study. Consider for example the following range of raw scores in fluency (Test 1):

England	11-50 (mean 27.8)
U.S.A.	9-58 (mean 27.6)
New Zealand	19-62 (mean 36.0)

and in flexibility (Test 1):

England	1-28 (mean 13.4)
U.S.A.	2-28 (mean 13.7)
New Zealand	1-31 (mean 15.5)

This wide range of individual differences is all the more graphic when contrasted with the rather narrow range of scores on traditionally-styled classroom and standardized achievement tests for such highly streamed pupils.*

The present study also differs to some extent from others (Getzels and Jackson, 1962; Wallach and Kogan, 1965) in examining fluency and flexibility as separate dimensions of thinking and not as later relabelled combined single scores representing divergent or creative thinking. It seems unwarranted to combine different measures into a single score before first examining the relationship between the specific components. Nor has the present study employed any combined single score measure against which a comparison with intelligence is made (cf. Anderson, 1968; Bradfield, 1969; Getzels and Jackson, 1962).

*Some teachers remarked that there was little point in giving these pupils science tests (including the present ones) since pupils in their class were so consistently bright that "there was nothing between them". The results, in terms of fluency and flexibility, seem to suggest that in fact there is a great deal between them.

Wallach and Kogan (1965) paused momentarily to examine fluency and uniqueness of associates as separate dimensions of thinking, but then proceeded to describe their scores in a unified way, and to refer to their tests as the "Ten Creativity Instruments" (Table 7, p.46) claiming that the scores represent a "unified, pervasive dimension of cognitive activity, which is properly characterizable as creativity" (p.25). Wallach and Wing (1969) further confound the issue by saying that the two indices of fluency and uniqueness were "strongly related ..... regardless of task context ...." (p.15). This seems to represent a surprising misinterpretation of Wallach and Kogan's actual results which indicated intercorrelations varying from .07 to .74. Such a range surely represents a considerable degree of 'task context'. Moreover, several intercorrelations were quite low, as can be seen from the following:
.07, .08, .20, .22, .24 (Table 7, p.46). Again these represent anything but strong correlations. It is not surprising that Wallach and Wing (1969) then state, "....we find ourselves facing an interpretational problem when we seek a psychological account of what may underly individual differences in these characteristics" (p.15).

Several correlational outcomes between fluency and flexibility might be predicted from such studies (including the present investigation): Individuals who score high on fluency (or on flexibility) on one test might also be expected to score high on other tests - this could certainly be so in the present study since all tests are centred on the same specific subject area, science. Such an interpretation postulates a relatively high correlation on the cognitive style traits of fluency and flexibility across tests but makes no prediction between the traits within tests.

Another outcome could be predicted whereby individuals who score high (or low) on one test in fluency might also be expected to score high (or low) in flexibility. This interpretation postulates a relatively high correlation between fluency and flexibility within tests but makes no prediction about traits across tests.

Finally a third position could be predicted in that individuals who exhibit fairly stable degrees of cognitive style across tests might also to a degree be sensitive to task context, the relative emphasis either way depending on many variables.

With correlations of .6, .7 and .8 between fluency and flexibility within tests and of .3, .4 and .5 for fluency and for flexibility across tests in the present study, the third position above seems to hold, with the weight of emphasis favouring a task-specific theory, but moderated by some stability of cognitive style across tests. Within this pattern, fluency was seen to correlate a little higher than flexibility across tests. Some weight can be given to these results since they were consistent across the three samples by country (England N = 182; U.S.A. N = 140; New Zealand N = 184).

While the pattern of relationship was consistent across the three samples by country, differences in cognitive behaviours were nevertheless apparent in the present sample. With the emphasis on divergent and creative thinking in American literature it might have been thought that pupils in the American sample (who were also involved in progressive science teaching programmes) would have been superior to the English and New Zealand samples on fluency and flexibility of thinking. It is perhaps surprising therefore that both the English and New Zealand samples obtained generally higher scores in these two dimensions of thinking.+

+New Zealand scored significantly higher than U.S.A. in six out of seven sets of scores on fluency, and four out of seven on flexibility; England scored significantly higher than U.S.A. in four out of seven on fluency and in two out of seven on flexibility.

It seems unlikely that these differences can be attributable to differences of intelligence, for, despite the English sample's superiority in this respect, correlations between intelligence and fluency and flexibility overall were low. Furthermore, the influence of intelligence on fluency and flexibility was examined by way of a first order partial r, holding intelligence (AH5) constant, with no significant changes in the fluency or flexibility relationship being observed.[+] Nor can sex differences in the three samples by country explain the difference in scores on fluency and flexibility, since t-tests (U.S.A. and New Zealand) yielded no significant differences between girls and boys in this respect. It is possible, however, that the greater use of multiple-choice tests which often require a simple check-type response has had the effect of reducing attainment in fluency and flexibility of thinking in the American sample.

Cutting across these results are some striking individual differences not only in level of scores but also in cognitive and response styles. Some individuals for example, displayed highly stable patterns of response style across all tests while

---

[+] Similar results were obtained by Yamamoto (1967), and Wallach (in press) in Renzulli (1970).

others (contributing to the task-specific nature of results) were somewhat erratic. In some ways the detailed analysis of these individuals may be more interesting than the general results but they are not detailed in the present investigation.

Fluency, Flexibility and Intelligence

The results of the study generally indicate a fairly substantial degree of independence between the cognitive style traits of fluency and flexibility and measured intelligence. Zero to low positive correlations were obtained between (a) intelligence and (b) fluency, and flexibility - with slightly higher coefficients of correlation being indicated for fluency. Cross-cultural differences were also small, suggesting a fairly pervasive trend. Thus the correlations between intelligence and the mean of all fluencies (Fl M) for each sample were as follows:

England .30; U.S.A. .24; and New Zealand .01.

The correlations between intelligence and the mean of all flexibilities (Fx M) were:

England .16; U.S.A. .22; and New Zealand .01.

Similar findings have been recorded by Wallach and Kogan (1965) and Wallach and Wing (1969) who report low coefficients of correlation between (a) intelligence and (b) fluency, and uniqueness of associates (e.g. .09, -.04, .14, .11 etc.). Wallach and Wing (1969) conclude that "Ideational fluency and fluency-dependent forms of uniqueness of ideas seem to define individual-difference characteristics that have relatively little to do with intelligence" (p.18). Other studies have reported similar correlations between intelligence and combined scores representing creative thinking (Dacey and Ripple, 1969; Dewing, 1970).

The charge is sometimes made of such studies that they have been carried out within a very narrow range of intelligence and when replicated on wider samples correlations as high as .54 between 'intelligence' and 'creativity' have been obtained (Yamamoto, 1964). While this may be so, the charge has less weight than may appear at first sight. In the first place restriction of range is always a matter of degree. Secondly, investigators are often not interested in 'people in general', nor in generalising their findings beyond the boundaries of their sample. And in the third place restriction of range is often more apparent than real. In this last respect, for example, Cropley (1968) was surprised by the wide range of intelligence scores (AL/AQ) from within what some critics might have charged was a selected sample (124 first year male university students, Australia). He concluded, "small correlations between creativity and intelligence tests cannot be dismissed as an artifact of reduced variance in the intelligence variables" (p.198). A similarly wide range of intelligence scores (AH5) were obtained from the selected sample of the present study. In these circumstances the need for correction for attenuation may be less necessary. Nevertheless no interpretation beyond the described sample is being made in the present investigation, and the issue of restriction of range is accepted as a warning against overstating the case of independence of intelligence and dimensions of divergent thinking. Certainly no simple relationship between intelligence and high and low scorers in fluency and flexibility of thinking was seen to exist in the present study. When high, low and a middle control group of scorers in fluency and flexibility were compared for significant differences on measures of intelligence it was found that high scoring fluency groups (and to a lesser but still significant extent, high scoring flexibility groups) in England and U.S.A. were superior on measured intelligence. Similar results were reported by Callaway (1968). The degree of independence between creativity and intelligence reported by Getzels and Jackson and others have also not been supported by other research workers (Rittmayer, 1968). Several factor analysis studies which have reworked Wallach and Kogan's (1965) results have also painted a less dramatic picture of independence between their 'creativity' scores and intelligence scores. Thus Fee (1968) while supporting the <u>relative</u> independence of the two cognitive modes added "this independence is clearly not as complete as Wallach and Kogan maintain. 'Creativity' is clearly not unidimensional and there is a hint that it may have relationships with factors of ability not explicitly taken into consideration by Wallach and Kogan and not specifically represented in their set of tests" (p.321). Similarly Cropley and Maslany (1969) reported that a factor analysis of Wallach and Kogan's study indicated the existence of "large loadings of creativity tests on the intellectual factor and of intelligence tests on the creativity factor" (p.395), concluding that "The tests measure a stable and internally consistent intellective mode, albeit one which is substantially related to general intelligence" (p.398).

<u>Factor Analysis, Cognitive Variables</u>

Employing a centroid analysis with samples by country three test-specific factors based on a composite fluency-flexibility component were disclosed in the present study. The same pattern occurred across the three cultures with the common priority of Test 3 (Uses of Chemical Substances and Physical Objects), as Factor I, accounting for the highest percentage of variance. The other two test fluency-flexibility factors were based on Test 1 (Scientific Vocabulary), and Test 2 (Definition of Scientific Concepts). Whether this suggests that the tests can also be considered at a higher level, dividing into a practical science (Uses of Chemical Substances and Physical Objects) versus a verbal science factor (Scientific Vocabulary, Definition of Scientific Concepts), is interesting to consider. Be this as it may, the factor analysis does refocus attention on the nature and demands of each task within the science domain. Apparently each test demands somewhat different abilities despite the overriding request to think fluently and flexibly. It is even likely that a different interest component relevant to the stimulus material of each test underlies the ability differential. Thus both cognitive and attitudinal aspects appear to be interacting to contribute to the task-specific nature of the results.

Finally, no factor that could be interpreted as a general intelligence or a general fluency or flexibility factor cutting across all tests appeared to result from the analysis. These results also seem to suggest some caution in referring to fluency or flexibility (or divergent or creative thinking) as a general trait existing outside a particular field of endeavour, or even as a unified trait <u>within</u> a discipline like science.

Until recently few studies in divergent and creative thinking have emphasized its task-specific nature. The beginnings of this research-trend saw a contrast in performance <u>between</u> specific fields as represented by the arts and sciences. More contemporary research, however, is highlighting task specificity <u>within</u> the one field of endeavour (Field and Cropley, 1969; Getzels and Csikszentmihalyi, 1966; Dewing, 1970; and Lovell, 1968). But apparently the observation is not commonly regarded in science, since Stenhouse (1971) recently commented that "few investigators appear to have recognised that 'creativity' is a fundamentally heterogeneous category even within the field of science" (p.171).

Other factor analysis studies which have relevance for the present investigation include those by Shouksmith (1970) whose k_3 factor (creative associating) contained major loadings in two ideational fluency measures, and Sultan (1962) whose Factors II and III contained loadings on both ideational fluency <u>and</u> spontaneous flexibility. As with the present study, no general factor nor broad verbal factor was reported by Sultan, his explanation being that Grammar school pupils are highly selected for verbal intelligence and attainment.

Employing a wide range of test variables with New Zealand pupils, Reid (1970) first employed an unrotated principal axes solution which disclosed ten factors. Factor I was interpreted as a general ability factor and Factor II a bipolar factor contrasting intelligence and creativity variables. However, he added that since the correlation between the two factors reached .313 any claim that the two dimensions are completely independent must be treated with caution (see also Cropley, 1968, p.200). Reid then applied an orthogonal rotation to the Varimax criterion which "revealed very clearly the pattern which had begun to emerge in the earlier solution. Each factor was defined by a particular test or test battery" (p.298). As within the present study, it was noted that "the finding was presaged to a degree by the strong within-test and weak across-test interrelationships revealed in the correlational analyses" (ibid. p.300).

Since few studies have employed factor analysis with divergent thinking variables in a cross-cultural setting it is difficult to find comparisons for the present study in this respect. However one factorial study with American and Irish Grammar School pupils (Dacey, Madaus and Allen, 1969) indicated similar factorial structures differentiating divergent thinking and intelligence across age and culture. They concluded that "while the magnitude of the scores might be uniformly affected within a culture by such variables the trait itself is not affected, hence the factor structures remain constant" (p.265).

B. PERSONALITY

The attitude scale designed for the present study incorporated six scales which were initially seen as being closely interrelated, and somewhat tentative - being partially dependent on planned subsequent factor analysis. Separate factor analyses for the three samples by country yielded four factors - a somewhat more parsimonious solution than the six empirical scales. However, there was a fairly substantial agreement between the two approaches in most cases. The factor analysis was nevertheless seen as contributing toward the validity of the original instrument and of value in a later analysis of high and low scoring groups on cognitive variables.

Some cultural differences in pupil self-perception were reflected in the factor solutions, although some marked similarities occurred between countries on item loadings for one factor, Respect-Obedience.+

Although three of the four attitude factors++ differentiated between high and low scorers on fluency and flexibility in seven cases, differentiation was neither dramatic, generally, nor substantial, cross-culturally. The most marked differentiation

---

+The four factors for each country were as follows: England: Factor I, Respect-Obedience, mainly loaded with items from the original Respect for Authority scale; Factor II, Self Control, mainly loaded with items from the original High Conformity, High Self Concept and Freedom of Emotional Expression scales; Factor III, Self Concept; and Factor IV, Risk-Taking, mainly loaded with items from the original High Risk scale.
U.S.A.: Factor I, Respect-Obedience; Factor II, Risk-Taking; Factor III, Self Control; and Factor IV, Serious Mindedness.
New Zealand: Factor I, Risk-Taking; Factor II, Conformity; Factor III, Self Control; and Factor IV, Respect-Obedience.

++Factor I (England and U.S.A.) Respect-Obedience; Factor II (U.S.A.) Risk-Taking, (N.Z.) Conformity; Factor III (England) Self Concept, (U.S.A.) Self Control.

occurred with the U.S.A. sample where three out of four factors indicated differences significant at the .01 level. For England, two of the factors indicated differences (both for flexibility) at the .01 and .05 level. For New Zealand, only one factor-difference in flexibility (.05) was significant. Nevertheless, all seven differences (as well as several others at the .08 level) were supportive of major research trends covering relationships between cognitive and attitudinal variables. Thus high scoring fluency and flexibility groups were generally more approving of such traits as running risks, doing new things, doing dangerous experiments, teasing people, making jokes, investigating the unusual, taking the lead, arguing with people and day dreaming; but less approving of being obedient, accepting expert advice, working hard, getting everything correct, having self-control, mixing well socially, doing one's duty, and not hurting other people's feelings. Although no factor-label or specific item used the terms 'independence' or 'self confidence', obviously the above cluster of differentiating items suggests that high scoring pupils in fluency and flexibility may indeed be characterised by greater independence and self confidence than low scoring groups.

If any general reason can be given for the present pattern of attitude trends between high and low scoring fluency and flexibility groups it may be that increased fluency of associates - referred to as 'cognitive energy' by Wallach and Wing (1969) - together with a similar measure of flexibility evidenced by enjoyment in ranging over many different ideas, excites these individuals to adventure more, and by so doing experiment, run risks, and at times question established authority. However, in this complicated network of connections, cause-effect relationships remain complex and mostly unknown.

Flexibility proved to be more significant than fluency in differentiating these personality traits (cf the cognitive domain where fluency showed slightly better discriminatory power). But the actual amount of directional change on the attitude scale, although significantly differentiating high and low scores on fluency and flexibility, was small. The difference certainly does not reflect a bipolar change from strongly approve to strongly disapprove.

A multiple-choice type personal preference questionnaire, with some items based on attitudinal and behavioural preferences said to differentiate creative scientists from non-creative ones, gave disappointing results.

Items such as the following did not differentiate significantly between high and low scorers on fluency and flexibility:

1. If an experiment failed would you rather
   (a) Find out why it failed? ☐
   or (b) Start a different one? ☐

2. If more time were given to science at school would you like to
   (a) Read about science? ☐
   (b) Try out new ideas of your own? ☐
   or (c) Work with the teacher on experiments? ☐

3. Do you think pupils should be encouraged to
   (a) Reach a good even standard of work? ☐
   or (b) Follow their special interests? ☐

Only the following three items significantly differentiated high and low scoring groups on fluency and flexibility:

1. Do you prefer to
   (a) Work out theories? ☐
   or (b) Observe results? ☐

   (The high scoring fluency scorers (England) favoured (a)).

2. Do you prefer to have
   (a) One or two friends? ☐
   or (b) Many friends? ☐

   (The high scoring flexibility scorers (England) favoured (a)).

3. In building something from plans do you prefer to
   (a) Make up your own plans? ☐
   or (b) Work from ready-made plans? ☐

   (The high scoring flexibility scorers (New Zealand) favoured (a)).

While these particular differences and their direction may not be surprising, they are too few to indicate anything worthwhile. It is quite possible that while some characteristics reflected in similar items to those employed in this preference questionnaire have differentiated adult creative scientists from non-creative ones, the same characteristics have either not yet crystallised out as differentiating factors amongst young secondary school pupils or have not been discernible on the narrower measurement basis of fluency and flexibility of thinking.

Several other studies have relevance for the present investigation. With a similar sample of pupils by country (104 Japanese first year high school students) Iwata (1968) reported that high scorers on ideational fluency were more confident, more dominant and more independent than low scorers. Employing flexibility as the major cognitive variable Carlier (1970) reported several significant differences between high and low scorers on Cattell's 16 PF with a group of undergraduates. High scorers on flexibility were positively related with C, H, E and Q_1 factors, but negatively with G. Taking the G factor first, Carlier's results indicated that high scorers on flexibility can be characterized as refusing to be bound by rules, being casual and lacking in effort for group undertakings. Low scorers can be characterized as having a strong sense of duty, perseverence and responsibility, and as being conscientious and moralistic - preferring hard working people to witty companions. Other factor descriptions for high flexibility scorers are: Factor C, emotionally stable, realistic about life, unruffled, possessing ego strength; Factor H, bold, independent, ready to try new things, spontaneous and abundant in emotional response, careless of detail; Factor Q_1, interested in intellectual matters, enquiring regarding ideas, doubtful on fundamental issues, and more inclined to experiment than moralise. Some striking similarities exist between Carlier's findings with high scoring flexibility groups on certain factor-scores of Cattell's 16 PF, and results between high and low scoring flexibility groups on factor attitude scales in the present study.

Other studies, although employing single combined scores derived from divergent or creative thinking tests, also relate in some details to the present investigation. Highly creative adolescents were described as displaying above average self-confidence (for boys only) in a study by Kurtzman (1967), and as being independent, non-conforming and seeking out change in their environment by Cashdan and Welsh (1966).

Another investigation (Taft and Gilchrist, 1970) which studied the relationship between high scorers on creative <u>attitudes</u> and creative productivity found the group to be impulsive, imaginative, and risk-taking.

Other research on attitudes closely related to findings of the present investigation is that of Hudson (1966, 1968). Hudson has reported a tendency for convergers to adopt authoritarian attitudes. In the present investigation similar outcomes were evident for pupils who scored below one standard deviation from the mean in flexibility (and was paralleled to some extent with fluency). In other words they appeared to behave as convergers. On the other hand similar attitudinal traits to Hudson's divergers were seen for the present sample of <u>science</u> pupils who scored <u>above</u> one standard deviation from the mean in flexibility (and to some extent with fluency).

Hudson (1968) also investigated two other traits that are relevant. The first was concerned with the question of 'yielding' to authority. In reporting tendencies for convergers to yield to authority Hudson suggested that "low mental fluency may well be linked to a more general susceptibility to pressure from authority: that the individual with a taste for the 'one right answer', both in answering mental tests and in his life's work is particularly susceptible to information about what his elders and betters think correct" (pp. 9-10). In so far as high scorers in fluency and flexibility can be aligned with divergers and low scorers with convergers, support for this relationship between authoritarianism and convergent thinking comes from the present study where Factor I, Respect-Obedience (with items loaded mainly from the original authoritarian scale) significantly differentiated high and low scorers on flexibility (England) and fluency (U.S.A.).

However the question of 'yielding' and 'suggestibility' associated with authoritarianism in Hudson's research was not studied in the present investigation. And, in fact, results with divergent and creative students in this connection seem to vary. For example in one study with ten-year-old pupils, those who were highly creative were also more open to suggestion (McHenry and Shouksmith, 1970). The authors, however, noted the influence of age as an important variable in explaining apparently contrary conclusions from some other studies. "So the creative child," they comment (p.159), "probably becomes much less suggestible by his mid teens". The developmental stage of an idea, as well as age, may be an important variable whereby creative individuals, as compared with non-creative, may be more perceptually open (in a way equivalent to suggestible) at the beginning stages of an idea, and more 'closed' (in a way equivalent to resisting suggestions from authority) at the concluding stages. This may be partly at the basis of differences in such studies as Hudson's (1968) and McHenry and Shouksmith's (1970), although major differences exist in terms of age of sample, definition of terms and test instruments employed.

Hudson's second study which has relevance for the present investigation was concerned with 'conscientiousness'. He reported that girls who were convergers rather than divergers tended to be more conscientious, divergers more rebellious and intellectually independent. However, Hudson suggested that wider sampling was necessary before the result could be considered in any way conclusive. Several items from factors in the present study can be aligned with conscientiousness e.g. 'showing respect for teachers', 'doing homework thoroughly', 'doing one's duty', 'being neat and tidy', 'working hard', 'taking things seriously', 'sticking to the truth' in Factor I, Respect-Obedience, (England), as well as similar items for Factor I (U.S.A.), and 'accepting expert advice', 'doing one's duty', 'being obedient', and 'getting everything correct' in Factor III, Self-Control (U.S.A.). In each of these cases low scoring groups in fluency and flexibility did tend to approve more of the above traits. However, the very scale (Factor IV, Serious Mindedness (U.S.A.)) which, by title at least, appears to come closest to a conscientiousness trait, did not significantly differentiate extreme scorers. The position therefore still seems open.

Of two studies more specifically related to the science domain, one (Pearce, 1968) examined a group of male high school students who showed promise in scientific research. Amongst characteristics which differentiated them from fellow science students were high levels of self sufficiency and introversion. These findings relate to one of the few items on the Personal Preference Questionnaire in the present investigation which did significantly differentiate high and low scoring science pupils on fluency and flexibility, i.e.

Do you prefer to have (a) one or two friends? ☐
or (b) many friends? ☐

Similar conclusions have been reported by

MacKinnon (1965, 1966, 1966a) and Roe (1953) with research scientists, and by Walberg and Welch (1967) with innovative physics teachers who were seen to have low needs for 'group affiliation'.

Associative responses and personality relationships related to introversion/extraversion have been described by Cramer (1968) in terms of the way the individual interacts with 'verbal' society. Cramer holds that the associative behaviour of extraverted individuals is characterized by high response commonality (low flexibility) thus reflecting social norms, whereas the associative responses of introverted individuals are characterized by more personal and idiosyncratic patterns (high flexibility). Cramer reports that the creative individual not only offers more responses but also more idiosyncratic responses (see also Mednick, 1962; and Wallach and Kogan, 1965).

Some results in the broader field of social behaviour nevertheless appear contradictory and difficult to resolve in the light of the above findings. In part this reflects the complexity of variables within the social context, and in part problems of definition. Thus while several studies (Barron, 1969; Cramer, 1968;

MacKinnon, 1962, 1966a) have suggested closer alignment between creative thinkers and introversion*, others have indicated that creative thinkers display more social skills and fewer withdrawing tendencies (Motlagh, 1968), and greater social participation and interest in helping others (Windholz, 1968).

One trait, complex in itself but considered only in a relatively simple way in the present study, remains to be discussed. This is risk-taking. The tendency for high scorers in flexibility to approve more of taking risks than low scorers was evident in the present study. Similarly Pankove and Kogan (1968) reported "a significant relationship between an associational fluency factor (one type of divergent thinking) and personality inventory measures of risk-taking (need for adventure) ..." (p.421).

Relationships between divergent thinking and risk-taking in terms of conservatism, accuracy or carefulness of responses on cognitive variables, and other aspects have been studied by several research workers (Cropley, 1969; Hudson, 1966; Kogan and Wallach, 1964), and suggest similar conclusions to those cited above.

---

*cf Hudson (1968) who reported <u>no</u> significant relationships on the introversion scale of the Maudsley Personality Inventory for convergers and divergers.

C. EDUCATIONAL IMPLICATIONS

These will be considered under two sections, (i) educational implications of the present study, and (ii) wider educational implications and related research.

(i) Educational Implications of the Present Study

Within the fairly narrow range of high science ability in the study, and within an area where conventional class tests usually result in a relatively narrow range of scores, fluency and flexibility of thinking highlighted a wide range of individual differences. However it is necessary to consider the value of such information. To many science teachers, the present set of scores on fluency and flexibility - being based on relatively unfamiliar criteria and lacking correlation with I.Q's and more convergent-type power scores in science, may appear to represent inconsequential, or discrepant results. But herein is to be found part of their value.

In the first place they differentiate pupils (already tightly clustered on more conventional tests) in terms of idea generation, so that high scorers are characterized by being highly productive and flexible. Such thinking is of value in scanning widely before arriving at a single solution, and in highlighting 'possibilities' before 'actualities'. These traits are of considerable value to the research scientist.[+] Outstanding pupils on present test results therefore represent a particular type of science student (quite possibly atypical) who holds some promise in furthering the development of scientific thinking. In the second place scores on these dimensions of divergent thinking have been shown to be of value in predicting higher stages of academic performance (Cropley and Field, 1969). Although no long-term predictive evidence has been compiled from the present investigation, other studies which will be detailed later (Cropley and Field, 1969; Field and Poole, 1970) have shown such dimensions of thinking to be more valuable than either I.Q. or beginning grades as predictors of final success at university for entering science students.

It is unknown whether scores in the present study have such predictive power for final grades in science at high school. And unlike groups of university science students who are fairly committed to the discipline, the present sample continues to receive a fairly general education for some years. Academic bias towards science may be in its formative stages,

---

[+] Kuhn, 1962; MacKinnon, 1962; Mackworth, 1965; Roe, 1953, 1953a.

but it is by no means set. Cognitive bias, interpreted here in terms of high/low scores on fluency and flexibility of thinking, is nevertheless clearly evident, whereby high scorers in fluency and flexibility can be aligned with divergent thinkers and low scorers with convergent thinkers. It appears that the present sample is at a critical period of development where style of teaching in science may influence further bias, away from or toward the subject. What begins as cognitive bias <u>within</u> science can easily grow to academic bias toward or away from it.

It is interesting to consider whether those with divergent or convergent thinking tendencies will excel at secondary school science or whether divergers will gradually need to change to convergers (or play a convergent role) if they are to prosper. It may be that the top scholars in both arts and science are indeed thoroughly flexible in the sense that they know when to be concise and convergent, and when to be flexible and divergent. They may have the ability to keep one or the other mode in check for relatively long periods, and to gauge how different teachers reward each mode. On the other hand some scorers who reflect divergent thinking tendencies in science may not wish to be so adaptive and consequently come to enjoy secondary school science less and less, even opting for more arts subjects in the upper forms.

The above discussion brings into focus differences in trends of approval and disapproval by pupils concerning certain attitudes and personality traits which differentiated high and low scorers on fluency and flexibility of thinking. On the whole, high scoring groups as a result of tending to approve rather more of running risks, questioning authority and rather less of being obedient, being accurate and mixing well socially, are likely to engender a certain degree of antagonism from some teachers. Differences of a personality kind between pupils and teachers however should not be allowed unduly to prejudice judgments of academic performance involving either objective or related affective criteria.

It would nevertheless be false to give the impression from the previous discussion that high scorers on fluency and flexibility in the present study stood in complete isolation when judged by other attainment criteria more well-known to teachers, or that they were perceived as inveterate trouble makers. Such was not the case. Teacher ratings on pupil originality and ability in science showed some alignment with high and low scorers on fluency and flexibility (favouring high scorers) although the level of significance was generally between .05 and .30. Peer ratings on originality in science, however, were more significant (.01) in differentiating high and low scoring groups in fluency and flexibility (see also Yamamoto, 1964). Teacher ratings which coupled both high originality <u>and</u> high ability (often with the addition of 'enjoy teaching most', i.e. O, A, B+ categories pp. 141-2) comprise a fairly common grouping. This tendency for originality and ability to be perceived together was also noted by Lovell (1968). Furthermore the above teacher groupings were often supported by peer groupings of either 'most original' or 'hardest worker' in science (see also Dewing, 1970). In general then a moderately supportive network of connections appeared to exist between high scorers on fluency and flexibility of thinking on the one hand, and teacher and pupil perception of those with most ability and originality in science on the other. There was no trend for high scorers in fluency and flexibility or for those rated as most original by teachers or peers to be assigned to the 'enjoy teaching <u>least</u>' (B-) category. This finding conflicts with those of other studies (Getzels and Jackson, 1962) which have reported a tendency for highly creative pupils to be troublesome, and to be disliked by teachers. Perhaps the narrower dimensions of fluency and flexibility did not reveal the strength of irritation apparently felt by teachers of pupils considered 'highly creative' in some other studies.

The extent to which particular classrooms in each sample by country contributed high or low scorers in fluency and flexibility of thinking carries with it some educational implications. Four out of nine classrooms and three out of five schools yielded significant differences in this respect within the English sample. In some cases a particular classroom contributed significantly more high scorers than expected by chance, in others significantly fewer. One classroom out of seven in the U.S.A. and none in the New Zealand sample reached significance at the .01 or .05 level. However, several cases at the .10 to .20 level for both U.S.A. and New Zealand indicate trends of significance in the distribution of high and low scorers by classroom and school in the sample as a whole.+

+In fact many of these are significant at the .05 level or beyond if the conventional formula for chi-square is employed rather than the present one with Yates' correction for continuity factor.

One reason for cross-cultural differences in this respect may be the greater local autonomy of schools in the English sample (Public and Grammar Schools), and contrasting approaches taken toward teaching science. Thus some secondary school science departments reflected an anti-, others a pro-Nuffield Foundation Science teaching philosophy. Such tendencies were not apparent in the U.S.A. or New Zealand samples to such a degree, although no doubt wider sampling would reveal similar tendencies.

Intelligence level, attainment in science, and to some extent syllabus coverage and socio-economic background are fairly common factors within samples by country, so that other factors such as teaching style (including type of questioning sequence), type of thinking rewarded in class discussion and examinations, and specific syllabus content areas that may or may not have matched test content areas, are likely to have accounted for some of the above classroom differences. The type of classroom management was also found to be significant in one study (Turner and Denny, 1969) which reported "As scores in teacher organization increase, changes in pupil ideational fluency ....decrease" (p.269).

No significant sex differences were found between high and low scorers in fluency and flexibility of thinking in the present study. This may be partly explained by the selective educational procedures which have occurred over a period of time resulting in equally high standards of science attainment for girls as for boys in the sample. The smaller number of girls in the classes also attests to this explanation. But if it still could be claimed that the boys were of superior academic ability of a convergent problem-solving kind in science, there is evidence from some studies that girls are superior to boys in fluency and flexibility of thinking thus compensating for a degree of less in science attainment (Abney, 1970; Gallagher and Jenne, 1967; McCutcheon, 1970; Motlagh, 1968). Furthermore, nearly all the girls in the present sample were members of coeducational classes receiving the same quality teaching as the boys. It could be argued that differences might occur between boys and girls drawn from single sex schools on open-ended tests employing science stimuli.

Other investigations, as well as the present study, have indicated no significant sex differences in dimensions of divergent thinking (Carlier, 1970;

Cashdan and Welsh, 1966; Dacey and Ripple, 1969; Dewing, 1970; Tibbetts, 1968). Others have indicated superiority for boys on originality (Coone, 1969) and on flexibility and category width (Field and Cropley, 1970). The complexity of these findings seems to indicate the influence of such variables as age, type of test, and selectivity of sample on sex. It also suggests that cultural rather than biological differences may be at the basis of differences that occur with increasing age and specialisation.

The final implication which will be considered in this section derives from the finding that fluency and flexibility tend to be fairly substantially related within tasks. There are two points to note here. The first is that pupils who were highly fluent in the present study were also amongst the most flexible. The second is that although there was evidence of a measure of stable cognitive style across tests, it was overshadowed by a tendency for traits to be task-specific. It would seem then that teachers should be cautious of being influenced by halo effects which may determine expectancies towards stable styles and set levels of performance rather than being sensitive to variation of performance within a subject area such as science, according to the nature of the task.

(iii) Wider Educational Implications and Related Research

It has been claimed by Crutchfield (1969) that, "of all cognitive skills the typical school does least with the nurturing of idea generation. Depressingly little attention is given to the stimulation of ideas in most schoolwork at most levels. This is partly because of the deadening influence of the conventional curriculum and teaching methods with their heavy stress upon the 'authority' of established knowledge" (pp.61-62). Crutchfield also holds that it reflects an uneasiness on the part of teachers to cope with the production of many different ideas, and of a belief that "'ideational fluency' is a deeply rooted individual trait and not at all susceptible to training" (p.62). Many educationists, however, claim that the school can encourage the development of fluency and flexibility of thinking, and indeed has a responsibility to do so (e.g. Williams, 1967). It has also been suggested that the school offer more autonomy in pupil thinking, and more variety and personal choice of curriculum (Cohen, 1968; Crutchfield, 1969; Lovell, 1968; MacKinnon, 1966; Suchman, 1964).

Questioning which focuses students' attention on hypotheses and predictions and which offers open-ended interpretation instead of prematurely insisting on the 'one right answer' or the answer the teacher 'wants', is seen to encourage the development of divergent thinking (Taba, 1963; Taba, Levine and Elzey, (1964) in Fredrick and Klausmeier, 1970). Teachers can often contribute more by skilful questioning sequences which assist such thinking processes as classifying, inferring and hypothesizing, than in serving ready-made knowledge. But unfortunately, says Karplus (1964), "Teachers rarely ask a question because they are really curious to know what the pupils think or believe or have observed. The pupils of course, quickly adapt to this situation. After a few years, answering questions is for them more a mind-reading proposition than a matter of reasoning about the substance of a scientific problem" (p.10).

Several studies have been concerned with the quality of questioning by science teachers and pupils. In one (Washton, 1967) a taxonomy of five basic types of questions for classification purposes and for assisting in the growth of creative thinking was devised. In another (Solomon, 1969), an observational instrument incorporating a taxonomy of cognitive behaviours of science teachers which was aimed at stimulating productivity of ideas in terms of fluency, flexibility and mental imagery was developed. In the development of such thinking not only teachers but also pupils should be encouraged to ask questions; not only research workers but also pupils should be encouraged to become problem-finders and not simply problem-solvers (Mackworth, 1965).

Nevertheless, pupils who are problem-finders, and who are questioning, flexible and original, run many risks in any educational system that emphasizes passivity and conformity in learning. Such pupils, remarks Cropley (1969), have a "disconcerting knack of producing highly unusual responses and completely unexpected solutions.. .... The teacher may find himself badly disconcerted and even deflated by an apparently irrelevant answer, at some vital point in a lesson" (p.75). Cropley continues, "Many teachers take the easy way out and treat all divergent behaviour as a threat and a nuisance, imposing severe sanctions against it" (p.79). Another unfortunate consequence for such pupils is the close parallelism between teaching style, educational objectives and type of examination assessment. At no point does there appear to be much reward for open-ended thinking; instead, convergence to a single correct solution is too often accepted as the dominant, if not only mode. In this respect Wallach and Kogan (1965) highlighted the contrast between research such as their own which studied and rewarded the production of many varied and unique associates, and the conformity and authoritarianism of many school examinations (p.332). Wallach and Wing (1969) go so far as to suggest ideational fluency as a useful measure in the "delineation of talent".

A number of teaching programmes and other experiments have been concerned with developing fluency and flexibility of thinking and nurturing creativity. The influence of teaching methods on ideational fluency, flexibility, originality, self evaluation, and achievement in elementary science was examined by McCormack (1970) who reported that the experimental group was found superior to the control group in gains in fluency, flexibility, originality and affective understanding of course objectives, but that no significant differences occurred in achievement in science nor in cognitive understanding of course objectives.

In a science teaching programme which emphasized divergent thinking processes rather than specific content, some significant differences in classifying with fluency and flexibility as the criteria occurred, although few significant gains in creative thinking were reported (Ransom, 1969). In another teaching project concerned with styles of categorization in science it was shown that significant gains in intuitive mode and in flexibility of classification occurred for pupils taught by the inquiry process (Scott, 1966).

Audio and video tape recordings were used by Trowbridge (1969) to assess the extent to which divergent thinking categories were employed by teachers and the extent to which these increased as the result of a training project. Teacher behaviours associated with divergent thinking categories rose from an initial eleven per cent to a concluding 25 per cent of total thinking category membership during a ten month period. This rise was significant beyond the .05 level over matched control groups.

Experiments employing association techniques as an aid to creative problem solving (Gall and Mendelsohn, 1967; Miller et al., 1970), and originality (Khatena, 1970), have not been so unequivocal. Similarly, answers to questions concerning the degree of structuring and non-structuring of knowledge in creative problem solving remain open (Smoke, 1970).

It is noteworthy that nearly all these studies emphasized thinking processes rather than mastery of content. Thus Field and Cropley (1970) stressed the need for examining a variety of thought processes in the study of science which included fluency, flexibility, originality and category width. They deprecated an over-emphasis on content while thinking processes were often left to fend for themselves.

Many of the studies cited above also stressed the complexity of the variables and a degree of task specificity, despite a hope for generality and transfer (Craver, 1969). They also emphasized, often by contrast with some teachers' expectations, that such creative productivity is achieved through a direct consideration and understanding of thinking processes which require a sense of effort and discipline in the teaching-learning situation. Creativity certainly does not appear to be 'all sweetness and light' nor does permissiveness seem to be the dominant teaching climate reflected in the research. "Thus far from adopting an easy, non-pressured learning system", remarks Shouksmith (1970), "the school which wants to produce creative thinkers should ensure that each child develops to the full his knowledge and powers of associating items of knowledge." (p.215).

Another fairly common finding from these studies was the contribution of fluency, flexibility and originality toward cognitive styles "which an individual employs rather persistently in a variety of different cognitive tasks" (Taba, Levine, Elzey (1964) in Fredrick and Klausmeier, 1970), and which are rather stable across time (Witkin and Oltman, 1967).

The academic bias of pupils towards arts or science reported by Hudson (1966, 1968) appears to become quite marked in many cases by the time pupils enter university. Progressively through university the results become more complex. In the first place it is claimed that a great deal depends on the particular university and its degree structure. Thus Hartley and Beasley (1969), found no significant differences in bias between 'A' level science students and 'A' level arts students at Keele. However, the impact of the case was reduced by the small sample size (N = 22 for science) and vague criteria. Christie (1969) also examining the position at Keele, reported a less striking difference in bias than Hudson, but reported that his conclusion did not contradict Hudson's basic findings. Secondly, it seems that complexity is evident in the progressive development through the degree structure with accompanying differential emphases in teaching style and type of thinking rewarded. For example, Field and Poole (1970) reported that despite the contrast in intellectual style between students entering the science and arts faculties, by the end of the first year it was convergers in both faculties who excelled in examinations, but that by the end of the second year divergers were catching up in science as well as in arts, suggesting a reversal of bias within the science faculty for top scholars. Field and Poole concluded that although the majority of science specialists entering university are convergers it is mainly the divergers among them who finally achieve the better results. Similarly Cropley and Field (1969) reported that "In a longitudinal study covering the four years required for an undergraduate honors course in science, it was shown that men graduating with honors came almost exclusively from among those who had been rated highly divergent in their style of thinking on entry to the university four years previously" (p.134). Neither grades nor I.Q. predicted this so well.

It is interesting to consider when and how such aspects of cognitive style develop. Wood (1970) attributed it partly to the mother's feeling of optimism toward the infant, which in turn assisted attitudes of self-confidence and adventuresomeness. Similarly Povey (1970) stressed the importance of early parent-child relationships in the formation of a child's early cognitive bias - long before specialist curriculum courses appear. Povey's study reported:

> "firm evidence that differences of cognitive bias exist between future arts students and scientists before they have received any specialist education. There is some indication that such differences increase with age and further academic specialization. These findings lend support to the view advanced by Hudson (1966, 1968) that arts/science differences amongst schoolboys are related to more general personality differences which are fairly clearly defined before a choice of specialism is made." (p.55)

Mackay and Cameron (1968) also pointed out that younger pupils who "have little 'natural' bias to convergence or divergence, having elected a particular course of study are influenced by the modes of thinking demonstrated by their tutors." These again usually reflect the dominant styles of thinking of the Department (Arts/Science). Support for this thesis also comes from a study by Lytton and Cotton (1969) who found 'Departmental climate' an appropriate criterion for studying and differentiating divergent/convergent thinking.

The studies outlined in this discussion illustrate the constant interaction of forces representing immense possibilities for change on the one hand with those of increasing crystallisation and specialisation on the other.

CHAPTER SEVEN

CONCLUSION: SOME LIMITATIONS AND, SUGGESTIONS FOR FURTHER RESEARCH

By comparison with the complex network of connections surrounding this general area of study, and the broad educational implications derived from it, the setting and scope of the present investigation may appear somewhat restrictive. In addition, efforts to avoid the overgeneralisations which have bedevilled studies of creative thinking, may have resulted in approaches and conclusions that are, at times, highly specific.

As a result of keeping the study closely centred on two related dimensions of thinking, and on seeking a match between the abilities and interests of the sample and the nature of the tasks, both advantages and disadvantages have been seen to accrue.

The main advantages appear to have been:

(i) some harmony between the presentation of science stimuli for open-ended thinking and the subject's interest and ability;

(ii) a wide range of individual differences on measures of fluency and flexibility for pupils otherwise closely grouped in terms of attainment in science; and,

(iii) a detailed analysis of the characteristics of the two modes of thinking including their relationship with intelligence and certain attitudes.

Disadvantages may appear to have resulted from:

(i) keeping the study rather sharply focussed;

(ii) analysing fluency and flexibility in statistical terms; and

(iii) establishing only limited connections between fluency/flexibility and the wider universe of teacher and classroom variables.

The statistical treatment of results in the present research design leaves on one side the analysis of the material in terms of the richness of individual responses, highlighting humour, aggression, and degree of abstraction. Furthermore, it is appreciated that pupils with similar statistical counts on fluency or flexibility sometimes differ widely in terms of other qualitative variables. Quantity of response should therefore not be confused with quality of response in the present investigation, although some studies have shown them to be positively correlated (Skager, Schultz, and Klein, 1965). The relationship between the production of a _single_ solution and _many_ solutions to a problem in terms of their rated quality has also been considered (Johnson, Parrott, and Stratton, 1964).

The lack of some acceptable grade or score in science attainment that might have acted as a bench-mark against which fluency and flexibility and other scores could have been correlated might appear as a limitation. However, no suitable cross-cultural test in science was available at the commencement of the investigation, although the I.E.A. Studies may make this more feasible in the future. If different standardized tests were selected for each country, difficulties would continue to exist. And if end of year grades in science were employed, difficulties over selecting the specific field in science (e.g. chemistry, physics, biology) on which the grade was to be based would remain.

Conclusions which related associations of high/low fluency and flexibility groups to intelligence and to attitudinal factors were of limited dimensions, and findings which indicated that performance tended to be task-specific suggest caution in interpreting the results of individuals. The complexity and brittleness of some responses become highlighted if individual rather than group results are examined. Individuals are often poised between balancing tensions, and are sensitive to mood, test stimuli, and social context. It would be interesting to examine the responses of some outstanding individuals in more detail and to relate them to some general findings of the study.

The centroid factor analyses of cognitive and attitudinal variables were mainly of confirmatory value, and of some assistance in terms of construct validity. A principal components analysis which included all variables (cognitive and attitudinal) in one matrix might have a different range of relationships.

The reliability and validity of the open-ended tests and the attitude scale point up limitations in their use with any particular individual, although they are satisfactory for small group and total sample analysis. Issues relating to reliability and validity, together with difficulties in establishing satisfactory interscorer reliability levels however, would seem to indicate that the test instruments could not be used by classroom teachers and should be limited to research purposes only.

Finally, in contrast to some studies on creativity where so-called creative thinking abilities are lauded and other thinking abilities devalued, it is necessary to be cautious about assigning predetermined values to high or low scorers on fluency and flexibility of thinking. It would be false to attribute only positive connotations to high scorers on fluency or flexibility of thinking and contrasting negative connotations to low scorers. The relative value of high or low scores on fluency and flexibility depends on many other variables and purposes. There are times when it may be considered appropriate to be non-fluent and non-flexible.

While the task requirements of the present study have sought maximal rather than minimal attainment levels along continua of response-fluency and flexibility, many other studies have sought quite the opposite. In fact the majority of studies in science attainment and thinking have probably sought a single correct response. Neither emphasis can claim any exclusively privileged position. Even in the wider area of creativity, Shouksmith (1970) points out that a balance is required: "The search for creativity will probably be most fruitful among individuals with a high level of general ability, a high degree of associative fluency and among those who also possess certain cognitive styles, not all of which are flexible and open. The associative fluency must involve 'meaningful' associations and not merely facility with words" (p.214). And Bray (1967) also warns that the associational flow of ideas may merely be one precondition among many factors that determines creativity.

FURTHER RESEARCH

Perhaps the foremost general research need is for a more detailed analysis of specific variables related to fluency and flexibility of thinking. The nature and relative interplay of these variables could be examined through control and experimental groups with sample size equivalent to three or four classrooms. Q-technique and analysis of covariance could be supportive ways of examining the data. In particular the following lines of endeavour seem worthwhile:

(i) Variation to the form of presentation of, and response to stimuli including comparisons of oral and written methods. Experiments employing the current tests with tape recordings of responses have already indicated some interesting differences from the present method of presenting and recording stimuli.

(ii) Variation in the presentation of stimuli along continua of so-called game-like and test-like conditions. (Kogan and Morgan, 1969; Nicholls, 1971; Vernon, 1971). Variation in task demand from a low key neutral position to a high key test situation will affect subjects in different ways so that the performance of some individuals may increase while others decrease, under the same conditions. Some general conclusions from contemporary research in this area are conflicting, partly because of different definition and use of the terms 'game-like' and 'test-like', and partly because of different tests and sample composition. Further experimentation with the style of presenting stimulus material, e.g. completely programmed in written form or orally by a self-chosen classmate, or orally by an outside adminstrator simulating certain moods (Hudson, 1968), could also produce some interesting findings in terms of response fluency, flexibility and originality.

(iii) Variation in test instructions and their influence on fluency and flexibility of response. Experimentation with the following three sets of instructions could prove interesting:

(a) 'Give as many responses as possible to ....'

(b) 'Give as many different responses as possible to..'

(c) 'Give as many different responses as possible to..'

(iv) Variation in the nature of the stimulus so that it moves toward, or away from, topics of personal interest in the science domain. It could be hypothesized that response fluency and flexibility will increase as the stimulus moves toward strength of interest and ability.

(v) Variation in time limits for responding, including upper ends of time limits for any one administration. Repeated administrations over several weeks could also be considered with corresponding increase in reliability of scores.

(vi) Variation in population sample including a wider range of interest and ability level. Such groups could afford the basis of comparisons for the correlational behaviour of specific variables.

(vii) Variation in scoring taxonomies. Different taxonomies may yield useful comparative information and ultimately lead to more reliable and valid scores.

(viii) The inclusion of a performance test, or at least a simulated performance test, in science.

Progress has already been made in this direction as part of the present study. Three tests: one in physics, one in chemistry and the other in biology have been designed. (See also Kinsbourne, 1968, who suggested designing "open-ended performance tests to elicit from scientists, less accustomed to random generation of words and sentences, a chain of non-verbal free associations" (p.462)).

(ix) More detailed analysis of personality traits associated with high/low scorers in fluency and flexibility. Such a study could move beyond self-report attitudinal traits to a more detailed examination of such dimensions as self-control and inhibition. From brief experimental use of the present tests in a personal interview situation the degree to which some subjects exercise tight control and censorship over their responses is striking. Such control appears to result in long pauses between responses and, of course, in a lowered fluency score. Such inhibition also seems to be related to a concern (almost obsession) for 'accuracy' of response - which has little relevance for open-ended situations - and to a loss of face and sense of embarrassment over suggesting a 'wrong' or 'inappropriate' response. The fear of 'making a mistake' seems to be so deep in some subjects that they remain fearful, even when it is virtually impossible to make a mistake. In this respect an index of carefulness or cognitive error derived from intelligence tests scores or other related scores, could be revealing correlate of fluency and flexibility of thinking. Humour and aggression are two other dimensions (not unrelated) that seem worthwhile studying in further depth.

(x) More detailed examination of certain teacher variables. The delineation of some aspects of teaching style including a taxonomy of questioning categories and sequences (Solomon, 1969; Washton, 1967), would seem to hold some relevance for studying pupil fluency and flexibility of response. It may be possible to employ audio and video tape recordings for the establishment of such taxonomies and then construct two or three contrasting taxonomies which could form the basis for studying the growth of fluency and flexibility of pupil response. Audio and video tape recordings might also give some insight into the behaviours and cues that teachers employ in forming judgments about, or assigning rating categories to, pupils who exhibit originality and other aspects of thinking. It would be necessary to obtain pre-test and post-test scores on many variables in such experiments.

(xi) Some evaluation by teachers of the quality of responses derived from fluency and flexibility of thinking. This is a most challenging area and fraught with many difficulties. However, teacher ratings on one or two categories, e.g. degree of abstraction of pupil response or level of concepts, could prove an interesting area of correlation with scores on fluency and flexibility.

Apart from the above suggestions which are premised on the general idea of a more detailed analysis of the interplay of specific variables, the other major research emphasis is concerned with a developmental or longitudinal approach. In this respect the following lines of endeavour seem profitable:

(i) An investigation of fluency and flexibility of thinking in the verbal responses of pre-school children. A preliminary study employing some of Wallach and Kogan's (1965) test stimuli with three and four year olds has been undertaken by the investigator. A striking range of responses and behaviour suggest that this is a most profitable and formative stage for the study of conditions that encourage or inhibit the growth of fluency and flexibility of thinking. The language development of the mother (and siblings) is likely to be a significant factor. But so too, are patterns of discipline in the home - balance of praise over punishment, encouragement over restraint, and independence over dependence.[+]

(ii) Longitudinal studies which follow pupils through several years of teaching. While it would be difficult to assign precise reasons for any changes in scores that might occur, evidence from outcomes of research suggested above might at least assist in throwing some light on the issues. Over and above this, an analysis of the extent of the changes (or stability of measures) would be worthwhile in itself.

(iii) Finally, in terms of longitudinal study and predictive validity, a study of the careers of high and low scorers in the present investigation. The students in the present sample will now be 17, 18 and 19 years of age. The intention is to follow up the New Zealand group in terms of:

(i) the final grades and subjects at secondary school;

(ii) the relationship between original scores on the open-ended thinking tests and current re-test scores, and/or between original scores and scores derived from some other parallel series of questions;

(iii) the university subjects taken and proposed - particularly whether there is any relationship

---

[+] In this connection some relevant questions for a parent interview questionnaire have been devised by the investigator.

between extreme scorers on fluency and flexibility and the type of course enrolled for, whether, for example, the highly fluent and flexible scorers in science show any indication of moving toward the biological and social sciences; and

(iv) the relationship between scores and academic attainment - bearing in mind the tendency for regression to the mean.

In conclusion, the investigation has, as anticipated, thrown up more problems for research than were originally formulated. The network of connections between variables is intricate. But it is precisely because of this that further research is enticing.

APPENDIX A

Taxonomies for Scoring the Remaining Responses for Test 2, 'TIME'; Test 3, 'ACID', 'OXYGEN', 'MAGNET' 'WHEEL'

Test 2. Guide Lines for Scoring Responses to 'TIME'

1A Measurement methods
e.g. clock
watch
sundial
candle clock
time machine

1B Measurement units
e.g. day
week
month
year
decade
purely a made-up measurement
light year

1C International time
e.g. different times in different countries
international date line
Greenwich

1D Astronomical time
e.g. time in space
earth orbits

2A Derivation of word
e.g. Greek - Chronos
contains four letters
tempus

3A Set times as sayings
e.g. spring time
wartime
daytime

3B Proverbs and sayings
e.g. time waits for no man
time is money
in no time
up with the times
time and tide
time flies

3C Specific hours
e.g. 12 noon lunch time

APPENDIX A contd

4A **Proper names**

e.g. 'Time Magazine'
'Time Tunnel'
'Time Square'

5A **Travel and direction**

e.g. travelling in time
can't go back or forwards in time

5B **Associated with race**

e.g. race against time

6A **Personification**

e.g. Old Father Time
old man and scythe

7A **Punctuality**

e.g. on time
timetables

8A **School timetable**

e.g. 40 minute period
interval in 10 minutes

9A **Age**

e.g. life time
age span
old in time

9B **History**

e.g. gives us history lessons
History - dates to learn
Christ born year 0
AD. BC.
Ice age

9C **Point in time**

e.g. past,
future
late
early

10A **Psychological**

e.g. loneliness
lost
time drags
very important to life
it almost controls our lives

APPENDIX A contd

11A Metaphysical

e.g. intangible
invisible
relative
infinite
never ends
abstract

12A Dimensional

e.g. fourth dimension
fifth dimension

Test 3. Guide Lines for Scoring Responses to
'ACID'
'OXYGEN'
'MAGNET'
'WHEEL'

ACID

1. Chemical changes and processes

e.g. neutralising alkalines, bases
neutralising a base to form a salt + water
for making salts
combines with bases for chemical reactions
liberates chlorine from bleaching powder

2 Industrial chemistry

3 Torture, violence, death

e.g. burning skin
acid thrown in face of policeman
disposing of bodies
getting rid of an enemy

4 Batteries

e.g. activator in battery
acid and water - electrical charge
to conduct electricity in

5A Burning

e.g. burning clothes
burning metals

5B Destructor

e.g. as a destructor instead of fire
as a weed killer

APPENDIX A contd

5C Cleaning
e.g. cleaning metals

5D Cutting
e.g. cutting glass
etching glass

5E Decomposing, decaying

6 Basic characteristics and definitions
e.g. disassociates into positive hydrogen ions
what is its usual chemical and atomic structure
turns litmus paper red
what does man need them for?
why do they have different colours?

7 Medicinal
e.g. used in some medicines
digestive promoter

8 Foods and beverages
e.g. in fizzy drinks (carbonic acid)
flavouring foods
citric acid
vinegar

9 As an additive
e.g. used in antifreeze

10 Explosives
e.g. picric acid - trinitrophenol

11 For experimenting with

OXYGEN

1 For breathing, for life support
e.g. we breathe it
forms about 20% of the air we breathe
plants breathe it
given off by plants
supports life in outer space
respiration
deep-sea diving
oxygen tents in hospitals
high altitude breathing
as a stimulant for athletes
used in our blood stream
in water for fish

APPENDIX A contd

2 In air, water

3A Industrial cutting equipment
e.g. oxy-acetylene cutter
welding

3B Compressed
e.g. compressing for tools
compressed for submarines

4A As basis for combining with other compounds and elements
e.g. combines with hydrogen to make water
for making water, for making air
helps make heavy water
making compounds, bases
exploding with two parts of hydrogen
creating new materials
making radicals
helps form hydroxides
used in many coumpounds e.g. lead oxide

4B Rusting (specifically, oxidation, fermentation, photosynthesis)

5 Common use in laboratory, used in chemistry experiments, used in chemical reactions, to study

6A For blowing things up
e.g. blowing up balloons
blowing up tyres

6B Explosives
e.g. used in TNT
for shattering bananas!

7 Named as a component of gases
e.g. carbon dioxide

8 Source of power, as a propellant for rockets

9 Combustion, burning
e.g. to keep fires burning with
supports burning
used for starting fires

10 Liquid oxygen
e.g. liquid oxygen can be used for deep freezing

11 Physiological/humour
e.g. for taking away from someone I know!

12 For spelling tests

APPENDIX A contd

MAGNET

1A Electrical motor equipment
e.g. electric motors
to manufacture electricity
dynamos
generators
electro magnets

1B Electrical radio and telegraph
e.g. telephones
radio
T.V.
loudspeakers
electric bells

2A Definitions and explanations
e.g. a metal object which attracts certain objects
can reverse poles on to other magnets

2B Testing the validity of metals

3A Attracting metals
e.g. picking up iron filings, scrap iron, clips, pins

3B As a method of joining metals

3C Looking for something lost
e.g. finding lost pins
finding things in nooks and crannies

4 In compasses

5 Experiments
e.g. seeing what things they attract and what they don't

6 Wrecking
e.g. wreck colour T.V. cathode ray tubes with
wreck watches
wreck helicopters
wreck a tape

7 For particle acceleration
e.g. alter the course of electron beams
deflect course of rapidly moving metal objects

8 Attaching things to a wall or notice board with, hanging things on the refrigerator with

APPENDIX A contd

9A Amusement
e.g. in toys
in games
as a plaything
keeping a child busy with
attracting boys with

9B Cheating with
e.g. fixing roulette
cheating at bowling with

10 Household uses
e.g. magnetic door catches
soap holders
keeps lids on tins out of the way after they have been opened

11 Making other magnets with

12 In outer space
e.g. in boots for spacemen

13 Psychological
e.g. an attractive likeable person
a business "magnet"
shipping "magnate"

14 Medical, surgical
e.g. removing metal splinters from eye
removing pins from stomach

WHEEL

1A Basis for transport, for moving objects, to roll or move heavy objects on (i.e. generic)

1B Specific naming of vehicles
e.g. cars
trucks
buses
bicycles
planes
roller skates

1C Wheelbarrow

1D Steering wheel

1E **Category based on number of wheels**
e.g. two wheeler
three wheeler
four wheeler
eight wheeler

1F **Transport associated with water**
e.g. paddle wheel
propellor

2A **Machinery (generic)**

2B **Gears, cogs, pulleys, flywheel, grinding wheel, conveyor belts, escalators**

2C **Tools**
e.g. electric saw
wallpaper roller
glass cutter
cutting wheel
pizza cutter

3 **Clocks, watches, balance wheel of a clock**

4 **Measuring and recording devices**
e.g. map measurer
ground distance measurer

5A **Sport and entertainment**
e.g. ferris wheel
roulette
for gambling
fireworks - catherine wheel
tops
making a swing out of
as a monocycle for a clown
winding things on to
discus throwing
as a target
as a shield
for rolling down hill
hulahoop
for playing quoits
making patterns with
drawing around
potter's wheel

6 **Torture**
e.g. wheel rack
torture in the latin club
for punishing Christians
spiked wheel for running people over with
break over someone's head

7A **As a trade mark, name of a motor magazine (Wheels)**

APPENDIX A contd

7B As a symbol in the Greek alphabet

8 Phrases
e.g. birds wheel about in the sky

9 Gyroscope

10 Providing work for man
e.g. garages

11 For coins and money

12 Such specifics as
e.g. for a locomotive turnabout
(drawing supplied 551)
for a roundabout
as a protection for trees when stacked upon each other
as a seat, as a table
as a man hole

APPENDIX B (See below)

No.	Test 1 Fluency	Test 1 Flexibility	Test 2 Fluency	Test 2 Flexibility	Test 3 Fluency	Test 3 Flexibility
165	33 32 33	14 17 17	19 18 19	18 14 13	12 11 12	10 9 9
			13 14 13	11 8 8		
			11 11 11	6 4 6	16 15 15	13 12 13
180	12 12 12	5 6 6	8 10 10	4 6 4	3 3 3	2 2 2
			6 8 7	3 4 3		
			4 5 5	3 4 3	5 6 5	3 4 4
195	36 34 35	26 27 16	12 13 13	11 12 10	4 5 5	3 2 3
			10 11 11	8 10 10		
			9 11 10	7 10 8	9 9 9	6 7 6
210	14 14 14	10 8 7	3 3 3	2 2 2	1 1 1	0 0 0
			2 4 4	1 2 2		
			3 3 3	2 2 2	3 3 3	2 2 2
225	18 18 18	9 9 7	10 11 10	8 9 8	3 4 4	2 3 3
			7 9 11	5 5 8		
			9 6 6	8 4 4	5 6 6	4 4 4
240	26 25 26	20 22 21	9 9 9	7 7 7	8 8 8	7 6 7
			6 7 8	5 5 4		
			8 9 10	7 3 3	14 15 14	9 7 8
245	30 30 30	20 14 17	16 17 16	11 10 11	13 13 13	12 9 12
			10 15 16	8 8 7		
			16 15 16	6 5 6	15 15 15	14 12 14
260	17 17 17	8 7 7	5 7 7	3 4 4	4 4 4	3 2 4
			4 5 4	3 3 3		
			3 4 4	2 3 2	7 7 7	2 3 2
275	42 39 42	25 17 22	11 10 11	8 6 8	10 11 11	9 8 9
			10 15 17	7 6 6		
			11 10 10	9 6 8	14 15 14	11 10 9
290	41 41 41	18 24 20	25 19 25	12 10 13	8 8 8	7 6 7
			29 14 29	17 17 16		
			16 14 16	11 9 9	7 6 6	6 5 5
297	30 30 30	18 22 17	13 13 13	6 7 7	3 3 3	2 2 2
			6 6 6	3 3 3		
			10 10 10	9 7 9	4 3 4	3 2 3
304	39 39 39	22 26 25	19 17 19	17 12 17	10 8 8	9 6 6
			9 8 9	5 4 6		
			8 7 7	4 5 4	4 4 4	3 3 3

Interscorer Reliability Data: Raw Scores for Three Tests by Three Scorers

	Test 1		Test 2		Test 3	
No.	Fluency	Flexibility	Fluency	Flexibility	Fluency	Flexibility
311	52 52 52	34 37 30	12 12 12	8 9 9	10 10 10	8 8 8
			8 10 8	4 4 4		
			5 6 6	4 3 4	11 11 9	8 5 7
318	19 19 19	12 15 12	6 6 6	3 4 4	4 4 4	3 3 3
			9 9 9	6 6 6		
			4 4 5	3 3 4	4 4 4	3 3 3
325	27 27 27	11 11 10	7 10 8	5 4 5	4 3 3	2 2 2
			16 16 15	4 4 4		
			23 18 21	12 7 9	8 11 6	3 4 3
332	41 41 41	20 25 23	6 7 6	4 3 4	2 3 3	0 1 2
			13 14 14	3 3 2		
			8 6 8	4 3 3	3 3 3	2 2 2
339	28 28 28	18 18 16	5 5 5	4 4 4	12 11 11	10 8 10
			7 7 7	3 4 3		
			4 3 4	3 2 3	11 11 11	6 6 6
340	28 28 28	16 18 14	12 12 12	5 7 6	7 6 6	5 3 2
			13 12 14	4 3 3		
			11 7 10	6 4 6	9 9 9	5 5 5
353	22 22 22	15 15 12	9 8 9	8 7 7	12 12 12	9 10 10
			5 4 5	3 3 3		
			6 7 6	2 4 2	12 12 12	10 8 9
360	17 17 17	6 7 7	9 9 9	7 8 8	4 4 4	3 3 3
			9 9 9	8 5 8		
			9 8 9	5 7 7	8 8 8	7 7 7
367	13 13 13	10 11 10	8 10 8	7 8 7	5 6 6	4 3 4
			7 5 7	6 4 5		
			7 6 7	4 5 4	8 8 8	7 7 7
374	22 22 22	11 13 12	12 13 13	7 7 7	9 9 9	8 8 8
			10 9 9	5 5 5		
			10 11 11	4 7 4	8 8 8	7 4 4
382	36 36 36	16 16 16	15 13 14	12 10 11	5 5 5	4 4 3
			13 14 14	7 7 7		
			9 9 9	5 6 7	5 6 6	4 4 4
388	15 15 15	10 13 13	9 6 8	6 4 5	5 5 5	3 4 3
			4 4 4	3 3 3		
			4 4 4	2 2 3	6 7 7	3 3 3

APPENDIX B contd

	Test 1		Test 2		Test 3	
No.	Fluency	Flexibility	Fluency	Flexibility	Fluency	Flexibility
394	20 20 20	9 12 12	7 9 9	6 5 6	2 3 3	1 2 1
			14 15 15	4 5 5		
			9 7 8	6 5 5	8 8 8	5 6 6
402	35 35 35	16 22 18	14 16 16	7 5 9	9 8 11	7 7 6
			6 7 7	4 4 4		
			12 11 12	7 9 8	13 13 14	9 7 8
409	26 26 26	11 12 12	14 8 14	13 7 12	12 12 12	9 8 9
			19 17 18	7 6 7		
			18 15 18	6 10 7	13 12 12	11 10 11
416	21 21 21	15 17 16	15 12 15	6 4 7	7 7 7	6 4 5
			17 16 16	2 2 2		
			13 11 12	6 6 5	13 14 14	5 7 8
423	23 23 23	17 19 17	11 10 11	9 9 9	8 8 8	7 7 7
			11 8 10	5 5 4		
			11 10 11	6 5 5	11 10 10	9 9 9
427	43 43 43	16 20 14	26 23 24	17 11 20	12 9 11	6 4 8
			25 24 23	10 7 10		
			26 26 23	17 11 13	15 12 14	7 7 8
439	27 27 27	18 15 14	9 6 8	7 2 5	6 6 6	3 2 3
			6 6 6	3 2 2		
			2 1 2	0 0 0	8 6 6	4 4 5
451	36 36 36	17 18 17	8 8 9	6 5 6	6 6 6	4 3 4
			9 9 9	8 8 8	11 9 9	7 6 7
			9 10 8	8 8 7		
463	36 36 36	20 20 21	17 18 18	16 10 14	10 9 9	5 6 6
			17 16 16	11 11 11		
			14 14 14	10 8 7	8 6 6	5 4 5
475	32 32 32	14 11 12	7 8 8	6 5 6	9 8 8	6 3 6
			9 9 10	8 7 9		
			9 8 8	7 3 7	8 7 8	7 4 7
487	37 37 37	13 15 11	10 10 13	5 5 6	7 5 6	5 2 5
			9 9 9	7 6 6		
			11 8 9	4 4 6	8 6 6	3 3 5
499	30 30 30	12 17 11	12 8 10	5 4 4	12 7 7	6 5 6
			13 10 11	3 3 3		
			8 7 7	4 3 3	11 8 8	3 3 3
512	40 40 40	25 20 24	13 10 13	11 8 11	12 11 11	9 6 10
			9 9 9	8 7 8		
			17 12 13	7 7 7	12 10 11	3 3 5

APPENDIX B contd

No.	Test 1 Fluency	Test 1 Flexibility	Test 2 Fluency	Test 2 Flexibility	Test 3 Fluency	Test 3 Flexibility
524	50 50 50	23 17 27	18 16 18	10 8 12	16 14 14	10 9 12
			16 15 17	8 8 11		
			20 12 17	10 5 11	27 17 25	11 5 10
536	37 37 37	17 18 17	21 18 18	9 7 10	12 10 11	7 6 9
			11 11 11	7 5 7		
			16 13 14	6 5 6	11 11 10	7 6 7
548	38 38 38	19 22 18	16 15 15	14 10 13	12 11 11	9 7 9
			10 8 9	6 4 7		
			18 16 16	7 5 10	26 25 25	23 18 21
560	42 42 42	16 13 15	20 17 19	17 12 17	11 9 11	5 5 5
			20 15 20	13 11 13	13 10 10	6 6 7
			15 13 13	12 11 12		
572	44 44 44	18 23 19	18 16 16	12 11 13	15 14 15	10 10 13
			16 11 13	5 2 6		
			27 9 13	6 3 6	17 16 15	12 10 11
584	29 29 29	12 15 12	5 5 5	2 1 2	10 10 10	5 7 7
			13 13 13	6 5 7		
			13 13 13	8 8 7	9 9 9	0 1 2
596	26 26 26	16 20 15	11 9 9	7 3 6	11 6 9	5 3 5
			12 11 11	4 4 4		
			16 14 16	7 5 7	14 11 9	6 5 5
608	29 29 29	18 17 16	18 17 18	8 8 10	6 5 5	2 3 4
			17 15 14	2 2 2		
			12 11 11	7 4 7	10 10 9	5 5 3
620	34 34 34	14 13 12	17 11 13	12 5 11	9 8 5	6 7 6
			13 9 12	5 4 6		
			13 9 11	8 7 9	7 5 5	4 3 3
632	26 25 25	11 10 12	9 9 8	6 5 5	4 4 5	2 3 3
			12 11 10	5 5 5		
			7 7 8	5 4 6	6 5 5	4 4 4
644	41 41 41	16 17 15	11 11 11	8 8 9	6 6 6	4 5 5
			6 6 6	6 3 5		
			10 11 10	5 6 6	8 8 8	4 4 4
656	26 26 26	18 17 17	14 14 14	9 8 11	9 9 9	7 6 7
			11 11 11	5 5 5		
			11 11 11	9 9 9	7 7 7	6 5 6

APPENDIX C

Computer Information and Statistical Formulae

Computers used in the investigation

I.B.M. 16/20

Olivetti Programma 101

Statistical formulae

t tests (employing two-tailed tests of significance)

$$t = \frac{M_1 - M_2}{\sqrt{\frac{\Sigma x_1^2 + \Sigma x_2^2}{N_1 + N_2 - 2} \quad \frac{N_1 + N_2}{N_1 N_2}}} \quad \text{(Pooled variance)}$$

M_1/M_2 = means of two samples

$\Sigma x_1^2/\Sigma x_2^2$ = sums of squares (variance)

$N_1 N_2$ = number

$$t = \frac{M_1 - M_2}{\sqrt{\frac{S_1^2 + S_2^2}{N_1 \; N_2}}} \quad (S_1^2 \; S_2^2 = \text{separate variance})$$

T scores $T = Z(10 + 50)$ where $Z = \frac{X - u}{\sigma}$

APPENDIX C contd

Correlation Coefficients

(Pearson Product Moment with normalized T scores for fluency/flexibility, test/retest and original/re-mark)

$$r = \frac{\Sigma xy}{N \sigma x \sigma y}$$

x = deviation of any X score from the mean in test X

y = deviation of the corresponding Y score from the mean in test Y

Σxy = sum of all the products of deviations, each x deviation times its corresponding y deviation.

$\sigma x \sigma y$ = standard deviations of the distributions of X and Y scores.

Partial r

$$r_{12.3} = \frac{r_{12} - r_{13} r_{23}}{\sqrt{(1-r_{13}^2)(1-r_{23}^2)}}$$

Spearman-Brown Prophecy Formula

$$r_{tt} = \frac{2r_{hh}}{1+r_{hh}}$$

The problem of speededness on equivalent forms is not relevant in the present circumstances since this issue is premised on a right-wrong difficulty gradient.

APPENDIX C contd

Chi-Square (including Yates' correction for continuity factor)

$$= \frac{\left([f_o - f_e] - 0.5\right)^2}{f_e} + \frac{\left([f_o - f_e] - 0.5\right)^2}{f_e}$$

f_o = observed frequencies

f_e = expected frequencies

Fisher's Exact Probability Test

$$\frac{(a+b)!\ (c+d)!\ (a+c)!\ (b+d)!}{N!\ a!\ b!\ c!\ d!}$$

Factor Analysis

Centroid programme on I.B.M. 16/20 employing orthogonal axes, with Tucker's Phi for the termination of factors.

APPENDIX D

F tests for Homogeneity of Variance on 15 test Variables for Three Countries, England, U.S.A. and New Zealand.

Test Variable	Group	F ratio	Significance P<
AH5	ENG/USA	1.0115	N.S.
	ENG/NZ	1.1279	N.S.
	USA/NZ	1.1408	N.S.
Fl 1	ENG/USA	1.1950	N.S.
	ENG/NZ	1.1320	N.S.
	USA/NZ	1.3527	.05
Fx 1	ENG/USA	1.0046	N.S.
	ENG/NZ	1.2335	N.S.
	USA/NZ	1.2278	N.S.
Fl 2A	ENG/USA	1.4591	.05
	ENG/NZ	1.6574	.01
	USA/NZ	1.1359	N.S.
Fx 2A	ENG/USA	1.0906	N.S.
	ENG/NZ	1.0341	N.S.
	USA/NZ	1.1278	N.S.
Fl 2B	ENG/USA	1.1974	N.S.
	ENG/NZ	2.0492	.01
	USA/NZ	1.7113	.01
Fx 2B	ENG/USA	1.2818	N.S.
	ENG/NZ	1.8835	.01
	USA/NZ	1.4694	.01
Fl 3A	ENG/USA	1.0845	N.S.
	ENG/NZ	1.5075	.01
	USA/NZ	1.3900	.05
Fx 3A	ENG/USA	1.2008	N.S.
	ENG/NZ	1.0623	N.S.
	USA/NZ	1.1304	N.S.
Fl 3B	ENG/USA	1.5076	N.S.
	NG/NZ	1.2128	N.S.
	USA/NZ	1.2430	N.S.
Fx 3B	ENG/USA	1.0772	N.S.
	ENG/NZ	1.1392	N.S.
	USA/NZ	1.0576	N.S.
Fl 3C	ENG/USA	1.2573	N.S.
	ENG/NZ	1.1640	N.S.
	USA/NZ	1.4634	.05
Fx 3C	ENG/USA	1.9198	.01
	ENG/NZ	1.0953	N.S.
	USA/NZ	1.7528	.01
Fl 3D	ENG/USA	4.4172	.01
	ENG/NZ	1.6993	N.S.
	USA/NZ	7.5063	.01
Fx 3D	ENG/USA	2.0737	.05
	ENG/NZ	1.5056	N.S.
	USA/NZ	3.1221	.01

APPENDIX E

Distribution Curves for Raw Scores for Tests 2A, 2B, 3A, 3B, 3C and 3D. (Figures 1 - 12)

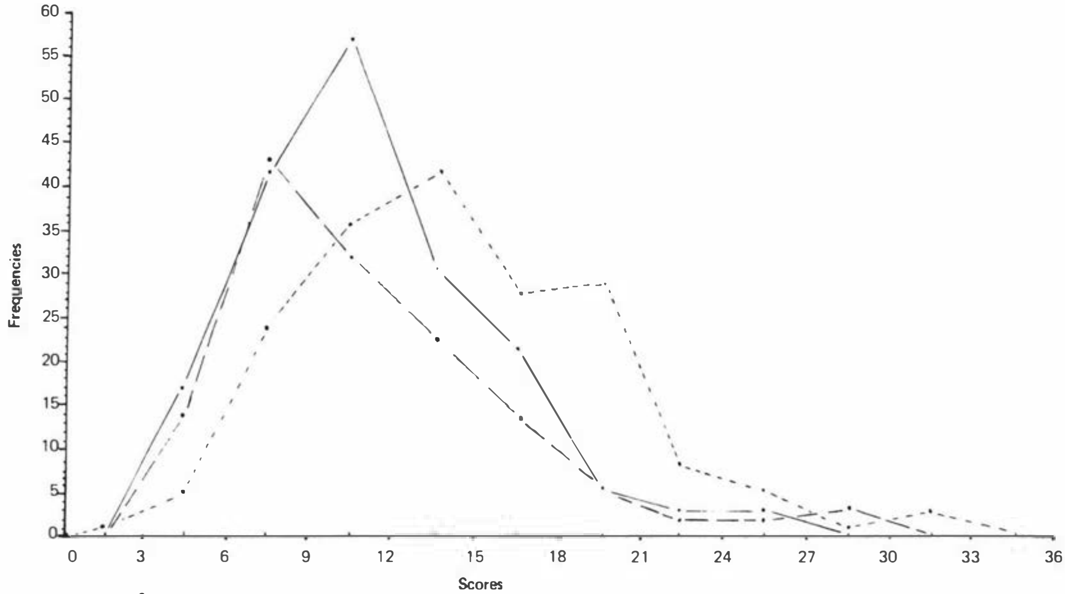


FIGURE 1 Frequency Polygon for the Distribution of Raw Scores, Test 2a, Fluency. England —(N = 181) USA – –(N = 140) NZ · · · · (N = 184)

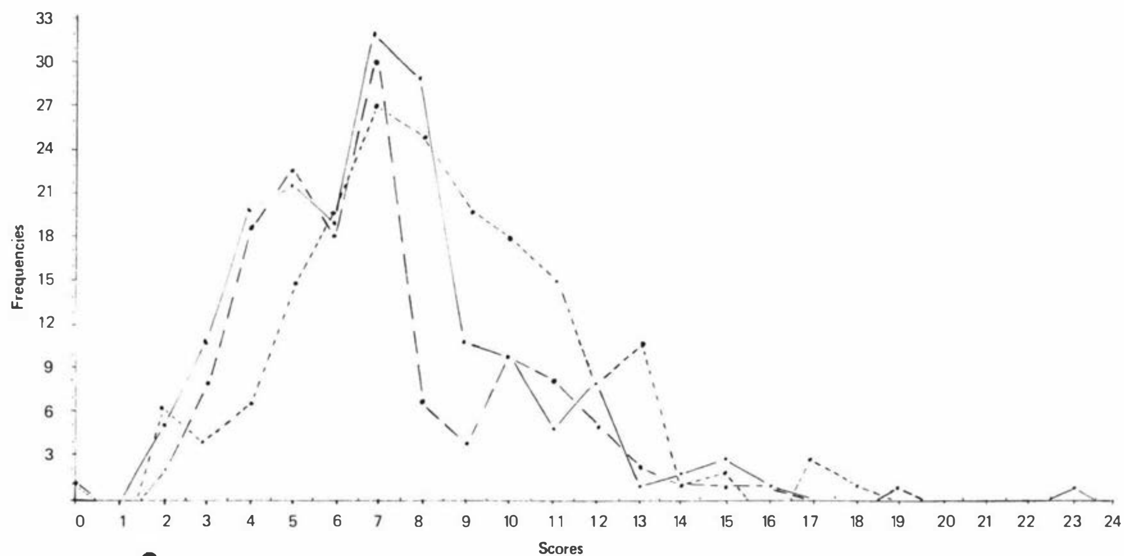


FIGURE 2 Frequency Polygon for the Distribution of Raw Scores, Test 2a, Flexibility. England — (N = 181) USA – –(N = 140) NZ· · · (N = 184)

APPENDIX B contd

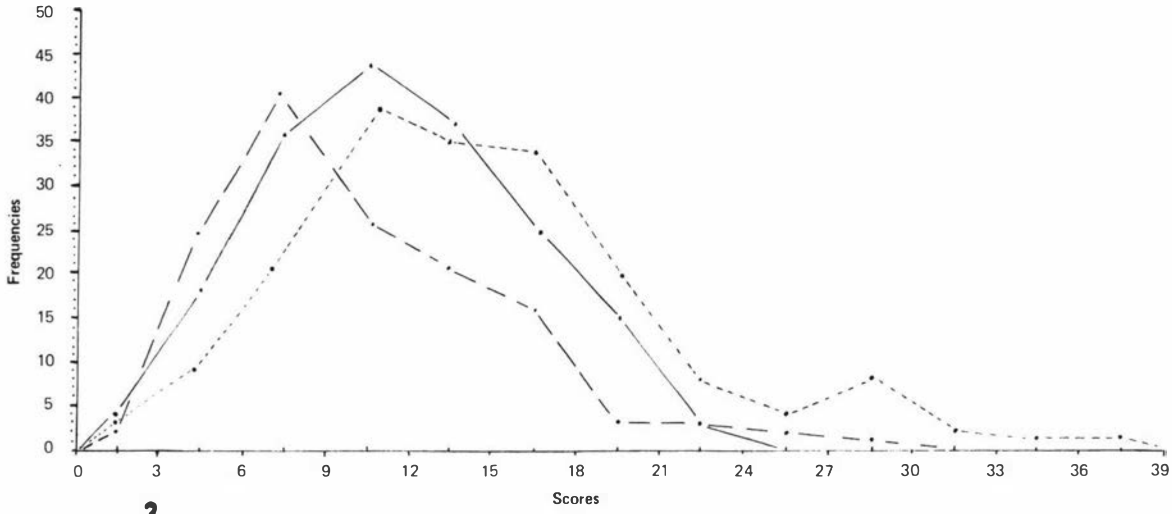


FIGURE 3 Frequency Polygon for the Distribution of Raw Scores, Test 2b, Fluency. England — (N = 182) USA – –(N = 140) NZ · · · (N = 184)

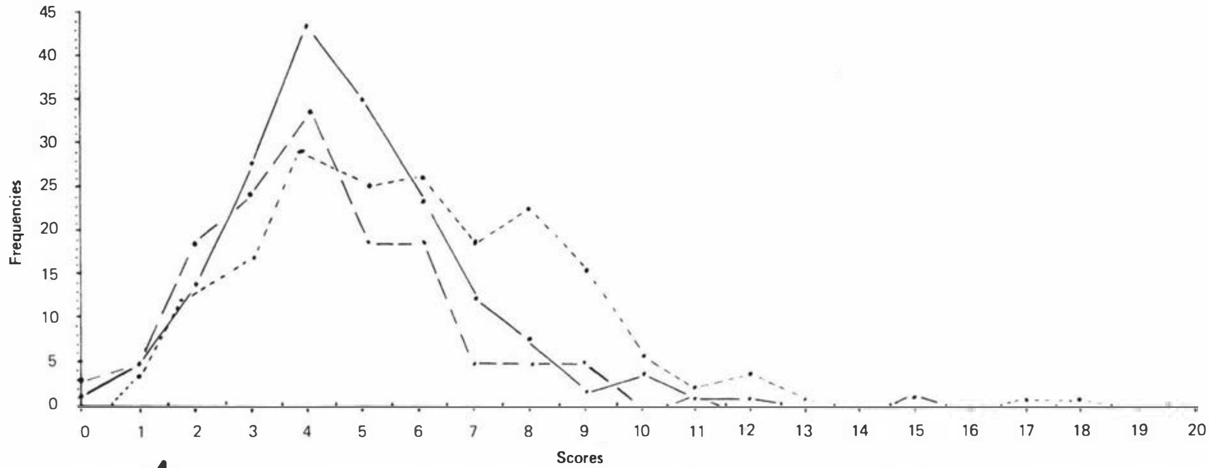


FIGURE 4 Frequency Polygon for the Distribution of Raw Scores, Test 2b, Flexibility. England — (N = 182) USA – –(N = 140) NZ· · · ·(N = 184)

APPENDIX E contd

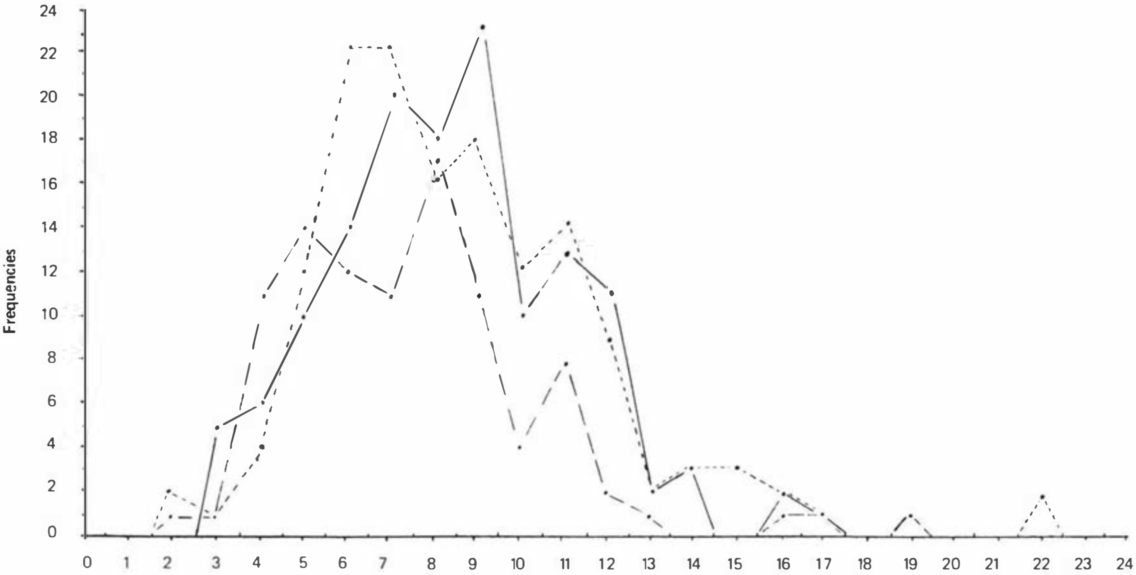


FIGURE 5 Frequency Polygon for the Distribution of Raw Scores, Test 3a, Fluency. England— (N = 138) USA– –(N = 96) NZ · · · (N = 145)

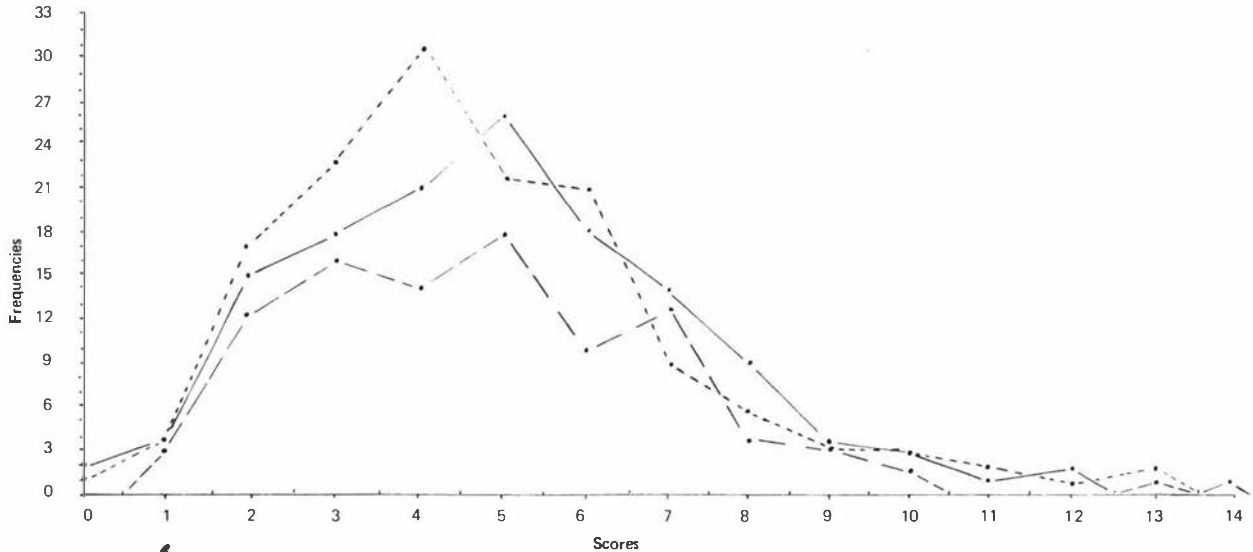


FIGURE 6 Frequency Polygon for the Distribution of Raw Scores, Test 3a, Flexibility. England— (N = 138) USA – –(N = 96) NZ . . .(N = 145)

APPENDIX B contd

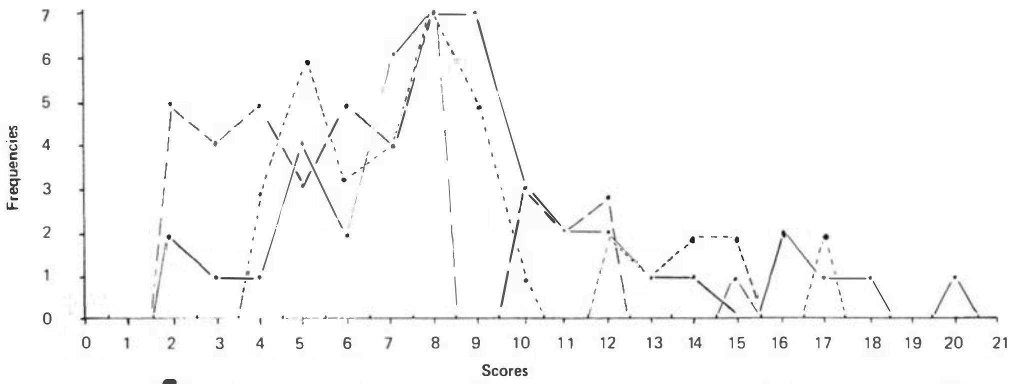


FIGURE 7 Frequency Polygon for the Distribution of Raw Scores, Test 3b, Fluency. England — (N = 44) USA – – (N = 42) NZ · · · (N = 38)

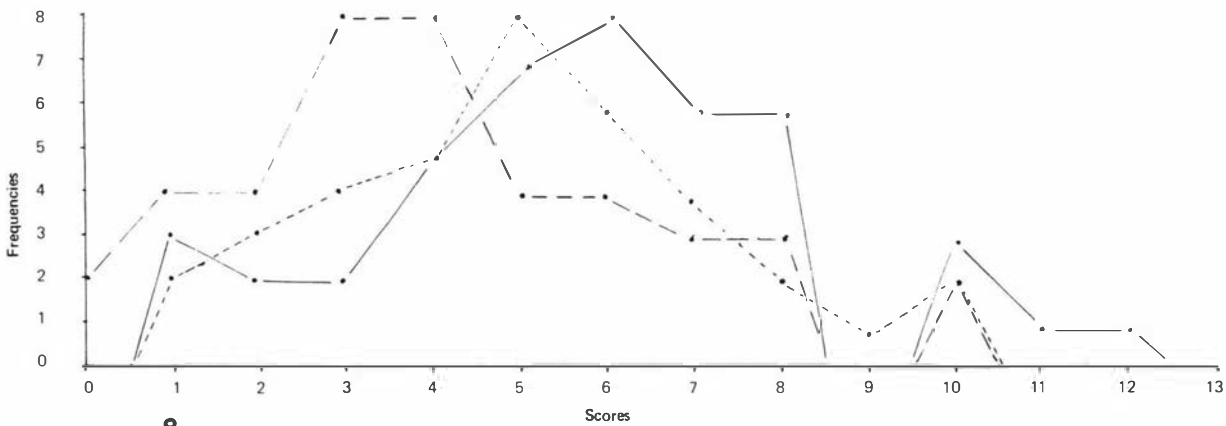


FIGURE 8 Frequency Polygon for the Distribution of Raw Scores, Test 3b, Flexibility. England — (N = 44) USA – – (N = 42) NZ · · · (N = 38)

APPENDIX B contd

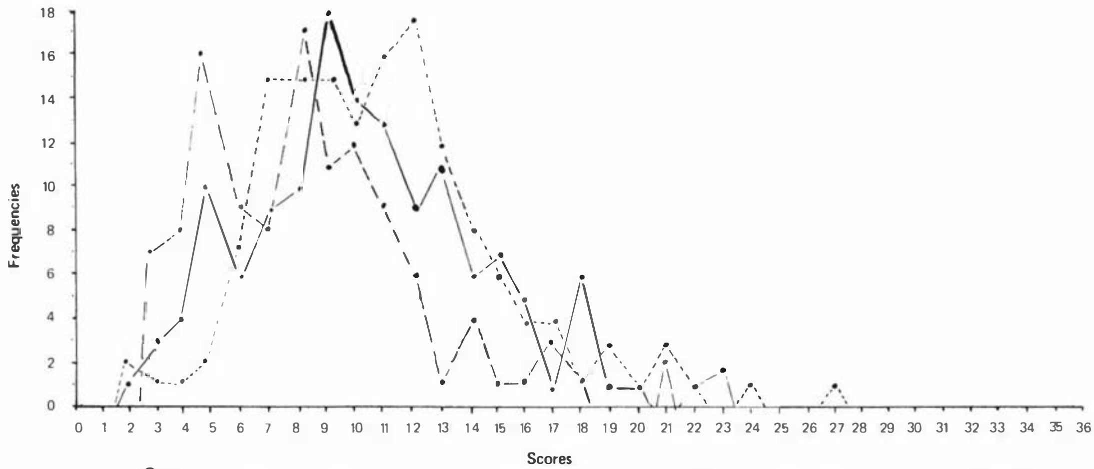


FIGURE 9 Frequency Polygon for the Distribution of Raw Scores, Test 3c, Fluency. England — (N = 138) USA – – (N = 116) NZ · · · (N = 151)

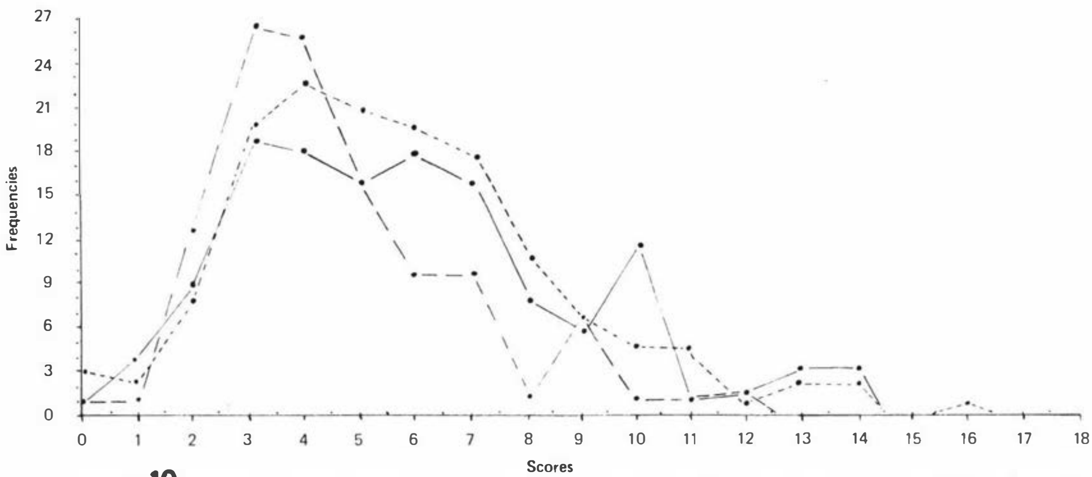


FIGURE 10 Frequency Polygon for the Distribution of Raw Scores, Test 3c, Flexibility. England — (N = 138) USA – – (N = 116) NZ · · · (N = 151)

APPENDIX E contd

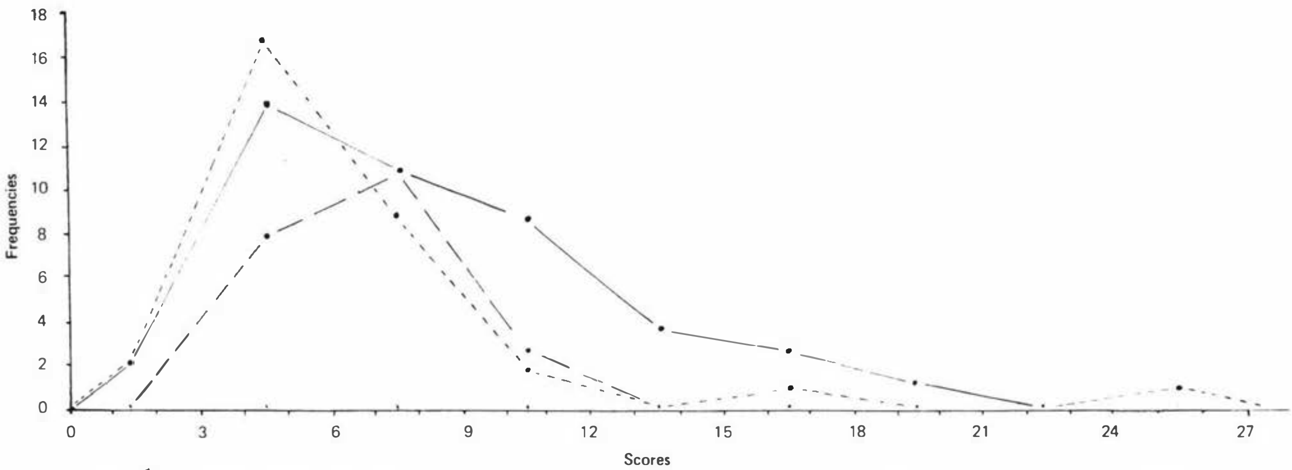


FIGURE 11 Frequency Polygon for the distribution of Raw Scores, Test 3d, Fluency. England— (N = 44) USA- - -(N = 22) NZ· · · ·(N = 33)

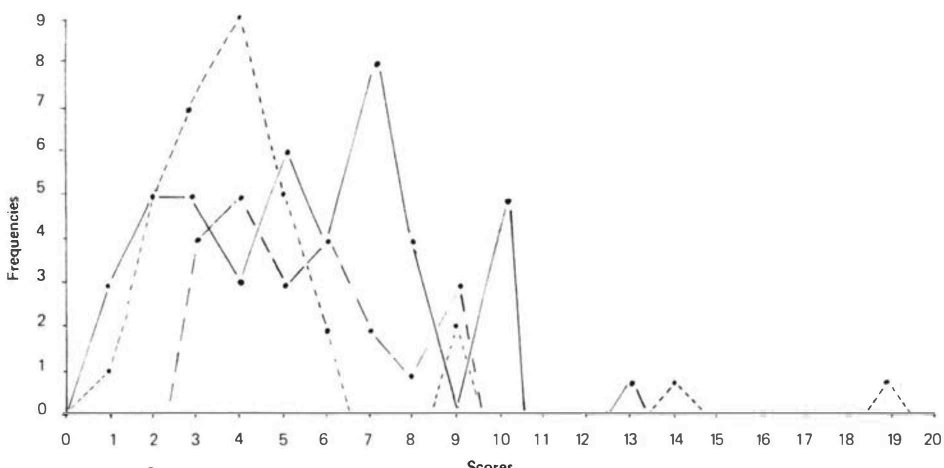


FIGURE 12 Frequency Polygon for the Distribution of Raw Scores, Test 3d, Flexibility. England — (N = 44) USA – – (N = 22) NZ· · · · (N = 33)

Between-Country Differences Analysis of Significance of Difference between Means (t-tests) on 45 sets of Scores for Test Variables (Intelligence = 1, Fluency = 7, Flexibility = 7) for Three Countries, England, USA, New Zealand.

Country	Test Variable	t	d.f.	Level of Significance
ENG-USA	AH5	7.4324	317	.001
ENG-NZ	"	5.2680	362	.001
USA-NZ	"	-2.3370	321	.02
ENG-USA	Fl 1	-.8508	320	N.S.
ENG-NZ	"	-10.0688	364	.001
USA-NZ	"	-8.1238	322	.001
ENG-USA	Fx 1	-.5543	320	N.S.
ENG-NZ	"	3.9511	364	.001
USA-NZ	"	-3.1250	322	.001
ENG-USA	Fl 2A	-.1877	319	N.S.
ENG-NZ	"	-6.1303	363	.001
USA-NZ	"	-5.1088	322	.001
ENG-USA	Fx 2A	.2630	319	N.S.
ENG-NZ	"	-3.6364	363	.001
USA-NZ	"	-3.9394	322	.001
ENG-USA	Fl 2B	2.2595	320	.02 to .05
ENG-NZ	"	-5.1230	364	.001
USA-NZ	"	-6.3559	322	.001
ENG-USA	Fx 2B	1.2458	320	N.S.
ENG-NZ	"	5.1201	364	.001
USA-NZ	"	-5.5287	322	.001
ENG-USA	Fl 3A	2.0866	232	.02 to .05
ENG-NZ	"	-.7951	281	N.S.
USA-NZ	"	-2.5516	239	.02
ENG-USA	Fx 3A	.6223	232	N.S.
ENG-NZ	"	1.0260	281	N.S.
USA-NZ	"	.3206	239	N.S.
ENG-USA	Fl 3B	2.9703	84	.01
ENG-NZ	"	.3447	80	N.S.
USA-NZ	"	2.6582	78	.01
ENG-USA	Fx 3B	2.8725	84	.01
ENG-NZ	"	1.0608	80	N.S.
USA-NZ	"	1.7851	78	.10, N.S.
ENG-USA	Fl 3C	4.2262	252	.001
ENG-NZ	"	-1.3303	287	.10, N.S.
USA-NZ	"	-5.2910	265	.001
ENG-USA	Fx 3C	3.9403	252	.001
ENG-NZ	"	.5443	287	N.S.
USA-NZ	"	-3.5440	265	.001
ENG-USA	Fl 3D	1.2548	64	.10, N.S.
ENG-NZ	"	.8983	75	N.S.
USA-NZ	"	-.1617	53	N.S.
ENG-USA	Fx 3D	.1427	64	N.S.
ENG-NZ	"	1.0705	75	N.S.
USA-NZ	"	.8207	53	N.S.

APPENDIX G

Sex Difference and Probabilities (Chi-Square) for Hi-Lo Scorers on Fluency and Flexibility, U.S.A., New Zealand

Country	Chi-squares and probabilities							
	Fl +1SD		Fl -1SD		Fx +1SD		Fx -1SD	
	χ^2	$p<$	χ^2	$p<$	χ^2	$p<$	χ^2	$p<$
USA	.043	-	.070	-	.001	-	.001	-
NZ	.219	-	.416	-	.002	-	.038	-

Appendix H

Observed Frequencies (o) and Expected Frequencies (e) by Classrooms and Schools for High (+1SD) and Low (-1SD) Scorers on Fluency and Flexibility, England, U.S.A., New Zealand

	School	Class Code	Class Total	School Sample Total	High Fluency Class O	High Fluency Class E	High Fluency School O	High Fluency School E	Low Fluency Class O	Low Fluency Class E	Low Fluency School O	Low Fluency School E	High Flex. Class O	High Flex. Class E	High Flex. School O	High Flex. School E	Low Flex. Class O	Low Flex. Class E	Low Flex. School O	Low Flex. School E
ENG	01	01	16	33	1	3.2	3	6.6+	7	3.0	13	6.2	2	3.0	4	6.2	4	2.2	6	4.5
		02	17		2	3.4			6	3.2			2	3.2			2	2.3		
	02	03	26	26	9	5.2	9	5.2	2	4.8	2	4.8	5	4.8	5	4.8	1	3.6	1	3.6
	03	04	28	57	2	5.6	6	11.4	7	5.2	13	10.6	1	5.2	7	10.6	2	3.8	8	7.7
		06	15		0	3.0			2	2.8			1	2.8			4	2.0		
		07	14		4	2.8			4	2.6			5	2.6			2	1.9		
	04	09	17	17	6	3.4	6	3.4	0	3.2	0	3.2	6	3.2	6	3.2	2	2.3	2	2.3
	05	10	23	49	8	4.6	13	9.8	2	4.3	6	9.1	6	4.3	12	9.1	3	3.1	8	6.7
		11	26		5	5.2			4	4.8			6	4.8			5	3.6		
		Σ =	182	182	37		37		34		34		34		34		25		25	
USA	06	12	22	22	4	3.0	4	3.0	3	3.0	3	3.0	6	3.0	6	3.0	4	3.1	4	3.1
	09	17	27	58	3	3.7	8	7.9	3	3.7	7	7.9	2	3.7	6	7.9	3	3.9	5	8.3
		18	31		5	4.2			4	4.2			4	4.2			2	4.4		
	08	16	28	28	2	3.8	2	3.8	4	3.8	4	3.8	2	3.8	2	3.8	4	4.0	4	4.0
	07	13	10	32	2	1.3	5	4.2	0	1.3	5	4.2	2	1.3	5	4.2	2	1.4	7	4.5
		14	10		0	1.3			5	1.3			0	1.3			5	1.4		
		15	12		3	1.6			0	1.6			3	1.6			0	1.7		
		Σ =	140	140	19		19		19		19		19		19		20		20	
NZ	10	19	31	31	2	5.0	2	5.0	2	3.7	2	3.7	5	5.1	5	5.1	3	4.5	3	4.5
	11	20	27	27	5	4.4	5	4.4	3	3.2	3	3.2	7	4.5	7	4.5	5	3.9	5	3.9
	12	21	20	20	6	3.3	6	3.3	2	2.4	2	2.4	4	3.3	4	3.3	1	2.9	1	2.0
	13	22	26	26	4	4.2	4	4.2	6	3.1	6	3.1	2	4.3	2	4.3	4	3.9	4	3.9
	14	23	21	21	4	3.4	4	3.4	1	2.5	1	2.5	4	3.5	4	3.5	4	3.0	4	3.0
	15	24	24	24	4	4.0	4	4.0	4	2.9	4	2.9	2	4.0	2	4.0	5	3.4	5	3.4
	16	25	35	35	5	5.7	5	5.7	4	4.2	4	4.2	7	5.8	7	5.8	5	5.0	5	5.0
		Σ =	184	184	30		30		22		22		31		31		27		27	

+ $6.6 = \frac{33}{182} \times \frac{x}{37}$

APPENDIX I

OBSERVED FREQUENCIES FOR TEACHER AND PEER RATING CATEGORIES* FOR HI-LO FLUENCY AND FLEXIBILITY GROUPS (England, U.S.A., N.Z.)

(*O = most original, A = highest ability, B+ = enjoyed teaching most, B- = enjoyed teaching least, W = hardest working)

			TEACHER RATING CATEGORIES																				PEER RATING CATEGORIES									
			O					A					B+					B-					W					O				
	SCH.	CLS.	TOT.	HFL.	LFL.	HFX.	LFX.	TOT.	HFL.	LFL.	HFX.	LFX.	TOT.	HFL.	LFL.	HFX.	LFX.	TOT.	HFL.	LFL.	HFX.	LFX.	TOT.	HFL.	LFL.	HFX.	LFX.	TOT.	HFL.	LFL.	HX.	LX.
ENG	01	01	3	0	0	1	0	4	1	2	2	1	1	0	1	0	1	1	0	0	0	0	19	5	6	9	3	15	1	8	5	3
		02	3	0	0	0	0	3	0	1	0	1	2	0	0	0	0	0	0	0	0	0	13	1	4	1	0	11	3	1	3	1
		ε	6	0	0	1	0	7	1	3	2	2	3	0	1	0	1	1	0	0	0	0	32	6	10	10	3	26	4	9	8	4
	02	03	4	2	0	1	0	4	3	0	2	0	4	3	0	1	0	4	0	0	1	0	40	16	1	3	0	37	26	1	10	0
	03	04	3	1	0	0	0	2	0	1	0	0	1	0	0	1	0	4	0	2	0	0	33	2	13	0	0	27	7	5	0	0
		06	3	0	0	1	0	3	0	0	0	0	4	0	0	1	1	0	0	0	0	1	29	0	0	1	3	34	0	0	5	0
		07	2	0	0	2	0	4	2	1	3	1	2	1	0	1	0	4	1	1	1	0	31	6	3	9	1	28	13	2	21	2
		ε	8	1	0	3	0	9	2	2	3	1	7	1	0	3	1	8	1	3	1	1	93	8	16	10	4	89	20	7	26	2
	04	09	4	1	0	1	0	4	1	0	0	0	4	1	0	1	0	0	0	0	0	0	27	5	0	8	0	24	5	0	4	0
	05	10	2	0	0	0	0	4	1	0	1	0	3	0	1	0	1	1	1	0	0	1										
		11	4	0	1	0	1	3	1	1	0	1	3	0	0	0	0	3	0	0	1	0	No. Inform.					No. Inform.				
		ε	6	0	1	0	1	7	2	1	1	1	6	0	1	0	1	4	1	0	1	1										
	ε	ENG.	28	5	1	6	1	31	9	6	8	4	24	5	2	5	3	17	2	3	3	2	192	37	27	31	7	176	55	17	48	6
USA	06	12	4	2	1	2	1	4	1	1	1	1	0	0	0	0	0	0	0	0	0	0	27	6	9	7	10	22	8	7	11	7
	09	17	2	0	0	0	0	4	1	0	1	0	4	1	1	1	0	3	0	0	0	0	46	34	0	27	0	27	14	0	13	0
		18	4	2	1	1	0	4	3	1	2	0	3	2	1	1	0	4	1	1	1	0	43	24	1	4	0	32	14	1	12	0
		ε	6	2	1	1	0	8	4	1	3	0	7	3	2	2	0	7	1	1	1	0	89	58	1	31	0	59	28	1	25	0
	08	16	4	0	0	0	0	4	1	0	0	0	4	0	0	0	0	4	0	1	0	1	37	8	0	0	0	36	0	1	0	1
	07	13	2	1	0	0	0	2	1	0	1	0	2	0	0	1	1	1	0	0	0	0	4	0	0	3	0	3	0	0	3	0
		14	0	0	0	0	0	1	0	0	0	0	2	0	0	0	1	1	0	1	0	1	12	0	0	0	6	8	0	0	0	4
		15	2	0	0	0	0	1	0	0	0	0	3	0	0	0	0	0	0	0	0	0	2	0	0	0	0	4	0	0	0	0
		ε	4	1	0	0	0	4	1	0	1	0	7	0	0	1	2	2	0	1	0	1	18	0	0	3	6	15	0	0	3	5
	ε	USA	18	5	2	3	1	20	7	2	5	1	18	3	2	3	2	13	1	3	1	2	171	72	10	41	16	132	36	9	39	12
NZ	10	19	4	1	0	1	0	4	1	0	1	0	4	1	0	1	1	1	0	0	0	0	56	21	0	36	2	49	2	1	4	2
	11	20	4	0	1	0	0	4	0	1	0	0	4	0	1	0	1	4	3	0	3	0	38	0	0	2	2	27	0	0	2	3
	12	21	4	2	0	0	1	4	3	0	1	1	0	0	0	0	0	0	0	0	0	0	35	20	0	3	3	29	7	4	6	0
	13	22	4	0	2	0	1	4	1	1	1	1	0	0	0	0	0	0	0	0	0	0	41	3	23	3	23	30	7	14	1	11
	14	23	4	2	0	1	0	5	3	0	2	1	5	3	0	2	0	4	0	0	0	1	32	21	0	19	6	27	19	2	19	0
	15	24	4	1	1	1	1	4	0	1	0	1	2	0	2	0	2	2	0	0	0	0	25	0	6	0	6	35	1	1	0	4
	16	25	3	1	0	2	0	3	0	0	0	1	3	0	0	1	0	4	0	0	1	0	42	0	4	2	23	37	7	7	14	7
	ε	NZ	27	7	4	5	0	28	8	3	5	5	18	4	3	4	4	15	3	0	4	1	269	65	33	65	65	234	43	29	46	27

Appendix J

Observed and Expected Frequencies for Teacher and Peer Rating Categories⁺ for Hi-Lo Fluency and Flexibility Groups (England, U.S.A., New Zealand)

⁺(O = most original, A = highest ability, B+ = enjoyed teaching most, B- = enjoyed teaching least, W = hardest working)

Country	Hi-Lo Fl, Fx	Teacher Rating Category								Peer Rating Category			
		O		A		B+		B-		W		O	
		o	e⁺⁺	o	e	o	e	o	e	o	e	o	e
ENG	H.Fl.	5	5.68	9	6.29	5	4.37	2	3.45	37	38.98	55	35.73
	L.Fl.	1	5.24	6	5.80	2	4.49	3	3.18	27	35.90	17	32.91
	H.Fx.	6	5.24	8	5.80	5	4.49	3	3.18	31	35.90	48	32.91
	L.Fx.	1	3.84	4	4.25	3	3.29	2	2.33	7	26.30	6	24.11
	Total Sample	28		31		24		17		192		176	
USA	H.Fl.	5	2.45	7	2.72	3	2.45	1	1.77	72	23.26	36	17.95
	L.Fl.	2	2.45	2	2.72	2	2.45	3	1.77	10	23.26	9	17.95
	H.Fx.	3	2.45	5	2.72	3	2.45	1	1.77	41	23.26	39	17.95
	L.Fx.	1	2.57	1	2.86	2	2.57	2	1.86	16	24.45	12	18.88
	Total Sample	18		20		18		13		171		132	
NZ	H.Fl.	7	4.40	8	4.56	4	2.93	3	2.44	65	43.85	43	38.14
	L.Fl.	4	3.24	3	3.36	3	2.16	0	1.80	33	32.28	29	28.08
	H.Fx.	5	4.54	5	4.70	4	3.02	4	2.52	65	45.19	46	39.31
	L.Fx.	3	3.97	5	4.12	4	2.65	1	2.21	65	39.54	27	34.40
	Total Sample	27		28		18		15		269		234	

⁺⁺e = proportion of H.Fl/L.Fl/H.Fx/L.Fx multiplied by sample on each category (O,A,B+,B-,W,O.)

E.G. "O" = H.Fl $= \frac{37}{182} \times 28$ "A" = H.Fl $= \frac{37}{182} \times 31$

L.Fl $= \frac{34}{182} \times 28$ L.Fl $= \frac{34}{182} \times 31$

APPENDIX K

Means, Standard Deviations and t-test Results for Retest Data. N = 74 (New Zealand) d.f. = 146

Var.	Mean	Std. Devn.	Test Group	T-Ratio	p<
Fl 1	36.6 38.5	7.934 9.386	Original Retest	-1.3491	.18
Fx 1	15.4 14.8	5.789 4.888	" "	.7314	.53
Fl 2M	14.7 14.1	5.137 4.352	" "	.8316	.59
Fx 2M	7.0 6.5	2.436 2.374	" "	1.2217	.22
Fl 3M	9.3 9.0	3.199 3.472	" "	.5501	.59
Fx 3M	5.3 5.0	2.537 5.588	" "	.3950	.70
Fl M	16.9 17.0	3.629 3.940	" "	-.0086	.99
Fx M	8.0 7.6	2.180 2.898	" "	1.0313	.30

APPENDIX L

Means, Standard Deviations, and t-test Results for Males (N = 40) and Females (N = 34) on Retest. (New Zealand) d.f. = 72

Var.	Mean	Std. Devn.	Test Group	T-Ratio	p
Fl 1	35.8	7.652	Male	-.8651	.60
	37.5	8.272	Female		
Fx 1	14.2	4.790	"	-1.9774	.05
	16.9	6.496	"		
Fl 2M	15.0	4.986	"	.3980	.69
	14.5	5.296	"		
Fx 2M	7.4	2.348	"	1.3900	.17
	6.6	2.469	"		
Fl 3M	9.3	2.977	"	-.1564	.87
	9.4	3.441	"		
Fx 3M	5.1	2.426	"	-.7678	.55
	5.5	2.639	"		
Fl M	16.9	3.728	"	-.2087	.83
	17.0	3.506	"		
Fx M	7.8	2.083	"	-.7686	.55
	8.2	2.270	"		

BIBLIOGRAPHY

ABNEY, C. (1970). 'A comparative study of creative thinking ability in three student groups at the University of Arkansas as measured by the Remote Associates Test.' Dissertation Abstracts Vol. 30, No. 7A, 2717-A

ANDERSON, R. (1968). 'Divergent thinking within superior, average and retarded subjects.' Dissertation Abstracts. Vol. 29, No. 4A, 1126-27.

ARVIDSON, G.L. (1970). Private communication.

BARRON, F. (1963). Creativity and Psychological Health. Princeton, New Jersey. D. Van Nostrand Company, Inc.

BARRON, F. (1969). Creative Person and Creative Process. U.S.A. Holt, Rinehart and Winston, Inc.

BENTLEY, J. (1966). 'Creativity and academic achievement.' The Journal of Educational Research. Vol. 59, No. 6, 269-272.

BERLYNE, D.E. (1965). Structure and Direction in Thinking. U.S.A. John Wiley & Sons, Inc.

BOGEN, J.E. and BOGEN, G.M. (1969). 'The other side of the brain: III. The corpus callosum and creativity.' Bulletin of the Los Angeles Neurological Societies. Vol. 34, No. 4, 191-220. (See Psychological Abstracts. Vol. 44, No. 6, 1970).

BRADFIELD, R. (1969). 'Divergent and convergent thinking: achievement motivated processes.' Dissertation Abstracts. Vol. 29, No. 12A, 4272.

BRAY, D.H. (1967). 'Reflections on creativity research.' Delta. No. 1, 1967. 43-47

BRODY, B.R. (1970). 'The effects of creativity and intelligence on teacher ratings.' The British Journal of Educational Psychology. Vol. 40, Part 3, 342-348.

BRUCE, D. and CROWLEY, J.J. (1970). 'Acoustic similarity effects on retrieval from secondary memory.' Journal of Verbal Learning and Verbal Behavior. Vol. 9, No. 2, 190-196.

BRUNER, J.S., GOODNOW, J.J. and AUSTIN, G.A. (1956). (Reprint 1961). A Study of Thinking. New York. John Wiley & Sons, Inc.

BRUNER, J.S. (1964). On Knowing - essays for the left hand. Massachusetts. Belknap Press of Harvard University.

CALLAWAY, W.R. (1968). 'A study of the relationship of a holistic conception of creativity to int lligence.' Dissertation Abstracts. Vol. 29, No. 2A. 475.

CARLIER, M. (1970). 'Une des modalites de la pensee divergente: la flexibilite et ses determinants personnels.' Revue Internationale de Psychologie Appliquee. Vol. 19, No. 1, 83-93.

CASHDAN, S. and WELSH, G.S. (1966). 'Personality correlates of creative potential in talented high school students.' Journal of Personality. Vol. 34, 445-55.

CATTELL, R.B. (1962). Manual for Forms A and B Sixteen Personality Factor Questionnaire. I.P.A.T.

CATTELL, R.B. (1964). 'Beyond validity.' Journal of Experimental Education. Vol. 33, No. 2, Winter 1964.

CATTELL, R.B. (1967). Objective Personality and Motivation Tests. University of Illinois Press.

CHRISTIE, T. (1969). 'Contrary imaginations at Keele: an alternative interpretation.' Universities Quarterly. Vol. 24, No. 1, 91-92.

CICIRELLI, V.G. (1967). 'Sibling constellation, creativity, I.Q., and academic achievement.' Child Development. Vol. 38, 481-90.

COHEN, D. (1968). 'Opportunities for creativity.' The Australian Science Teachers Journal. Vol. 14, No. 2, 6-19.

COONS, J. (1969). 'A cross-cultural study of sex differences in the development of selected creative thinking abilities.' Dissertation Abstracts. Vol. 29, No. 12B, 4828-29.

CRAMER, P. (1968). Word Association. New York. Academic Press.

CRAVER, A. (1969). 'Convergent and divergent creative thinking as a function of stimulation.' Dissertation Abstracts. Vol. 29, No. 12B, 4842.

CROPLEY, A.J. (1968). 'A note on the Wallach-Kogan tests of creativity.' British Journal of Educational Psychology. Vol. 38, Part 1, 197-201.

CROPLEY, A.J. (1969). Creativity. London, Longmans, Green and Co. Ltd. Education Today Series.

CROPLEY, A.J. (1970). 'S-R psychology and cognitive psychology.' Creativity. Penguin, edited P.E. Vernon. 116-125

CROPLEY, A.J. and FIELD, T.W. (1969). 'Achievement in science and intellectual style.' Journal of Applied Psychology. Vol. 53, No. 2, 132-135.

CROPLEY, A.J. and MASLANY, G.W. (1969). 'Reliability and factorial validity of the Wallach-Kogan creativity tests.' British Journal of Psychology. 60, No. 3, 395-398.

CRUTCHFIELD, R.S. (1965). 'Creative thinking in children: its teaching and testing.' Lecture delivered at Stanford University, in a lecture series on "Explorations in Intellectual Capacity" sponsored by Harcourt, Brace and World in commemoration of Lewis M. Terman and Arthur S. Otis.

CRUTCHFIELD, R.S. (1969). 'Nurturing the cognitive skills of productive thinking. Life skills in school and society.' ASCD Year Book. Louis J. Rubin (ed.) NEA. Washington D.C. 53-71.

DACEY, J., MADAUS, G., and ALLEN, A. (1969). 'The relationship of creativity and intelligence in Irish adolescents.' British Journal of Educational Psychology. Vol. 39, Part 3, 261-266.

DACEY, J. and RIPPLE, R. (1969). 'Relationships of some adolescent characteristics and verbal creativity.' Psychology in the Schools. Vol. 6, No. 3, 321-324.

DALRYMPLE-ALFORD, E.G. (1970). 'Measurement of clustering in free recall.' Psychological Bulletin Vol. 74, No. 1, 32-34.

DAMM, V.J. (1970). 'Creativity and intelligence: research implications for equal emphasis in high school.' Exceptional Children. Vol. 36, 565-569

DEWING, K. (1970). 'The reliability and validity of selected tests of creative thinking in a sample of seventh-grade West Australian children.' British Journal of Educational Psychology. Vol. 40, Part 1, 35-42.

DI VESTA, F.J. and BERNSTEIN, M.E. (1969). 'The effect of serial structure on children's and adults' paired-associate learning.' Journal of Experimental Child Psychology. Vol. 8, 63-73.

EDWARDS, A.L. (1970). The Measurement of Personality Traits by Scales and Inventories. Holt, Rinehart and Winston, Inc.

ENTWISLE, D.R. (1966). Word Associations of Young Children. Baltimore, Maryland. The Johns Hopkins Press.

FEE, F. (1968). 'An alternative to Ward's factor analysis of Wallach and Kogan's 'C eativity' correlations.' British Journal of Educational Psychology. Vol. 38, Part 3, 310-321.

FELDHUSEN, J.F., TREFFINGER, D.J. and ELIAS, R.M. (1970). 'Prediction of academic achievement with divergent and convergent thinking and personality variables.' Psychology in the Schools. Vol. 7, No. 1, 46-52.

FIELD, T.W. and CROPLEY, A.J. (1969). 'Cognitive style and science achievement.' Journal of Research in Science Teaching. Vol. 6, Issue 1, 2-10.

FIELD, T.W. and POOLE, M.E. (1970). 'Intellectual style and achievement of arts and science undergraduates.' The British Journal of Educational Psychology. Vol. 40, Part 3, 338-341.

FOX, P.W. (1970). 'Patterns of stability and change in behaviors of free association.' Journal of Verbal Learning and Verbal Behavior. Vol. 9, 30-36.

FREDRICK, W.C., and KLAUSMEIER, H.J. (1970). 'Cognitive styles: a description.' Educational Leadership. 668-72. April, 1970.

FRUCHTER, B. (1954). Introduction to Factor Analysis. Princeton, New Jersey. D. Van Nostrand Inc.

FULGOSI, A. and GUILFORD, J.P. (1968). 'Short-term incubation in divergent production.' American Journal of Psychology. Vol. 81, No. 2, 241-246.

GAIER, E.L., (1970). 'Identification of creativity: the individual.' Psychological Bulletin. Vol. 73, No. 1, 55-73.

GALL, M. and MENDELSOHN, G.A. (1967). 'Effects of facilitating techniques and subject-experimenter interaction on creative problem solving.' Journal of Personality and Social Psychology. Vol. 5, No. 2, 211-216.

GALLAGHER, J.J. and JENNE, W.D. (1967). 'The relationship of cognitive style to classroom expressiveness and associated variables.' The Journal of Educational Research. Vol. 60, No. 6, 273-278.

GERJUOY, I.R. and SPITZ, H.H. (1966). 'Associative clustering in free recall: intellectual and developmental variables.' American Journal of Mental Deficiency. Vol. 70, 918-27.

GETZELS, J.W. and CSIKSZENTMIHALYI, M. (1966). 'Portrait of the artist as an explorer.' Reprint from Transaction.

GETZELS, J.W. and CSIKSZENTMIHALYI, M. (1966a). 'The study of creativity in future artists: the criterion problem.' Reprint from Experience, Structure and Adaptability. Springer Publishing Company, New York.

GETZELS, J.W. and CSIKSZENTMIHALYI, M. (1967). 'Scientific creativity.' Reprint from Science Journal. 80-84.

GETZELS, J.W. and JACKSON, P.W. (1962). Creativity and Intelligence. London and New York. John Wiley and Sons, Inc.

GHISELIN, B. (1954). The Creative Process. A Symposium. Berkeley and Los Angeles: University of California Press.

GINSBURG, G.P. and WHITTEMORE, R.G. (1968). 'Cr ativity and verbal ability: a direct examination of their relationship.' British Journal of Educational Psychology. Vol. 38, 133-139.

GOTZ, I.L. (1968). 'Creativity and conformity' Educational Theory. 18. 262-71.

GRAY, J. (1968). 'An investigation of the relationship between primary process thinking and creativity.' Dissertation Abstracts. Vol. 28, No. 12B, 5206.

GROVER, B.L. (1966). 'Prediction of achievement in divergent and convergent learning situations.' Journal of Educational Research. Vol. 59, No. 9. 402-405.

GRUBER, H.E., TERRELL, G., and WERTHEIMER, M. (eds.) (1964). Contemporary Approaches to Creative Thinking. New York. Atherton Press.

GUILFORD, J.P. (1950). Fundamental Statistics in Psychology and Education. New York. McGraw-Hill Book Company, Inc.

GUILFORD, J.P. (1959). 'Traits of creativity', in H.H. Anderson (ed.) Creativity and its Cultivation. Harper. 142-61.

GUILFORD, J.P. (1959a). 'Three faces of intellect'. American Psychologist. 14. 469-79.

GUILFORD, J.P. (1963). 'Potentiality for creativity and its measurement.' In Proceedings of the 1962 Invitational Conference on Testing Problems: Educational Testing Service.

GUILFORD, J.P. (1964). 'Progress in the discovery of intellectual factors.' Widening Horizons in Creativity. (ed.) C.W. Taylor. John Wiley and Sons. 261-97.

HALL, W.B. and MACKINNON, D.W. (1965). 'The prediction of creativity from personality inventories.' Paper presented at the annual meeting of the Western Psychological Association, Honolulu, Hawaii.

HAMILTON, D. (1970). 'A comparison of school achievement, teachers' ratings, self ratings, and a personality score as predictors of creative thinking potential.' Dissertation Abstracts. Vol. 30, No. 7A, 2905.

HAMMOND, W.D. (1970). 'Certain differences in the syntactic structure of creative writing at four elementary grade levels.' Dissertation Abstracts. Vol. 30, No. 8A, 3189.

HARRIS, R. and HALL, A.E. (1970). 'Creativity and the need for associative novelty.' Canadian Journal of Psychology. Vol. 24, No. 2, 90-97.

HARTLEY, J. and BEASLEY, N. (1969). 'Contrary imaginations at Keele.' Universities Quarterly. Vol. 23, No. 4, 467-71.

HELSON, R. (1968). 'Generality of sex differences in creative style.' Journal of Personality. Vol. 36. 33-48.

HELSON, R. (1968a). 'Effects of sibling characteristics and parental values on creative interest and achievement.' Journal of Personality Vol. 36, No. 4, 589-607.

HENRY,, R. and VOSS, J.F. (1970). 'Associative strength growth produced via category membership.' Journal of Experimental Psychology. Vol. 83, No. 1, 136-140.

HOEPFNER, R. (1969/70). 'Intellectual aptitude involvement in thinking skills.' Journal of Research and Development in Education. Vol. 3, 43-52.

HOWELL, D. (1970). 'Free association reliability as a function of response strength.' Journal of Experimental Psychology. Vol. 85, No. 3, 431-433.

HUDSON, L. (1960). 'A differential test of arts/science aptitude.' Reprinted from Nature. Vol. 186, No. 4722. 413-414.

HUDSON, L. (1962). 'Intelligence, divergence and potential originality.' Reprint from Nature. Vol. 196, No. 4854. 601-602.

HUDSON, L. (1963). 'Personality and scientific aptitude.' Nature. Vol. 198, No. 4883, 913-914.

HUDSON, L. (1963a). 'The relation of psychological test scores to academic bias.' Reprint from The British Journal of Educational Psychology. Vol. XXXIII. 120-131.

HUDSON, L. (1964). 'Future open scholars.' Reprint from Nature. Vol. 202, No. 4934, 834.

HUDSON, L. (1966). Contrary Imaginations. London. Methuen and Co. Ltd.

HUDSON, L. (1966a). 'Research on intellectual types: artists, scientists and technologists.' Advancement of Science. 612-616.

HUDSON, L. (1967). 'The stereotypical scientist.' Reprint from Nature. Vol. 213, No. 5073, 228-229.

HUDSON, L. (1968). Frames of Mind. (Ability, perception and self-perception in the arts and sciences.) Methuen and Co. Ltd. London.

IWATA, O. (1968). 'Some relationships of creativity with intelligence and personality variables.' Psychologia. 11, 211-220.

JACKSON, P.W., and MESSICK, S. (1965). 'The person, the product, and the response: conceptual problems in the assessment of creativity.' Journal of Personality. Vol. 33, No. 3, 309-329.

JACKSON, P.W. and MESSICK, S. (1967). 'Creativity.' Research Memorandum. E.T.S.

JACOBSON, L.I.; ELENEWSKI, J.J.; LORDAHL, D.S.; and LIROFF, J. (1968). 'Role of creativity and intelligence in conceptualization.' Journal of Personality and Social Psychology. Vol. 10, No. 4, 431-436.

JEWITT, M. (1966). A Secondary Book of Chemistry. Sampson Low, Marston and Co. Ltd.

JOHNS, L.D. (1970). 'Scientific thinking in biology by pupils who have completed GCE O-level courses in biology.' The British Journal of Educational Psychology. Vol. 40, Part 3. 355

JOHNSON, P.E. (1969). 'On the communication of concepts in science.' Journal of Educational Psychology. Vol. 60, No. 1. 32-40

JOHNSON, D.M.; PARROTT, G.R. and STRATTON, R.P. (1967). 'Productive thinking: produce one solution or many.' Reprinted from the Proceedings, 75th Annual Convention APA. 299-300.

KARPLUS, R. (1964). 'Science in the elementary school.' New Developments in Elementary School Science: A Conference. (Oklahoma City: The Frontiers of Science Foundation of Oklahoma Inc.)

KATZ, J.J.(1967). 'Recent issues in semantic theory.' Foundations of Language. Vol. 3, 124-194.

KEATS, J.A. (1959). 'Tests AH4 and AH5.' Review by J.A. Keats, Senior Lecturer in Psychology, University of Queensland. The Fifth Mental Measurements Year Book, 1959. (ed.) O.K. Buros. 391.

KHATENA, J. (1970). 'The training of creative thinking strategies and its effects on originality.' Dissertation Abstracts. Vol. 30, No. 8A, 3325.

KINSBOURNE, M. (1968). 'The contrary imaginations of arts and science students: a critical discussion.' Developmental Medicine and Child Neurology. Vol. 10, No. 4, 461-464.

KNELLER, G.F. (1965). The Art and Science of Creativity. U.S.A. Holt, Rinehart and Winston, Inc.

KOGAN, N. (1967). Private discussion, Cambridge, England.

KOGAN, N. and MORGAN, F.T. (1969). 'Task and motivational influences on the assessment of creative and intellective ability in children.' Genetic Psychology Monographs. 80. 91-127.

KOGAN, N. and WALLACH, M.A. (1964). Risk Taking. A Study in Cognition and Personality. New York. Holt, Rinehart and Winston.

KROP, H.D.; ALEGRE, C.E.; and WILLIAMS, C.D. (1969). 'Effect of induced stress on convergent and divergent thinking.' Psychological Reports. Vol. 24, 895-898.

KUHN, T.S. (1962). The Structure of Scientific Revolutions. The University of Chicago Press, Chicago.

KURTZMAN, K.A. (1967). 'A study of school attitudes, peer acceptance, and personality of creative adolescents.' Exceptional Children. Vol. 34, No. 3, 157-162.

LAKOFF, S.A. (1967). 'The way of the scientist.' (Book Review). (eds.) Simon and Schuster, 1966. Harvard Educational Review. Vol. 37, 659-664.

LAWLEY, D.N. and MAXWELL, A.E. (1963). Factor Analysis as a Statistical Method. Butterworths. London.

LICHTENWALNER, J. and MAXWELL, J. (1969). 'The relationship of birth order and socio-economic status to the creativity of preschool children.' Child Development. Vol. 40, (4). 1241-1247.

LIGGETT, J. (1965). 'Tests AH4 and AH5'. Review by J. Liggett, Senior Lecturer in Psychology, University College of South Wales. The Sixth Mental Measurements Yearbook, 1965. (ed.) O.K.Buros. 506.

LOIGMAN, B. and GOSS, A. (1969). 'Attributes of unfamiliar, long, scientific names.' Psychonomic Science. Vol. 17, No. 6, 327-328.

LOVELL, K. (1968). 'Some recent studies in cognitive and language development.' Merrill-Palmer Quarterly. Vol. 14, No. 2, 123-138.

LYTTON, H. and COTTON, A.C. (1969). 'Divergent thinking abilities in secondary schools.' British Journal of Educational Psychology. Vol. 39. 188-90.

McALPINE, D.M. (1970). 'The creativity enterprise.' Delta Six. 16-25.

McCORMACK, A. (1970). 'The effect of selected teaching methods on creative thinking, self-evaluation and achievement of students enrolled in an elementary science education methods course.' Dissertation Abstracts. Vol. 30, No. 10, 4311-4312.

McCUTCHEON, N. (1970). 'A study of relationships among creativity, intelligence and test anxiety of middle-class fourth grade boys and girls.' Dissertation Abstracts. Vol. 30, No. 11a, 4833.

McHENRY, R.E.; and SHOUKSMITH, G.A. (1970). 'Creativity visual imagination and suggestibility: their relationship in a group of 10-year-old children.' British Journal of Educational Psychology. Vol. 40, Part 2, 154-160.

MacKAY, C.K. and CAMERON, M.B. (1968). 'Cognitive bias in Scottish first year science and arts undergraduates; convergent and divergent thinking.' British Journal of Educational Psychology. Vol. 38, 315-18.

MacKEAN, D.G. (1962). Introduction to Biology. Jarold and Sons. England.

MacKINNON, D.W. (1958). 'Implications of personality assessment for the identification of the effective teacher.' An address delivered before the California School Supervisors' Association, Bay Area Section, Sonoma Mission Inn.

MacKINNON, D.W. (1960). 'The highly effective individual.' Reprint from Teachers College Record. Vol. 61, No. 7, 367-378.

MacKINNON, D.W. (1962). 'Intellect and motive in scientific inventors: implications for supply.' Reprint from The Rate and Direction of Inventive Activity. Economic and Social Factors. National Bureau of Economic Research. 361-384.

MacKINNON, D.W. (1962a). 'The nature and nurture of creative talent.' Reprint from American Psychologist. Vol. 17, No. 7, 484-495.

MacKINNON, D.W. (1963). 'Creativity and images of the self.' Reprinted with permission of Atherton Press from The Study of Lives, edited by Robert W. White. New York.

MacKINNON, D.W. (1964). 'Creativity and transliminal experience.' Address sponsored by Division 8 of the American Psychological Association and given at its 72nd Annual Convention in Los Angeles, California.

MacKINNON, D.W. (1965). 'Personality and the realization of creative potential.' American Psychologist. Vol. 20, No. 4, 273-281.

MacKINNON, D.W. (1966). 'Educating for creativity - a modern myth.' A lecture given at the Conference on Education for Creativity inthe American College, sponsored by the Extension Division, University of California, Berkeley.

MacKINNON, D.W. (1966a) 'Some reflections on the current status of personality assessment with special reference to the assessment of creative persons.' A talk to the Faculty Symposium, Department of Psychology, University of California, Berkeley.

MacKINNON, D.W. (1966b) 'Instructional media in the nurturing of creativity.' Offprints from Instructional Media and Creativity. Edited by C.W. Taylor, published by John Wiley and Sons Inc.

MacKINNON, D.W. (1967). Private communication. I.P.A.R.

McNEILL, D. (1966). 'A study of word association.' Journal of Verbal Learning and Verbal Behavior. Vol. 5, No. 6, 548-557.

McNEMAR, Q. (1964). 'Lost: our intelligence? Why?' Address of the President to the Seventy-Second Annual Convention of the American Psychological Association, Los Angeles, September 5, 1964.

MACKWORTH, N.H. (1965). 'Originality.' American Psychologist. Vol. 20, No. 1, 51-66.

MARIN, M. (1968). 'A comparison of the performance of high school physics students using closely directed experiments with that of students using open ended experiments.' Dissertation Abstracts. Vol. 29, No. 2A, 515-516.

MASON, D.K. (1960). 'Earth-space concepts.' The Instructor. 69-70.

MASSIALAS, B.G. and ZEVIN, J. (1967). Creative Encounters in the Classroom. New York. John Wiley and Sons, Inc.

MASTERS, J.C. and ANDERSON, G.W. (1970). 'Delay and instructions as determinants of originality in word associations.' Journal of Verbal Learning and Verbal Behavior. Vol. 7, 432-433.

MASTERS, J.C.; MESIBOV, G.B.; and ANDERSON, G.W. (1970). 'Word association type and the temporal stacking of responses.' Journal of Verbal Learning and Verbal Behavior. Vol. 9, No. 2, 207-211.

MATTHEWS, W.A., and HOGGART, K. (1970). 'Associative grouping and free recall.' British Journal of Psychology. Vol. 61, Part 3, 345-358.

MEARIG, J.S. (1967). 'Sex and cultural differences in performance on certain verbal creativity measures.' The Journal of Educational Research. Vol. 61, No. 3, 114-117.

MEDNICK, S.A. (1962). 'The associative basis of the creative process.' Psychological Review. Vol. 69, No. 3, 220-232.

MEEHL, P.E. (1967). 'Theory testing in psychology and physics: a methodological paradox.' Philosophy of Science. June. 103-115.

MEHRYAR, A.H. and SHAPURIAN, R. (1970). 'The application of a high-grade intelligence scale in Iran.' The British Journal of Educational Psychology. Vol. 40, Part 3, 307-313.

MENDELSOHN, G.A. and GRISWOLD, B.B. (1964). 'Differential use of incidental stimuli in problem solving as a function of creativity.' Reprint from Journal of Abnormal and Social Psychology. Vol. 68, No. 4, 431-436.

MENDELSOHN, G.A. and GRISWOLD, B.B. (1966). 'Assessed creative potential, vocabulary level, and sex as predictors of the use of incidental cues in verbal problem solving.' Journal of Personality and Social Psychology. Vol. 4, No. 4, 423-431.

MESSICK, S. and FRITZKY, F.J. (1963). 'Dimensions of analytic attitude in cognition and personality.' Journal of Personality. Vol. 31, No. 3, 346-370.

MESSICK, S.J. and KOGAN, N. (1963). 'Differentiation and compartmentalization in object-sorting measures of categorizing style.' Perceptual and Motor Skills. Vol. 16, 47-51.

MESSICK, S.J. and SOLLEY, C.M. (1957). 'Word-association and semantic differentiation.' The American Journal of Psychology. Vol. LXX, 586-593.

MILLER, B.J.; RUSS, D.; GIBSON, C.; and HALL, A.E. (1970). 'Effects of free association training, retraining, and information on creativity.' Journal of Experimental Psychology. Vol. 84, No. 2, 226-229.

MOROZOV, M.N. (1968). 'Tvorcheskii protsess kak poisk optimal' nogo upravleniya.' (The creative process as search for optimal control.) Vestnik Acadernii Meditsinskikh Nauk. 23 (II) 76-89. USSR

MORRISON, D.E. and HENKEL, R.E. (1969). 'Significance tests reconsidered.' The American Sociologist. May, 1969. Vol. 4, No. 2.

MOTLAGH, H. (1968). 'Selected non-academic factors related to creativity.' Dissertation Abstracts. Vol. 29, No. 5A, 1453.

MURDOCH, P. and VAN BRUGGEN, Y.O. (1970). 'Stability, generality, and change of category width.' Journal of Personality. Vol. 38, No. 1, 117-33.

NEMAN, R. and DIXON, T.R. (1970). 'Stimulus generality in word association sets.' Journal of Experimental Psychology. Vol. 83, No. 3, 506-507.

NICHOLS, J.G. (1971). 'Some effects of testing procedure on divergent thinking.' Child Development (in press)

NUFFIELD FOUNDATION SCIENCE TEACHING PROJECT, 1967. Biology, Chemistry, Physics. Longmans/Penguin Books Ltd.

OSGOOD, E.C. (1953). Method and Theory in Experimental Psychology. New York, Oxford University Press.

OSGOOD, C.E., SUCI, G.J. and TANNENBAUM, P.H. (1957). The Measurement of Meaning. Illinois: University Press.

PANKOVE, E., and KOGAN, N. (1968). 'Creative ability and risk-taking in elementary school children.' Journal of Personality. Vol. 36, Part 3, 420-439.

PARLOFF, M.B.; DATTA, L.; KLEMAN, M.; and HANDLON, J.H. (1968). 'Personality characteristics which differentiate creative male adolescents and adults.' Journal of Personality. Vol. 36, (4) 528-552.

PARNES, S.J. (1963). 'Education and creativity.' Reprint from Teachers College Record, Vol. 64, No. 4, 331-339.

PARNES, S.J., and HARDING, H.F. (eds.) A Source Book for Creative Thinking. New York. Charles Scribner's Sons.

PEARCE, C. (1968). 'Creativity in young science students.' Exceptional Children. Vol. 35, (2) 121-126.

POLLERT, L.H.; FELDHUSEN, J.F.; VAN MONDFRANS, A.P.; TREFFINGER, D.J. (1969). 'Role of memory in divergent thinking.' Psychological Reports. Vol. 25, 151-156.

POLLIO, H.R.; DEITCHMAN, R. and RICHARDS, S. (1969). 'Law of contrast and oppositional word associates.' Journal of Experimental Psychology. Vol. 79, No. 2, 203-212.

POPHAM, W.J. (1967). Educational Statistics: Use and Interpretation. Harper and Row. New York.

POVEY, R.M. (1970). 'Arts/science differences: their relationship to curriculum specialization.' British Journal of Psychology. Vol. 61, No. 1, 55-64.

RANSOM, W. (1969). 'Effect of 'Science - a process approach' on creative thinking and performance in selected processes of science in the second grade.' Dissertation Abstracts. Vol. 30, No. 2A, 625-626.

REID, N.A. (1970). 'Creative Thinking: A factorial study of creative thinking abilities in intermediate school children.' Unpublished master's thesis, University of Auckland, N.Z.

ROE, A. (1953). The Making of a Scientist. New York: Dodd, Mead.

ROE, A. (1953a). 'A psychological study of eminent psychologists and anthropologists and a comparison with biological and physical scientists.' Psychological Monograph. 67, No. 2 (whole No. 352).

RENZULLI, J.S. (1970). 'Problems and issues in the measurement of creativity.' Paper presented at the Annual Meeting of the American Educational Research Association, Minneapolis, Minnesota, March, 1970. Symposium, Assessing Creativity: Progress in Both Directions. 1-8.

RITTMAYER, J. (1968). 'Relationships among high verbal high non-verbal and high total creativity scores and intelligence, academic achievement, socio-economic status and teacher judgements.' Dissertation Abstractsm Vol. 28, No. 12A, 4913.

RORER, L.G. (1963). 'The great response-style myth.' Psychological Bulletin. Vol. 63, No. 3, 129-156.

ROTBERG, I.C. (1968). 'A method for developing comprehensive categories of meaning.' Journal of Verbal Learning and Verbal Behavior. Vol. 7, 589-592.

SCHLICHT, Jr, W.J.; ANDERSON, D.L.; HELIN, W.D.; HIPPE, D.L.; LISTIAK, R.L.; MOSER, R.J.; and WALKER, J.L. (1968). 'Creativity and intelligence: Further findings.' Journal of Clinical Psychology. 24, (4), 458.

SCHMIDT, B. (1968). 'The influence of the space age on written language.' Elementary English. Vol. 45, 643-47.

SCHULMAN, D. (1966). 'Openness of perception as a condition for creativity.' Exceptional Children. Vol. 33, 89-94.

SCOTT, N.C. Jr. (1966). 'The strategy of inquiry and styles of categorisation.' Journal of Research in Science Teaching. Vol. 4, 143-153.

SHAPIRO, R.J. (1968). 'Creative research scientists.' Psychologia Africana. Monograph supplement No. 4. National Institute for Personnel Research. 1-180.

SHIGAKI, I. (1970). 'The effects of teacher strength and sensitivity and pupil intelligence and creativity on the production of divergent responses.' Dissertation Abstracts. Vol. 30, No. 9A. 3864.

SHOUKSMITH, G. (1970). Intelligence, Creativity and Cognitive Style. B.T. Batsford Ltd. London.

SINAY, R. (1968). 'Creative aptitude patterns of college honors students.' Dissertation Abstracts. Vol. 28, No. 12B, 5212.

SKAGER, R.W.; SCHULTZ, C.B.; and KLEIN, S.P. (1965). 'Quality and quantity of accomplishments as measures of creativity.' Journal of Educational Psychology. Vol. 56, No. 1, 31-39.

SMITH, M.S. (1960). Modern Physics. Longmans, Green and Co.

SMOKE, J.G. (1970). 'To structure or not to structure: a divestiture of two aspects of creativity.' Art Education. Vol. 23, No. 4, 19-21.

SOLOMON, G. (1969). 'The classification and measurement of image provoking cognitive behaviors of science teachers.' Dissertation Abstracts. Vol. 30, No. 1A, 75.

STEINBERG, D.D. (1970). 'Analyticity, amphigory, and the semantic interpretation of sentences.' Journal of Verbal Learning and Verbal Behavior. Vol. 9, 37-51.

STENHOUSE, D. (1971). 'Scientific creativity: 'normal' or 'revolutionary'.' The Australian Journal of Education. Vol. 15, No. 2, 171-184.

STONE, P.H.; DUNPHY, D.C.; SMITH, M.S.; OGILVIE, D.M. with associates. (1966). The General Inquirer: A Computer Approach to Content Analysis. Cambridge, Massachusetts. The M.I.T. Press.

SUCHMAN, J.R. (1964). 'The child and the inquiry process.' in Intellectual Development: Another Look. A.S.C.D.

SULTAN, E.E. (1962). 'A factorial study in the domain of creative thinking.' British Journal of Educational Psychology. Vol. 32, 78-82.

TABA, H. (1963). 'Opportunities for creativity in education for exceptional children.' Exceptional Children. Vol. 29, No. 6, 247-256.

TAFT, R., and GILCHRIST, M.B. (1970). 'Creative attitudes and creative productivity: a comparison of two aspects of creativity among students.' Journal of Educational Psychology. Vol. 61, No. 2, 136-143.

TAYLOR, D.W. (1963). 'Creative thinking among scientists.' Reprint from Creativity and College Teaching published as Vol. XXXV (1963, pp.27-44) of the Bulletin of the Bureau of School Service (College of Education, University of Kentucky, Lexington).

TAYLOR, D.W. (1963a). 'Variables related to creativity and productivity among men in two research laboratories.' Offprint from Scientific Creativity: Its Recognition and Development. Taylor and Barron (eds). John Wiley and Sons, Inc.

TAYLOR, C.W. and BARRON, F. (1963). Scientific Creativity: Its Recognition and Development. New York. John Wiley and Sons, Inc.

TAYLOR, Bob. L., and McKEAN, R.C. (1968). 'Divergent thinkers and teacher education.' Journal of Educational Research. Vol. 61, No. 9. 417-418.

TAYLOR, C.W. (ed.) (1964). Widening Horizons in Creativity. New York. John Wiley and Sons, Inc.

TAYLOR, C.W. (ed.) (1964a). Creativity: Progress and Potential. New York. McGraw-Hill Book Company.

TAYLOR, D.W.(1960). 'Thinking and creativity.' Reprinted from Annals of the New York Academy of Sciences. Vol. 91, Article 1, 108-127.

TAYLOR, D.W. (1961). 'Environment and creati ity.' Paper presented to a symposium on The Identification and Utilization of Talent at the XIV International Congress of Applied Psychology held at Copenhagen, Denmark, August 13-19, 1961.

THOMSON, Sir G. (1956). The Factorial Analysis of Human Ability. London, University of London Press.

TIBBETTS, J. (1968). 'Relationships of creativity to socioeconomic status, race, sex, I.Q., age and grade-point average in an adolescent population.' Dissertation Abstracts. Vol. 29, No. 4A, 1174-5.

TORRANCE, E.P. (1962). Guiding Creative Talent. Prentice-Hall, Inc., Englewood Cliffs, N.J., USA.

TORRANCE, E.P. (1966). Torrance Tests of Creative Thinking. Directions Manual and Scoring Guide Verbal Test, Booklet B. Research Edition. Personnel Press, Inc., Princeton, New Jersey.

TROSMAN, H. (1969). 'The cryptomnesic fragment in the discovery of free association.' Journal of the American Psychoanalytic Association. 17, (2), 489-510.

TROWBRIDGE, N. (1963). 'Cross cultural study of creative ideas in children.' Comparative Education Review. Vol. 12, 80-83.

TROWBRIDGE, N. (1969). 'Creativity criteria for evaluating teachers and students in classrooms.' Reprint from the Proceedings, 77th Annual Convention, APA. 1969. 639-640.

TURNER, R.L., and DENNY, D.A. (1969). 'Teacher characteristics, teacher behavior, and changes in pupil creativity.' Elementary School Journal. Vol. 69.

UHES, M.J., and SHAVER, J.P. (1970). 'Dogmatism and divergent-convergent abilities.' Journal of Psychology. Vol. 75, 3-11.

VERNON, P.E. (1964). 'Creativity and intelligence.' Educational Research, 6, No. 3, 163-169,

VERNON, P.E. (1969). Intelligence and Cultural Environment. London. Methuen and Co. Ltd.

VERNON, P.E. (Editor) (1970). Creativity. Penguin modern psychology readings.

VERNON, P.E. (1971). 'Effects of administration and scoring on divergent thinking tests.' The British Journal of Educational Psychology. Vol. 41, Part 3, 245-257.

VICK, O.C., and JACKSON, D.N. (1967). 'Cognitive styles in the schematizing process: a critical evaluation.' Educational and Psychological Measurement. Vol. 27. 267-286.

VILLAS-BOAS, M.C.C. (1968). 'A study of the motivational role of self-concept and locus of control in creative children.' Dissertation Abstracts. Vol. 29, No. 1B, 400.

VYGOTSKY, L.S. (1962). Thought and Language. (ed. and translated by E. Hanfmann and G. Vakar) Cambridge, Massachusetts. The M.I.T. Press.

WADDINGTON, D.J. and FINLAY, H.S. (1965). Organic Chemistry Through Experiment. Mills and Boon Ltd.

WALBERG, H.J. and WELCH, W.W. (1967). 'Personality characteristics of innovative physics teachers.' The Journal of Creative Behavior.

WALLACH, M.A. (1959). 'The influence of classification requirements on gradients of response.' Psychological Monographs - General and Applied. Vol. 73, No. 8, 1-21.

WALLACH, M.A., and KOGAN, N. (1965). Modes of Thinking in Young Children. New York. Holt, Rinehart and Winston, Inc.

WALLACH, M.A., and WING, Jr, C.W. (1969). The Talented Student. New York. Holt, Rinehart and Winston, Inc.

WARD, W.C. (1969). 'Rate and uniqueness in children's creative responding.' Child Development. Vol. 40, No. 3, 869-78.

WASHTON, N.S. (1967). 'Teaching science creatively: a taxonomy of pupil questions.' Science Education. Vol. 51, No. 5, 428-431.

WECHSLER, D. (1949). Wechsler Intelligence Scale for Children Manual. A.C.E.R. Melbourne, Australia.

WILLIAMS, F.E. (1967). 'Intellectual creativity and the teacher.' The Journal of Creative Behavior. Vol. 1, No. 2, 173-180.

WINDHOLZ, G. (1967). 'Divergent and convergent abilities of semantic content as related to some personality traits of college students.' Educational and Psychological Measurement. 27, 1015-1023.

WINDHOLZ, G. (1968). 'The relation of creativity and intelligence constellations to traits of temperament, interest and value in college students.' Journal of General Psychology. Vol. 79, 291-299.

WITKIN, H.A., and OLTMAN, P.K. (1967). 'Cognitive style.' International Journal of Neurology. Vol. 6, No. 2, 119-137.

WOOD, R.H. (1970). 'Three environmental press variables and their relationship to coping behavior and creativity in children.' Dissertation Abstracts. Vol. 30, No. 12, 5306-5307.

WOOD, C.W. and HOLLIDAY, A.K. (1964). Physical Chemistry. Butterworths.

YAMAMOTO, K. (1962). 'Intercorrelations between measures of creative writing and measures of intelligence, reading, language, study work and arithmetic skills.' Research Memo, BER-62-4 Minneapolis: Bureau of Educ. Research, University of Minnesota.

YAMAMOTO, K. (1964). 'Evaluation of some creativity measures in a high school with peer nominations as criteria.' Journal of Psychology. 58. 285-293.

YAMAMOTO, K. (1965). 'Effects of restriction of range and test unreliability on correlation between measures of intelligence and creative thinking.' British Journal of Educational Psychology. Vol. 35, Part 3, 300-305.

YAMAMOTO, K. (1967). 'Creativity and unpredictability in school achievement.' The Journal of Educational Research. Vol. 60, No. 7. 321-325.

YAROSHEVSKII, M.G. (1969). 'Logika Razvitiya Nauki.' Voprosy Filosofii. Vol. 23, No. 3, 44-55.