EFFECTS OF DIFFERENT METHODS OF ADMINISTRATION ON PERFORMANCE IN CONVERGENT AND DIVERGENT "TESTS"

by

Vivienne Jayne Renner

B.A. (Auckland) 1968

Diploma of Education (Adelaide) 1970

M.Ed. Qualifying Exams (Adelaide) 1971

Advanced Diploma of Education (Adelaide) 1971

B.Sc. (Adelaide) 1974

A thesis presented to the

BOARD OF RESEARCH STUDIES

UNIVERSITY OF ADELAIDE

In Fulfillment of the

Requirements for the Degree

DOCTOR OF PHILOSOPHY

Department of Education

January 21, 1978

TABLE OF CONTENTS

CHAPTER

LIST OF TABLES

TABLE

TABLE

TABLE

TABLE

TABLE

SUMMARY

The intention of this study was to investigate the effect of varying the conditions of administration and the order of presenting different tests of convergent and divergent thinking on levels of performance. In addition, the research assessed whether the relationship between convergent and divergent "tests" differed under these experimental regimes.

A survey of the literature in the field of convergent and divergent thinking indicates that there is controversy regarding their nature. Furthermore, there are conflicting research reports as to whether formal or informal conditions enhance levels of performance on divergent tests and also whether convergent and divergent thinking are distinct factors of the intellect. There have been no prior extensive investigations into the effect of different conditions on convergent thinking or the effect of order of testing.

Eight different experimental regimes involving the condition of test administration and order were applied to 493 Grade 5 pupils with an average age of 120.2 months. There were two types of condition, either formal, i.e., test and informal, i.e., play, and presentation of the order of the convergent and divergent tests was varied

alternately between any two schools having the same conditions. A battery of four convergent and four divergent tests were given.

Analysis of the results showed that:

(a) Children perform better on convergent tasks under test conditions.

(b) Children perform better on divergent tasks under test conditions.

(c) Order of "tests" exerts an effect on divergent scores. Children perform better on the divergent tasks when the divergent battery precedes the convergent battery.

A preceding convergent battery imposes a limiting set so that such tasks, particularly under formal conditions, have an inhibitory effect on subsequent divergent thinking processes.

Generally, the relationships between tests of convergent and divergent thinking are low, although three of the eight schools had values which were significant. Neither test or game conditions appeared to affect this relationship. Order of testing appeared to influence the relationship in two cases. The lowest r value occurred in that school having formal conditions for both types of tasks, but having the divergent battery prior to the convergent battery.

The study concluded that convergent and divergent "tests" should be given in a formal context with the divergent stimuli presented first. This procedure also provides the most meaningful separation of the two cognitive modes of thought. Finally, divergent tests seem more responsive to contextual effects of measurement than convergent tests. This may reflect the possibility that divergent abilities are more dependent upon the transient states of the organism; e.g., set or arousal.

This thesis is based upon original research of the undersigned. No part of the research data or the text have been used for or accepted for the award of any degree from any university. To the undersigned's knowledge and belief, the thesis contains no material previously published or written by another person except as otherwise cited in the text and given due reference.

V. Jayne Renner

ACKNOWLEDGMENTS

This Ph.D. thesis for me is a symbol of my intellectual growth. My intellectual life, beyond that of my parents, began with my favourite grandfather who I simply called Bumpa because it was easier to say. So it is to Bumpa I must first give acknowledgment for introducing me to words and to ideas. This thesis is thus, for me, a significant outcome of his teachings. To the rest of my family, and particularly my parents, thanks are due for not only giving me such an excellent upbringing, education and lots of love, but also for never failing to answer my questions. With such a background, therefore, it seems appropriate that I should have sought to study some aspects of the structure of intellect and how best this could be expressed.

The planning and conduct of this research was carried out at the University of Adelaide under the supervision of Dr. Jack Rowell. The analysis of the data and the writing was done in the University of Southern California with Dr. James E. Birren as external supervisor.

To Dr. Jack Rowell I give special acknowledgment for giving me not only the initial concepts of this thesis but also the theoretical and empirical background, in addition to the essential methodological tools for allowing me to carry out this investigation. He was not only a great teacher but a true friend during my residence in the University of Adelaide, South Australia. To the other members of the Department of Education, University of Adelaide, I also give my grateful thanks for their encouragement and support. I also thank Professor Welford of the Department of Psychology, University of Adelaide, for his advice to me in Southern California, a few weeks before my final oral examination.

Maryjo Walsh and Garst Reese were invaluable advisers and friends during the difficult statistical analyses. Not only did they help to provide meaning to a vast amount of data but provided stoic support during the long days of writing up. To the members of the University of Southern California examining committee, Dr. Birren, Dr. Schaie and Dr. Ragan I also give thanks for their constructive criticism and advice which allowed the final draft to come to fruition. Particular thanks go to Eleanor Hudson, my invaluable typist and friend, who not only typed the two drafts of this thesis, but patiently tolerated my indecisions and acted as a final arbiter on sentence construction.

There were numerous friends who provided support and encouragement, particularly when the inevitable "giving up" moods engulfed me. Not the least of these was Eleanor

James, who with her special type of magic gave me the will to continue. Dr. Ian Renner, my husband, saw me leap, stumble and sometimes fall during the entirety of my graduate career. His endurance is to be commended. His love and support, I could never have done without. The fact that he always cared about my work is a rare gift and one for which I am especially grateful.

This thesis represented a personal struggle for me perhaps because, as my father once said, its completion represented a stage of adulthood I seemed unwilling to accept. Appropriately, therefore, I thank Dr. Rameljack who increased my personal understanding. I also give particular thanks to Dr. Kiely.

Final thanks go to Dr. James Birren of the University of Southern California, my external supervisor, my mentor my teacher, my colleague, and my friend. Special acknowledgment is due to him for not only advising me and guiding me through the analyses and writing up of this thesis, but also promoting my intellectual and personal growth. He once said to me, on one of my emotional "balking" days, that he would drag me into maturity kicking and screaming. The completion of this thesis represents the fact that, to a very great extent, he has succeeded in that goal and much to <u>my</u> surprise, I am smiling.

CHAPTER I

INTRODUCTION

Purpose

The purpose of this research was to examine how varying the conditions of administration would affect the level of performance children display on convergent and divergent tests. Specifically, the research was designed to study the effects of either an evaluative or a play-like atmosphere on the levels of performance on "tests of convergent and divergent thinking." An additional purpose was to determine whether the relationship between convergent and divergent scores changed in response to the conditions under which the "tests" were taken.

This research was planned in a context that was concerned with the appropriate manner of evaluating the intellectual abilities of children. It has been suggested that children might perform better on tests of intelligence and creativity if such tests were given under play-like conditions, rather than under the more formal atmosphere that usually surrounds the taking of tests in schools. Presumably if more relaxed conditions could be devised then children might show much better levels of performance. The intent of this research was to determine how best to test children so as to bring out the optimum expression of their divergent (or creative) and convergent (or intellectual) abilities. The measurement of intelligence is still a widespread activity in schools motivated by many humanitarian goals for children, such as proper placement in classes to develop children's latent capacities. It is usual that at several junctures in the child's schooling, tests are given to guide the planning of the individual's school career. While tests have been subject to much criticism in recent years, there remains the issue about how best to assess the capacities of children lest we anticipate too much of some children and neglect the potential abilities of others who may go unnoticed as a result of retiring personalities in the classroom and at home.

Another problem also arises, at least in Australia and New Zealand, where, despite the emphasis on regular assessment such as IQ tests, primary schools have introduced game-like contexts for many "formal subjects," such as mathematics. Open class units in which children are not assigned regular places in the traditional sense for every lesson, share the same large room with children from other grades, work at their own pace, as well as being taught according to applied Piagetian methods designed to encourage discovery of principles, may not be appropriate for many children. Inadequate levels of achievement in core subjects at the primary school level causes problems in the early years of secondary school, at the very least, since children starting high school may not show sufficient knowledge and skills in such subjects. The problem has grown to such magnitude that a serious question arises as to whether a game-like classroom atmosphere really facilitates assessment and learning, and whether children perform better, when it becomes necessary to assess them, under informal or formal conditions.

Since 1950, in addition to the long-standing interest in measuring intelligence, creativity has become a subject of interest. Attempts have been made to devise tests which purportedly measure creativity as a different attribute from intelligence. Guilford in his Presidential address to the American Psychological Association (1950) lamented the fact that researchers had traditionally avoided investigating the area of creativity, yet he emphasized the necessity for systematically understanding what it was that allowed some individuals to exhibit acknowledged creative behaviour to a noteworthy degree. In 1972(a) Guilford again, this time while addressing the National Association for Gifted

Children, pointed out that the flight of the first Russian Sputnik in 1958 and America's felt threat to world leadership had accelerated research into all aspects of education of which the drive to produce more creative scientists and scholars was one. In addition to attempting to increase people's intellectual performance of which creative productivity was part, questions began to be posed during the 1960's as to whether intelligence and creativity were independent or related capacities and whether some situations would improve or decrease their level of performance. Since the answers to such questions are still ambiguous, the search legitimately continues. Related questions arose as to whether children were maximally creative under test or play conditions and whether the relationship between creativity and intelligence altered if creativity tests were given under game-like conditions. Finally, investigators examined the importance of certain personality traits in relation to performance on these cognitive tasks. An additional question, not yet investigated on any large scale, is whether intelligence test performances also alter under informal as against formal conditions. Furthermore, the statistical relationship between intelligence and creativity under different administrative regimes is still controversial. It was with some of these problems with which the present study was concerned.

In this thesis the terms convergent and divergent thinking are used interchangeably with the respective terms intelligence and creativity. This is a common practice in the literature. Although the terms can be distinguished, the degree of overlap is an issue beyond the scope of this thesis. Intelligence or convergent thinking is taken to mean the level of performance on intellectual tests which require one right or correct answer (as on an IQ test). Creativity or divergent thinking is taken to mean the level of performance on intellectual tests which are open-ended and have no one correct answer but require subjects to give as many different responses as they are capable of giving.

Intelligence as measured by IQ tests can be looked upon as the general factor which is common to standardized tests of intelligence (Spearman, 1904; Jensen, 1969), or a broad multivariate collection of variables most of which are independent, and of which divergent thinking is a part (Guilford, 1967). IQ is generally regarded as a standard score derived from performance on intelligence tests, as different from learning ability which is the capacity for acquiring, retaining and producing new information. However, certain IQ tests are regarded by some to be a good predictor of a person's learning ability (Wechsler, 1958; Jensen, 1969).

Creativity has been a very popular modern subject, and the distinction between creativity and intelligence has been the topic of a large number of investigations which will later be reviewed in this thesis. In terms of general results from previous investigations, intelligence and creativity can be regarded as different cognitive modes, although their relationship is controversial. For the purpose of this thesis, creativity or divergent thinking and intelligence or convergent thinking are regarded as two aspects of cognitive functioning. The relationship between the two is presumed to change under different conditions of administration. Thus, creativity or divergent thinking as represented by a battery of creativity tests is seen in this investigation to represent one generalized cognitive mode, whereas it is assumed that the more commonly used IQ tests represent another generalized cognitive mode labelled intelligence or convergent thinking.

Furthermore, due to the controversy surrounding the nature of intelligence and intelligence test results, and also due to the fact that conventional administrative procedures were in some circumstances altered, the term intelligence is not used in its usual traditional sense but only as far as it pertains to the type of thinking demanded in the tasks given. Indeed, the term convergent thinking is preferred and will be more generally adopted. Thus, convergent thinking test results were those derived from otherwise traditional IQ tests that demanded only one single correct answer. Convergent thinking ability was seen as the level of performance derived from such tests.

In the same line of thinking, due to the controversy that still surrounds the nature of creativity and also due to the fact that these tests were given under a variety of different administrative conditions, the term creativity is used in the sense of divergent thinking abilities, and the latter term is preferred. Divergent thinking capacity implies the level of ability to perform on open-ended tests which do not require one correct answer. Divergent thinking test results are those derived from answers from such open-ended questions which were evaluated in terms of:

(a) Fluency; i.e., the quantity or number of appropriate responses given, excluding repetitions.

(b) Flexibility; i.e., the number of different categories of responses or the number of shifts in thinking.

(c) Originality defined as the rareness, unusualness or uniqueness of response, operationalized as a statistical concept.

Thus, the working or operational definition of convergent thinking or intelligence was that it is an ability or collection of abilities required to find the one correct and presumably logical answer to a variety of questions on both traditional, standardized verbal and non-verbal tests. The working or operational definition of divergent thinking or creativity was that it was an ability or collection of abilities to generate as many and different and rare appropriate responses to a battery of open-ended verbal and non-verbal tests.

Since no tests were timed, fluency did not mean, in Guilford's 1967 terms, ideational fluency, which refers to the speed at which appropriate responses can be given, but merely referred to the total number of appropriate responses which could be given, excluding repetitions. The term flexibility merely looked at shifts in category of thought, thereby partially incorporating Guilford's idea of spontaneous flexibility which implies a shift from one class to another within the same larger superordinate category, and his idea of adaptive flexibility whereby there is a complete change of strategy, and a major change from one category to another is made. In other words, flexibility reflects the individual's ability to change set. A divergent individual would be expected to change set with ease, while his opposite counterpart would be expected to perseverate and exhibit rigidity with respect to departing from a certain pattern of ideas. Originality as defined was seen as a statistical concept in terms of rareness of response. Cleverness of a response was not measured, nor was originality seen in terms of transformations which are the alteration of a product of thought, cognition or information whereby such a product changes from its initial state to another state due to "changes, revisions, redefinitions or modifications," (Guilford, 1967, p. 64).

In terms of changing the context under which convergent and divergent tasks are given, Wallach and Kogan (1965) point out that,

> One context that our culture makes available for the definition of situations that are to be evaluation-free, concerns play and other activities engaged in for their intrinsic enjoyment. A situation closer to that context than to the context of a test or an examination hence would seem to be necessary if we are to assess adequately a person's ability to generate many associations and to generate many that are unique, (p. 19).

They are, of course, making special reference to creativity tests.

In their conclusions, Wallach and Kogan (1965) point out that in order for the expression of intelligence to be more reflective of creativity, one approach would be to administer orthodox intelligence tests in "a permissive, non-evaluative context," (p. 306). However, they speculate that such an approach would not have much of an effect on intelligence results because they consider that children are "probably sensitized to the readily perceptible cues that an intelligence test offers in abundance. One such fundamental cue is the convergent type of thinking (the one correct answer) demanded by intelligence test items," (p. 306). However, if intelligence tests were administered in a relaxed, game-conscious atmosphere in an unstructured way, as Lundsteen (1966) and Boersma and O'Bryan (1968) suggest, it is possible that the correlations between intelligence and creativity might be very much higher than those found by the above investigators when intelligence and creativity were given under test and play conditions, respectively.

Yet it is now a decade since Wallach and Kogan (1965) claimed that a better index of creative behaviour was obtained when children did these "tests" under informal relaxed play conditions. And, again, it is a decade since the question was initially raised as to whether the relationship between intelligence and creativity would be as distinctive as Wallach and Kogan had reported if intelligence tests were also given under informal conditions, and whether intelligence scores would be higher or lower if such tests were given in the same evaluative-free play-like context. Since 1965, there have been a number of studies which have investigated the effect of administration on the results from creativity tasks under formal or informal conditions, and under timed or untimed conditions. But there has been no systematic large studies done to examine the effect of administration on intelligence tests and their results on groups of people at any age level.

To summarize this section, therefore, the following statements can be made. The general purpose of this thesis was to determine whether there is a significant effect of the conditions of administration upon the performance of cognitive-type tests which were of two types: convergent and divergent. A related purpose was to see if the conditions of administration influence the level of correlation between convergent and divergent "tests."

The context of this research lies in the issues of the validity of measures of convergency and divergency taken under varying conditions of test administration. There has arisen in the research literature not only the question of the relative validity of convergent and divergent tests and of the construct of intellect, but also the extent to which performance on such tests is influenced by conditions of administration. Two types of conditions of administration are thought to be important: test conditions in which the student is aware of the evaluative aspect of the results and a formal structured relationship is established; in contrast a game condition may be established in which the students are by design relaxed, and the evaluation context is eliminated.

The importance of this research lies in the fact that since cognitive tests are so widely used, school and university administrators, educators and research psychologists must be better informed about the influence of the conditions of administration upon the level of performance and the validity of the tests. There is also the general scientific issue of the relationship between convergent and divergent "tests" taken under different conditions of administration.

In the next section, a comprehensive review of the relevant literature will be given, leading to the specific hypotheses and questions which were pursued in the research.

CHAPTER II

BACKGROUND

Convergent Thinking/Intelligence

Statement of Problem

The primary purpose of this thesis was to examine the effects of evaluative or play-like conditions on performance levels on certain cognitive tasks. These tasks concerned tests involving convergent and divergent thinking. However, the term convergent thinking is commonly used with intelligence, and the term divergent thinking is commonly used with creativity. Thus, it seems appropriate that this literature review should begin by what some other investigators have discovered about these two processes. First, for historical reasons, it is pertinent to discuss what has been meant by intelligence, and what its nature is in order to demonstrate that because of the controversy that surrounds this topic, the term convergent thinking may be more appropriate to use in studies looking at performances of individuals on tests which require only one right answer.

Definitions of Intelligence

According to Burt (1955), the term intelligence is derived from the Latin word "Intelligentia" meaning knowing or understanding. Psychologists often attempt to define intelligence operationally; that is, "intelligence is what intelligence tests measure," which refers to the fact that we have a scientific measure which we apply to people in order to access their ability in certain ways, relative to other people. Different tests may not measure exactly the same abilities, and correlations between tests fluctuate.

Although intelligence has been measured for over 70 years, psychologists have never formally agreed on a definition of it. Binet (see, for example, Binet & Simon, 1908) and the early group of testers regarded intelligence as a collection of faculties - initiative, adaptation to circumstances, etc. - which were located on the test scale according to the predilections of the test constructors. It has since been variously defined as a composite of the capacity to learn or to think abstractly, or functional adaptation (Freeman, 1962), or the capacity for abstract thinking (Terman, 1925; Terman & Merrill, 1937), or innate all-round cognitive ability (Burt, 1955), or a cluster of high grade mental skills (Hearnshaw, 1964) to note a few; but in the absence of clear definitions of abstract thinking, adaptation and the like, there has been an understandable hesitancy to equate such definitions with intelligence, or to judge the extent to which such skills or abilities are measured by current intelligence tests.

Much of the confusion arises from the following sources:

1. Most of the important definitions are derived from the fields of biological observation and animal experimentation (the ability to learn, for example); from the clinical study of the mentally retarded (the deficient capacity to think abstractly); or from educational measurement (the capacity to adapt means to ends). Neglecting for a moment that there is considerable disagreement over the meaning of the term adaptation (compare the views of Stern & Keislar, 1967, with those of Porteus, 1965, for example), the clinician will rightly object that the biological notion of intelligence as adaptability to the environment is far too broad to be useful in clinical measurement. Essentially, what differentiates low grade from high grade intelligence is efficiency in manipulating symbols. On the other hand,

investigations such as those of Harlow (1949, 1958) are sufficient to indicate that such a criterion is just as inappropriate in the field of animal experimentation since tasks involving even the rudiments of symbolic functioning are nearly impossible tasks for all except primates. It is equally obvious that neither criterion fits particularly well with the special demands of educational measurement. The amount of overlap between such views is very slight, and the theoretical interplay is negligible.

2. In recent years, there has been an almost complete switch from the qualitative study of intelligence and the theoretical construction of tests to the elaboration of statistical methods of analysis. Hearnshaw (1964) points out that whereas Spearman's work on the nature of intelligence has been largely ignored, his factorial methods have flourished. It is more than half a century since Thurstone wrote "The Nature of Intelligence" (1924), yet we still have no adequate current theory of intelligence, and test content is determined partly by administrative convenience, partly by analogy

with practical tasks, partly by uncoordinated intuitions, and partly by the vestiges of past theories. Psychologists have preferred to set up practical criteria derived directly from measurable activities of the individual, to measure and analyze differences in individual performance without concerning themselves with theoretical propositions. Intelligence has become a concept we find useful in describing human behaviour. It is used as a theoretical construct to describe mental operations involved in problem-solving, or as the most general cognitive ability. Such a development has been valuable, but the wholesale defection from the theoretical to an empirical standpoint has resulted in the repudiation of the need for constructive thinking about intelligence and its functioning.

3. There has been a widespread and careless use of the qualification "general" when referring to intelligence. Such is implicit in most of the Binet type or Wechsler "global" definitions, where test constructors are little concerned with the unidimensionality of their scales, but

are concerned merely with whether the tests work tolerably well in certain well-defined areas, notably that of educational prediction. Such an assumption of a general ability seems legitimate here, since it is clear that the ability required to solve problems of the type encountered in the Binet or Wechsler scale is roughly the same as that required to do well in business, at school or university; and that inability to do such tasks is indicative of mental deficiency or educational subnormality. There has been empirical verification of this; a considerable amount of evidence indicates a good relationship between success in school or college and high scores in intelligence tests with correlations generally in the vicinity of 0.60 to 0.65 (see for example, Butcher, 1968; Guilford, 1967; Vernon, 1971). Undoubtedly, many factors such as persistence, initiative, emotional stability, quality of teaching, etc., influence success, and these ought to be considered in any predictive battery.

Even such an empirical approach, however, leads inevitably to assumptions and theorizing about the nature of educational abilities and the structuring and identification of these abilities into narrower categories in adulthood. In view of this, there is little justification for psychologists neglecting theoretical issues and merely concerning themselves with the nature of correlations between tests and certain kinds of life success. As writers such as Butcher (1968) point out: from the theoretical point of view this is offensive; practically there is the point that intelligence or mechanical aptitude or verbal ability, etc., are never the sole factors in a real life situation, as distinct from a controlled experimental situation.

Another fundamental objection is that achievement in different areas may not be strictly comparable. School A, for example, may use a slightly different criterion from School B, group X from group Y, and so on. On the other hand, it is extremely simple to re-evaluate success from one situation to another in terms of a few unchanging reference factors such as clerical aptitude, typing aptitude, or mechanical aptitude, etc., for which there are tests of known reliability and validity.

Where the criterion area is clearly demarcated by special aptitudes (or abilities presumed to exist), or is only vaguely defined, the omnibus test is obviously of extremely limited value. The recognition in recent years of the need for a changed orientation in such situations has resulted in the production of a large number of tests yielding measurements of special aptitudes and/or abilities.

Kuhlmann (1939) and Garrett (1946), for example, approved the measurement of disparate functions, as in sub-tests of the Wechsler tests or Thurstone's PMA (Primary Mental Abilities) battery. For these investigators, from both the practical and theoretical standpoint, such an approach is vastly superior to that taken by makers of omnibus tests who hope, by averaging scores on a hodge-podge of functions, to obtain a measure of some worthwhile ability. This approach is essentially naive, it is neglectful of basic theoretical problems as that in which the sum of a conglomerate of tests is accepted as a single measurement of general intelligence. For example, is there a primary mental ability of fluency, or is the spontaneous flow of associations subsumed under the more general verbal factor (in other than very homogeneous groups)? For that matter, is it more appropriate to judge a person's general efficiency in a number of intellectual tasks using a profile of scores from a "Primary Mental Abilities" battery, or to judge from a battery of highly saturated tests incorporating different types of material? It is impossible to attempt to solve the problem of measuring special aptitudes and abilities without first solving the problems of the existence and the nature of general ability. Any sort of answer depends upon an embracing theory of intellect designed to serve as a basis for testing in all areas. Under present conditions, however, we have no assurance that our tests are even partially adequate, much less that they are complete.

Mapping the Intellect

The most important single step towards exploring the structure and organization of human abilities by breaking down the global concept into its unidimensional variables, has been that of factor analysis. On the question of the number, nature and scope of these primary components, however, there has been considerable disagreement between the American and British schools of factor analysis. Here is a quote by Wiseman (1973), an Englishman, that seems appropriate. On page 10 of his introduction to a book of readings on 'Intelligence and Ability,' he writes,

> We must undoubtedly acknowledge our great indebtedness to America for the massive scale of her research effort in psychology, and not least in the cognitive field. The development of psychology must have been very much slower and more hesitant without this powerful attack. But we must also recognize that the foundations on which it was built were - to a very considerable extent - British. . . .

Wiseman (1973) continues with the comment that since,

> our university and college libraries carry ever-increasing proportions of American textbooks in psychology and educational psychology, it is all the

> more important to emphasize this country's fundamental contributors. I believe that this is not mere chauvinism, but a necessary process in understanding some of the basic conflicts. . . .There are both American and British schools of thought in the cognitive field, and particularly so in theories of the structure of abilities. Eventually they will come together: but this will happen later rather than sooner if the peculiarly British contributions receive less than their deserved share of attention. If I were to hazard a prediction (and 'hazard' is the right word here) I would guess that the final solution - if there is to be one - will lie nearer the British than the American line of thought, (p. 10).

Most of the British factorists, represented most vigorously by the views of Burt, Thomson and Vernon, favor a hierarchial interpretation of ability factors. They accept _g_; i.e., a general intellectual factor. In addition, they recognize such group factors as _W_ meaning word fluency, _N_ meaning numerical ability, _K_ meaning spatial ability, _R_ meaning reasoning ability, and _P_ meaning perceptual ability or perceptual speed; since, in addition to _g_ and specifics, tests invariably measure some group factor abilities. The stability of these factors, however, varies greatly with the ease of the test, presuppositions of the factorists, and the selection of subjects. In addition, the form of the test items, and the conditions under which the tests are administered, seem to introduce formal factors, which have some effect on the operation of content factors. Many present day group factorists favor the view that after the removal of _g_, tests tend to fall into two main groups:

a. the numerical -- verbal -- educational -- or _V:ed_ factor.

b. the practical -- mechanical -- spatial -- physical or _k:m_ factor.

These are considered to be two very strongly unified groups, not easily broken down into separate _V_, _N_, _K_, etc. factors until the onset of adulthood, with its attendant differentiation and structuring of abilities into narrow minor factors under the influence of selective occupation, education, and so on.

Most group factorists consider that most of the variance of ability is due to _g_ and to highly specific or small group factors. Vernon (1971) maintains that _g_ variance should account for appreciably more variance than all the other group factors together, and that the emergence of large group or primary factors is due essentially to selectivity of the sample. He reports that, when the same tests which, among unselected recruits, gave _g_ and group factor variances of 50 and 20-25 percent, respectively, were analyzed amongst a more highly selected group, _g_ often fell to 15 percent and the group factors rose to 35 percent.

These would correspond roughly with the general size of the factors extracted under similar conditions by the group factor method.

Many of the factorial studies after 1935 followed the multiple factor methods developed by Thurstone (see, for example, Thurstone, 1938, 1948; Thurstone & Thurstone, 1941). Such methods were in reality an extension of the simple summation techniques originally advanced by Burt (1909, 1917). Thurstone subdivided intelligence by factor analytical techniques into 10, later reduced to 7 primary factors. These are:

a. number ability;

b. word fluency (concerned with fluency and dealing with single and isolated words);

c. verbal meaning (concerned with grasping ideas and word meanings);

d. reasoning;

e. spatial relations;

f. perceptual speed;

g. memory.

These tests are each designed to measure primary abilities and little else, but they are not the only factors in intelligence. However, they are the most clearly identifiable ones from all the tests Thurstone conducted. Several other factors have recently been identified and some of these represent a further breakdown of Thurstone's reasoning factor. From these factors identified by Thurstone, a single general factor emerged, called a second-order general factor, probably equal to Spearman's g factor. This second-order factor is essentially linguistic, because it correlates most highly with verbal ability and word fluency. The highest correlation was between the general factor and reasoning.

In other investigations, however, such as those of Guilford and his colleagues (see, for example, Guilford, 1956, 1959, 1967, 1972b; Guilford & Hoepfner, 1963, 1966) the general factor is renounced and the variance attributable to g is spread over the remaining factors by a process of rotation to simple structure or components. By applying the statistical method of factor analysis to his experimental data, Guilford (1959) claimed to have discovered 50 intellectual factors, although his theoretical model predicted 120 factors.

Guilford called this system, whereby he organized the intellect into primary abilities, "the structure of intellect." Every cell in his theoretical model represented a unique component or factor, which in turn represented a particular and unique ability. Thus, Guilford would claim intelligence is made up of at least 120 discrete, distinct and different abilities, each of which are required in a certain class of tasks. Although factorially distinct, these components, according to Guilford, can be classified into three different categories, because they are related to each other in certain ways. Such classification is made according to:

1. The type of process or operation performed on certain tasks. Such classification yields five groups of abilities involving factors of cognition, memory, convergent thinking, divergent thinking and evaluation.
2. The type of content involved which may be figural, symbolic or semantic. To cover the general area of "social intelligence," Guilford had theoretically proposed, without empirical evidence, a fourth type of content which he called behavioural.
3. The type of product involved which occurred when a particular type of operation was applied to a certain type of content. Six kinds of products are involved, which Guilford labelled as units, classes, relations, systems, transformations and implications.

Each factor, representing a certain ability can, therefore, be described in three different ways.

Subdividing intelligence into so many different abilities, and labelling such abilities as Guilford did, can give rise to a certain confusion in terminology. For instance, Guilford predicts 24 convergent production abilities which involve tasks requiring logical deductions or inferences where "the input information is sufficient to determine a unique answer" (Guilford, 1967, p. 171).

This differs from the more general definition of convergent given in this thesis, where convergent thinking is that involved in any task where only one right answer is required. Thus, for example, a vocabulary test involving verbal comprehension, where giving the meaning of words is required, would be classified by Guilford as the ability to cognize semantic units. Guilford would claim that such a vocabulary test involves the <u>operation</u> of cognition factors, semantic <u>content</u> factors which together generate a <u>product</u> factor of units. While not denying the different components involved in any one intellectual task, such as a vocabulary test, the opposing school of thought would maintain that these different abilities have common to them one underlying general ability.

Indeed, the wide divergence between the two factor analytic interpretations may be more apparent than real.

The strict demands of the single discrete structure solution are too rarely met in matrices of observed correlations to be of much significance as long as the orthogonal pattern of factor axes is retained. Thurstone, for example, permitted his factors to be correlated; and once the factors in the matrix are found and rotated, another matrix of correlations is found between the factors. This, in turn, can be analyzed and in the cases of tests of general ability, generally obeys the tetrad criterion, giving rise to a second-order factor which corresponds very closely to Spearman's g.

As Eysenck (1939, 1967) has pointed out, within the multiple factor framework the simplicity and orderliness of Spearman's original picture may be recovered, without neglecting the additional group factors which made his simple model seemingly inapplicable.

There also have been other experimental results which have shown that factors discovered through the simple structure solution correlate in a hierarchial pattern and that the emergence of a general factor is prominent (Vernon, 1971). Guilford's "structure of intellect" model is also under sharp scrutiny and Cronbach (1970a, 1970b) in reanalyzing the "structure of intellect" model (albeit in only limited areas, so far) claims to have found high correlations between certain factors which Cronbach concludes reflects the existence of general group factors. Although Guilford (1972b) defended his model, it would appear that at least some of his factors may not be as unique as originally claimed and that distinguishing a number of orthogonal factors from test matrices does not exclude the possibility of one or more broader underlying common factors.

The existence of a general factor is now generally admitted, albeit reluctantly by many American factorists. The question then becomes why should there be such divergence of opinion about its importance, its size in different experimental populations, its generality and its relationship with much narrower primary factors?

Vernon (1971) claims that investigations finding a small amount of g do so because they test homogeneous samples.

> It is because the majority of American investigations are conducted with college students, aircraft pilots, high school pupils and other selected groups, that their results so readily fall into independent primary factors, instead of g and group factors. But when more heterogeneous adult groups have been studied, a g has usually appeared, (p. 31)

It is clear that the selection of the sample has an important effect on the nature of the factors that are eventually extracted from the test battery. The correlation found between tests, or between tests and external criteria, depend to a very large extent on the heterogeneity of the sample. An increase in heterogeneity increases the correlations and consequently affects the size of the factor loadings. Thus, Thomson (1939, 1951) went so far as to express doubt as to whether separate factors are meaningful in the absolute sense since they are apparently dependent on the population from which measurements have been taken.

Vernon (1971) concluded that invariance of the underlying factor structure could be shown to exist despite sampling differences. He based this on the fact that many American investigations have produced invariant primary factors rather than g and group factors such as v:ed and K:m. These samples, however, have generally been selected from college, university, armed services, etc., type populations, and where more heterogeneous populations have been studied, g has generally appeared as a common factor.

The present author would agree with investigators such as Vernon (1971), that it is imprudent and mathematically difficult to belittle "g" which seems the best hypothesis for theoretical, empirical and practical purposes. Such a hypothesis was assumed in the present research.

The Nature of the General Intelligence Factor

Spearman (1904) introduced the concept of "g" or general intelligence. He introduced the term to refer to the correlations which exist in varying degrees between the most diverse sorts of cognitive performance.

Finding that non-IQ tests not only correlated with each other but also correlated with IQ tests, he assumed that there was a common ability which enters into all test performances and should be substituted for the concept of intelligence. Three of his conclusions are, to quote him, the following:

> By this same new system of methodics, there is also shown to exist a correspondence between what may provisionally be called 'General Discrimination' and 'General Intelligence' which works out with great approximation to one of absoluteness . . . this phenomenon appears independent of the particular experimental circumstances; it has nothing to do with the procedure selected for testing either Discrimination or Intelligence, nor with the true representativeness of the values obtained by these tests, nor even with the homogeneousness of the experimental reagents; if the thesis be correct, its proof should be reproducible in all times, places, and manners - on the sole condition of adequate methodics.
>
> . . . all branches of intellectual activity have in common one fundamental function (or group of functions), whereas the remaining or specific elements of the activity seem in every case to be wholly different from that in all the others . . .

> As an important practical consequence of this universal Unity of the Intellectual Function, the various actual forms of mental activity constitute a stably interconnected Hierarchy according to their different degrees of intellective saturation, (Spearman, 1904, as quoted in Wiseman, 1973, p. 6).

According to Butcher (1968), the second of these quoted paragraphs is the first statement of Spearman's concerning his two-factor theory. The third statement indicates that each cognitive task will have a specific factor loading on the "g" factor or factor of general intelligence, and such a loading will indicate to what extent general intelligence plays in that particular mental task. Butcher (1968) states the following regarding this:

> According to this celebrated 'two-factor theory,' which might have been called a 'one-factor theory' (since it depends on the existence of only one common factor), the performance of every cognitive task depends only on general intelligence and on another factor entirely specific to the particular task, (p. 45).

Later, Spearman and Wynn Jones (1950) admitted that the two-factor theory "only indicates the initial degree of analysis; certainly not the ultimate," (p. 10). They acknowledged the existence of group factors which influenced some, but not all tests, but maintained that these group factors differed from "g" which influences all tests.

Spearman actually preferred not to identify g with intelligence but rather suggested it depended on a general mental energy with which every individual is endowed. He considered *g* to be the amount of general mental energy, and the specifics the efficiency of specific mental energies. The general factor was regarded to be innate and ineducable while the specifics were regarded as being largely affected by education and training.

High *g* tests display the characteristics that involve seeing relationships or what Spearman labels as "the eduction of relations and correlates."

In addition, Spearman dealt at some length with the question of whether *g* involves the power of abstraction. He pointed out that intelligence as measured by usual tests is essentially characterized not by the nature of the content cognized, but rather by the fact that this content is cognized abstractly. Much of the evidence for a greater intercorrelation between abstract, than between perceptual tests, and therefore the greater *g* saturation of the former, is spurious, in that the level of difficulty in the two types of tasks is not equalized. When this is allowed for, the average correlations are shown to be about equal.

The Nervous System and Intellectual Processes

Thomson offered an alternative theory to explain the zero-tetrad differences that Spearman took as evidence of a single general factor. He suggested that each cell calls upon a sample of the "bonds" a mind can produce, or on a sub-pool of the mind's bonds. If two tests are very much alike, they may be fairly described as sampling the same region of the mind.

What these "bonds" of the mind are Thomson does not state. He ventures the suggestion that they include inherited responses such as reflexes, acquired habits or associations, and that they are associated with the neurones or nerve cells of the brain. Thomson describes thinking as the excitation of these neurones in patterns in a continuum from the simplest instinctive patterns to the more complex acquired ones.

Intelligence is probably a function of the number and complexity of the patterns which the brain can make. Intelligence tests, however, do not call upon brain patterns of a high degree of complexity, for tests are always associated with acquired knowledge and with the educational environment of the subjects tested. Tests may differ in their richness or complexity and if a miscellaneous set of tests, or tests of less rich material are given, the extensive sampling of the bonds which is expected to occur, probably results in the positive correlations. Such positive correlations do tend to appear, particularly if the bonds tend to be all-or-nothing in their nature, as the action of neurones is known to be.

For concreteness, it is convenient to identify the elements, on the mental side, with something in the nature of Thorndike's "bonds" (Thorndike, 1931) and on the physical side with neurone arcs. Thomson is inclined to make a distinction between Spearman's g and the various other common factors, mostly, if not all, of lesser extent than g. The former measures the whole number of bonds rather than a single unitary power or organ of the mind, as Spearman suggested; the latter indicates the degree of structure amongst the bonds. Some of this structure is, no doubt, innate. But more of it is probably due to the environment, education and experience with life. The actual organization of sub-pools is likely to be extremely complicated, and its categories are probably interlaced and interwoven.

From Thomson's theory, therefore, g is not solely a fixed, inherited quantity, but represents "the total number of bonds" (Vernon, 1971, p. 33). While these bonds may be largely dependent on inherited properties of the central nervous system they are also modifiable by experience, schooling or occupation, by acquired organic conditions such as brain injury or even ageing, and by deterioration of mental efficiency in certain psychopathological conditions.

The physiological aspect of the organization of bonds is clearly developed in Hebb's attempt to arrive at a system which is faithful to present-day physiological facts. His theoretical model, which is applied initially to the analysis of the development of simple visual perceptions, is elaborated to account for the actions of attention, motivation, learning and intelligence in the cerebral control of behavior (Hebb, 1949, 1972). Activities are aroused by the cell assembly, the basic variable, and the phase sequences which result from the successive actions of a series of cell assemblies which mutually interact. Hebb refers to these as conceptual or ideational activity, and the various phenomena of behaviour are analyzed in terms of this general schema.

For Hebb, intelligence has two supplementary components which he labels as Intelligence A and Intelligence B. Intelligence A represents the individual's innate potential, his inherited capacities and is defined in terms of neural metabolism and brain structure. Intelligence B represents the developmental component, is the actual level of intellectual function and is dependent on the individual's past or present experience and training. While the relative importance of these two components or factors varies with the particular intellectual test, Hebb maintains that both factors are involved in any test performance. Intelligence B is dependent on Intelligence A and is limited by it for the latter represents our inherited potential. Intelligence B, being dependent on both hereditary and environmental influences, represents the observable lasting changes in perceptual, intellectual and behavioural organization which occur during development and are induced by the first factor, Intelligence A.

Particularly interesting from the psychometric point of view is Hebb's interpretation of data dealing with mental capacities and brain damage. He generalizes to the extent that localized brain damage in children produces a general impairment, rather than specific functional loss, while in adults a corresponding injury produces a more specific loss of powers and a less obvious general impairment. The question arises as to whether the latter represents g. Hebb tentatively suggested that the greatest deterioration occurred on speeded tasks, abstract problem-solving and unfamiliar tasks. The deterioration occurred least in vocabulary, information and verbal comprehension, such tests having been shown to be most highly loaded with g in an adult population. This seems to imply the presence of a purely general factor which increases to maturity, is dependent on the organization of the higher nervous system, and is associated with the action of the cerebral cortex.[4]

Vernon develops the point that factors over and above *g* or general intelligence, although partly due to hereditary and neurophysiological influences, arise primarily as a consequence of an individual's upbringing and education. Such consequences, as he puts it, impose "a certain grouping on his bonds," (Vernon, 1971, pp. 31-32). The organization of abilities is in part culturally determined. Since performance on tests is also in part culturally determined, Vernon added a third dimension to the concept of intelligence which he called *Intelligence C* (which he added to Hebb's Intelligence A and B). Intelligence C is meant to be a particular sample of Intelligence B which is measured by a particular intelligence test, the performance level being culturally bound.

Ferguson (1954) would expand such a notion even further maintaining that not only does the level of performance on a particular task differ from culture to culture, depending on the learned cultural dictum, but, as a logical outgrowth of this, it is likely that the factor structure of any one test will also considerably differ from one culture to another. It is expected that for any one identical problem, individuals from diverse cultural backgrounds will utilize very different abilities to find a solution.

The Development of Intelligence

The importance of the purely developmental rather than cultural influences on g and the broader group factors is nowhere more explicitly stated than in the genetic system of Piaget. The psychology of thought or intelligence, which he regards as synonymous, treats, in the process of growth, the ordering and dimensioning of a child's experience from the egocentric and phenomenal world through the stages of sensori-motor intelligence and the pre-conceptual stage to the final development of abstract conceptual thought (Piaget, 1950). In the course of growth, the internalized actions attain a continually greater degree of differentiation and organization, mobility and combined ability, until, at the full operational level, complete generality, power of abstraction and reversibility is reached. This is the level of intelligence proper, where the child can effectively manipulate and deduce formal logical relationships, can freely put them together, multiply them with one another, link and unlink them, substitute, invert, and in general, exploit all their formal possibilities. This is similar to Spearman's eduction of correlates.

Thus, intelligence for Piaget is a continuous construction process. Intellectual development also involves active, rather than passive, processes whereby the individual attains different and higher order forms of mental organization of structures termed "schemas." Schemas then are the intellectual products developed by two interrelated processes which Piaget called assimilation and accommodation. Assimilation involves an incorporation of environmental data or information into the individual's own mental organization, whereas accommodation involves the modification or application of the individual's inner mental organization to particular environmental stimuli. It is the product of these two processes that result in more effective schema, the psychological unit of intelligence.

While maturation, experience with the environment and learning or education are all regarded to be important factors in fostering intellectual growth, the most important force is what Piaget calls equilibration. Equilibration is an auto-regulatory process whereby the individual strives to reach for balance or homeostasis between contradictions or conflicts which an individual experiences with himself or with the environment. Although Piaget does not talk in terms of general or specific factors of intelligence, his concept of intelligence involving the building up of ever-increasing mental units or schema seems similar to those theorists advocating broad general intellectual abilities underlying all cognitive behaviour.

Rather similar in vein to Piaget's point of view is that of Bruner. However, unlike Piaget, Bruner emphasizes the importance of cultural influences, since to a great extent intelligence is the "internalization" of the actions, the skills, the images and the symbols of a given culture, (Bruner, 1964, 1966). Combined with his concern with the importance of culture, Bruner (1964, 1966) was also concerned with the evolutionary aspects. Intellectual growth is seen as the way individuals "gradually learn to represent the world in which they operate - through action, image and symbol," (Bruner, 1966, p. 6). Such representations of the world can only be understood with reference to those "tools" which exist in the culture which serve to amplify the individual's powers and with reference to evolution which provide man with his heritage.

The Search for Factors of Intelligence

So far in this discussion, the concept of a general factor of intelligence has been accepted. Investigators like Thomson differentiated between and emphasized both the neurological and developmental aspects of *g*. The neurological aspect is suggested by positing a system of neurone functioning and patterning as the neurological basis of a bond structure. The developmental aspect is put forward by noting that *g* is the number of bonds the mind is capable of synthesizing, and that the sub-pools, represented probably by the group factors are structured systematically be education, training and similar influences.

The cultural and purely developmental influences affecting the organization of bonds have been stressed by Vernon and Piaget, respectively. Bruner has also promoted the importance of cultural factors. The neurophysiological aspects have been particularly stressed by investigators like Hebb. His concept of initial activity aroused by the cell assembly, the organization of phase sequences and the tentative acceptance of a general ability factor associated with the action of the cortex is rather close to Thomson's theoretical foundation.

Although the stance in this chapter has been to support the concept of general abilities and broad group factors, the existence of specific abilities is not denied. The limitations of the approach such as that of Guilford and his colleagues, where intelligence is broken down into discrete components is stressed, since the final validity of this comprehensive work is doubtful.

Vernon (1971) summarizes the following major weaknesses of this discrete component approach:

> 1. There is no good proof of the independence of anything like such a large number of factors, even in highly-selected groups. Most of the separate researches cover only some half-dozen of the new factors at a time, and if even a dozen could

be studied simultaneously (together with reference factors), it is probable that several would coalesce or mutually modify one another. Again, if broad verbal, spatial and numerical (i.e., 'material' factors) were first removed, one might hope that the three columns would usually coalesce into one. Guilford does not deny that there may be second-order factors running through sets of several of his listed factors, but he has not yet published any study of these because they are liable to vary so markedly with the selectivity of the tested population.

2. No other laboratory or research institution seems to have been convinced of the validity of Guilford's scheme, nor (with few exceptions) to have used his factors as a basis for fresh experimentation. And although the consistency of findings from one research by Guilford to another is quite striking, investigations by others seldom provide much confirmation. Several large-scale studies . . . have yielded results which can be only partially reconciled with one another or with Guilford's classifications. A less elaborate scheme based on fewer, more distinctive, factors might gain wider acceptance, show greater stability from one research to another and greater practical utility; though Guilford would no doubt answer that it would give a less complete picture of the complexities of intellect.

3. There is a serious dearth of external validatory evidence to show that the new factors give additional information about thinking in everyday life. Certainly this is difficult to come by, but one would hope for proof that each new factor could contribute to the selection of people with thinking

> capacities needed for particular jobs or courses of study. . . . A small-scale research by Hills (1955) into the relations between college mathematics grades and 9 tests highly loaded on 9 of the factors led to the disappointing conclusion that 'there is no particular ability or set of abilities or traits which is universally associated with success in mathematics.' Certain tests appeared to be predictive of some courses at one institution, but not of similar courses given by other instructors, or in other institutions. Single tests are not, of course, the same as factor measurements. But, until some external or 'real-life' meaning can be attached to more of the factors, the criticism can hardly be refuted that they represent not so much thinking abilities as abilities to do the various kinds of psychological tests, (pp. 144-145).

This does not mean, of course, that the search for intellectual factors should desist, but rather that the inquiry into their part in some general cognitive framework should not be neglected. Wechsler (1958), for instance, had this in mind, in constructing his own battery of tests.

With respect to the number and importance of factors, Wechsler (1958) criticizes a model such as Guilford's where large numbers of factors are extracted. Wechsler (1958) claims that the primary purpose of factor analysis is to account for the major variance in a battery of tests with the minimum number of main factors. Those components which contribute a large part of the variance should be regarded as either general or broad factors depending upon the amount of variance each contributes while those that account for less than 3 percent can be regarded as representing specific abilities which may be of less importance.

Guilford (1967) criticizes this approach to limiting the number of factors, not only on the grounds that components accounting for only a small part of the variance may be of great importance in intelligent behaviour but because on scientific grounds any factor is important. In addition, the apparent splitting of narrower factors from an apparent broad factor may be a misconception since the latter may be "a confounding of a number of basic factors," (Guilford, 1967, p. 36). While this is obviously a valid statement it is also true to say that isolation of numerous factors must also be balanced with studies that examine their possible relationship, in order that the configuration of the map of the intellect can be better understood.

Since this contextual atmosphere was part of Wechsler's rationale for constructing his intelligence tests, which suited this investigator's purpose of measuring general convergent ability, two of the sub-tests from the Wechsler battery were used in this investigation. It is worthwhile, therefore, to examine the issue of what Wechsler meant by intelligence.

Wechsler (1958), on viewing the literature on this subject to that date, came to the unhappy conclusion that psychology was in the paradoxical position of having devised and used tests for measuring intelligence and then disclaiming their relationship to the subject matter in hand by asserting that nobody really knows what intelligence means. Wechsler claims that the main difficulty lies in the fact that general intelligence, like the concept of energy is not a tangible material entity but a construct which is limiting and abstract. For this reason, Wechsler is not concerned so much with what intelligence is, but rather what it involves and distinguishes or differentiates.

He accepts that intelligence includes having the ability to learn, to adapt or to educe relationships, but claims that it also involves much more. Intelligence manifests itself in a variety of ways as in learning, adapting or reasoning tasks, but in addition there should be a degree of commonality between those forms of behavior which one identifies as intelligent. In defining intelligence, Wechsler points out that three points should be considered:

(a) that identifying basic factors or common elements of intelligence constitutes only part of the problem in a definition;

(b) that general intelligence is not synonymous with intellectual ability; and,

(c) that general intelligence is not a discrete entity but is rather part of the total organism of which the personality is an integral related component.

Intelligent behaviour is goal-directed. Necessary conditions, therefore, for such behaviour are purposiveness and motivation. Keeping these things in mind, Wechsler (1958) then defines intelligence as

> . . . the aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his environment. It is aggregate or global because it is composed of elements or abilities which, though not entirely independent, are qualitatively differentiable. By measurement of these abilities, we ultimately evaluate intelligence," (p. 7).

However, intelligence is not synonymous with the total sum of abilities. Although intelligence is a function of a number of abilities, intelligence also reflects the configuration of these abilities or the way they are combined. An excess of any one ability may not effect intelligent behaviour as a whole. In addition, factors other than intelligence *per se*, such as drive, play an important role in intelligent behaviour. But Wechsler found that the only way to measure intelligence quantitatively was to assess aspects of abilities, although he felt that the overall process can reflect the factor *g*. This, then, was to be validated by his statistical analysis. Because he found that significant correlations existed between items, Wechsler claimed that these items were not independent of one another. However, Wechsler does warn that psychometricians should desist from confusing the identities of general intelligence and intellectual ability. In this way the measurements are not invalid and through such tests we can learn the effects of intelligence, what it does, although not necessarily the exact nature of it.

Thus, the hypothesis assumed in the construction of the Wechsler intelligence scales is that intelligence is the ability to exercise mental ability in contextual situations, situations that have content and purpose as well as form and meaning.

Wechsler significantly disclaims the power to measure all that makes up general intelligence in his tests. But the thing Wechsler does claim from an intelligence scale is that it measures sufficient portions of intelligence to enable us to use it as a fairly reliable index of the individual's global capacity. As we have seen, not all investigations into intelligence have taken this holistic approach. Yet, at the same time it should be appreciated that other contributions in addition to the ones already mentioned have also added significantly to our knowledge of this aspect of cognitive function. Some of these will now be considered.

Historically, Binet and his colleagues made a major breakthrough in the field by providing us with tools by which individual children could be compared with each other, and the normal could be distinguished from the mentally retarded or abnormal. But Binet's major contributions were to that of measurement rather than that of theory. His concept of mental age provided psychologists with a meaningful measure by which a score could be interpreted with reference to other members of the population. Although he defined intelligence as the ability for goal-directed behaviour, the ability to make adaptations to changing circumstances, or, as mentioned earlier, a collection of mental faculties which include initiative, judgment and so on (Binet & Simon, 1908), the criterion for measuring such was largely subjective.

However, Wechsler (1958) claims that one of Binet's greatest contributions lay in his assumption that from a battery of tasks, selection of the specific task was of less importance as long as it was a measure of general intelligence.

> This explains in part the large variety of tasks employed in the original Binet scale. It also accounts for the fact that certain types of items which were found useful at one age level were not necessarily employed at other age levels.

> More important than either of these details is the fact that for all practical purposes, the combining of a variety of tests into a single measure of intelligence, ipso facto, presupposes a certain functional unity or equivalence between them, (Wechsler, 1958, p. 10).

This functional equivalence of test items then makes it crucial that each subtest has its own statistical validation.

Such an approach, of course, assumes the existence of underlying general factors. However, the difficulty with intelligence tests is that different tests do not measure the same things in any very precise way, nor do all intelligence tests measure just one kind of principal ability. Those belonging to the school of factorists would stress the idea that the existence of one general ability is not a self-evident necessity, and that other alternatives are possible, whereas those supporting the existence of broad general factors would claim that although specific abilities exist they are related to more general factors. The difficulty with factor analysis is that it does not provide us with a unique description of what is present in the matrix table but rather provides us with the amount of variance which can be attributed to certain spatially related sets or factors. Because an indefinite amount of mathematically equivalent answers can be found, other criteria, besides those of statistics have to be used, such as those to determine what is psychologically meaningful.

Hearnshaw (1951) was critically aware of the incompleteness of the factorial approaches in mapping human abilities. He defined intelligence as "a cluster of high-grade skills concerned with problem solving" probably, though not necessarily, at the symbolic level.

The three basic intellectual skills, Hearnshaw describes might usefully be considered properties of g, though this may be taking undue liberties with the author's views, as he makes no implications of a unitary faculty. He maintains that a theory of intellect will be derived from learning theory, and that intellectual skills are differentiated from learning skills by a minimization of repetition and a maximization of relevant generalization.

Hearnshaw claims that a number of important intellectual skills are ignored in present intelligence tests. Three of these skills are:

1. The capacity to relate or integrate events over a period of time. Such skills involve the temporal, rather than the spatial, integration of experience, relating such experiences to the past and a future, such skills requiring at least the rudiments of symbolic functioning. He notes that in all

present intelligence tests, the materials are presented simultaneously, rather than sequentially, for example, there is not one of the tests of the Thurstone PMA battery where the time dimension is involved, excepting perhaps in the rote memory tests. The suggestion is made that tests ought to involve the successive presentation of data. A number of such experimental tests were later described (Hearnshaw, 1956) where data was presented successively. Interestingly, Hearnshaw found that such tests are not highly loaded with a general ability or a verbal ability factor.

2. The ability to form concepts, to make relevant generalizations, to order or classify experiences and events. For Hearnshaw, most current tests were poor indicators of the processes of conceptualization, particularly novel conceptualization, and of the abilities to grasp principles of order.

3. The ability to appreciate the relevance or significance of experiences or events. This

involves a progressive sophistication leading through richer schemata of organized experience to the attainment of judgments or "wisdom." The obvious question is whether a broad general capacity is involved in all judgment situations, or whether such a cognitive act is specific to types of experience. Again, we have the problem of how meaningful are specific factors.

Burt (1958) points out that factors are useful categories to use when classifying results. Other evidence, beyond mere statistical unity is required, however, in order that we can regard them as real entities. Furthermore, Burt (1955) regarded such mental activity as ability which was general, cognitive as opposed to being affective or conative and innate. The fact that the term general cognitive ability was used reflected his findings that the highest common factor in a battery of tests was a general one. He found

> in nearly every factorial study of cognitive ability, the general factor commonly accounts for quite 50% of the variance (rather more in the case of the young child, rather less with older age groups) while each of the minor factors accounts for only 10% or less, (Burt, 1958, p. 5).

He concludes that

> for purposes of prediction - forecasting what this or that individual child is likely to do in school or in afterlife - the general factor is by far the most important, though admittedly not our only guide, (Burt, 1958, p. 5).

Burt (1955, 1958) also discusses another of his initial propositions, that of the innate or genetic component of intelligence, a feature we shall return to later.

One final word must be said about Burt's approach to the existence of general factors. His approach does not mean that he refuted the significance of other factors. Rather, he claimed that intermediate factors are required to provide a complete account of the structure of abilities. His theory of cognitive ability embraced four main kinds of factors which are arranged in a hierarchical manner. Thus, at the top of the hierarchy was the general factor, followed by the major group factors which were common to large groups of performances (namely, the verbal-educational factor or verbal abilities and the spatial-mechanical factor or non-verbal abilities). Next in the hierarchy were the minor group factors which still had something in common with the major group factors, and finally, there were the numerous specific factors, those factors which were specific to a particular performance and made up of true and error variation.

Although the greater part of the variance was, as already stated, contributed to that produced by the general factor, it is by no means insignificant that the factors produced in this theory here are orthogonal; i.e., uncorrelated. Yet, at the same time as Burt (1958) points out even if we resort to so-called second-order analysis, as Thurstone, himself, attempted, we do end up with a so-called "super-factor" which can only be interpreted as a factor of general intelligence.

Other psychologists have attempted to circumvent the difficulties presented by using the general term intelligence and to overcome the difficulties by arguing to what extent this intellectual capacity is inherited or the result of learning and environmental experience. They do this by splitting the concept and distinguishing different meanings of intelligence. We have already discussed the contributions of Hebb and Vernon, particularly Hebb who likes to distinguish between two different kinds of intelligence, Intelligence A and Intelligence B. As Burt (1969) points out, such a differentiation only serves to indicate that there are "two different ways of using the word intelligence," (p. 199). Not unreasonably, Burt prefers to talk of "genotypic" and "phenotypic" manifestations of intelligence, qualities reflected in the way Cattell and his colleagues conceptualized intelligence (Cattell, 1963; Catell, 1967; Horn & Cattell, 1966).

The theory of fluid and crystallized general intelligence denies that there is only one factor of general ability, as originally put forward by Spearman, but rather that there are what (Cattell (1963, 1967) calls two broad factors. Fluid intelligence as revealed in factor analysis is that ability found on typical subtests of traditional intelligence scales and is associated with "insightful performances in which individual differences in learning experience play little part," (Cattell, 1967, p. 209). Fluid intelligence, then, is similar to Hebb's concept of Intelligence A (Hebb, 1949, 1972) in that it is innate or inherited and unmodifiable. Crystallized intelligence, on the other hand, represents those abilities acquired by learning, environmental experiences, cultural opportunities and interests. Cattell (1967) defines it as "the culturally acquired judgmental skills," (p. 209), which are reflected in and measured on tasks which are heavily loaded on verbal and numerical skills. A parallel with crystallized intelligence can be made with Hebb's (1949, 1972) concept of Intelligence B.

Cattell's approach, however, differs in an important way from those other theoreticians who split intelligence into an innate component and an acquired component.

Innate intelligence or Intelligence A, as defined by

Hebb, is by definition inaccessible to present measurement while Intelligence B or the Intelligence C of Vernon (that which is actually measured) may have little relation to Intelligence A. Cattell's approach, on the other hand, claims to be able to evaluate both factors which go to make up this total double structure of intelligence, and which are factorially separable. Importantly, Cattell (1967) emphasizes that while his fluid and crystallized general abilities are functionally independent they are also significantly correlated with an r of 0.47. Attempts, therefore, to fit factors such as verbal-educational ability and spatial mechanical ability into Cattell's findings of fluid and crystallized intelligence are inappropriate since the former are orthogonal factors and the latter oblique. Vernon (1950, 1971) points out that the most representative tests of fluid intelligence are those containing problems involving matrices, abstractions and similarities, which are all highly loaded with g, and the most representative tests of crystallized intelligence are those involving the v:ed and k:m factors, such as comprehension, vocabulary and arithmetic. Because of this he finds little justification going beyond the factor hierarchical model previously described.

Cattell (1967) justifiably attacks Vernon's stance. Directly referring to this endeavour to fit these two broad abilities to the concepts of fluid and crystallized intelligence he states,

> . . . this attempt to creep from one conceptual position to the other, without clear rejection of false assumptions is merely asking for confusion. The resemblance of the two dualities is only superficial and gross. Fluid and crystallized abilities are oblique factors: _k_ and _v:ed_ are orthogonal. The _gf_ and _gc_ theory supersedes and abolishes the need for the Spearman _g_ altogether, whereas the _k_ and _v:ed_ theory retains _g_. The crystallized ability factors, _gc_, covers alike mechanical, practical, numerical and verbal skills, whereas the _k_ and _v:ed_ divide them into two different categories. There are, besides, a whole set of connotations of the fluid and crystallized ability theory in physiological, cultural and developmental fields which do not belong to the theory of the _k_ and _v:ed_ dichotomy, (p. 223).

With reference to this last statement it should be noted that Cattell and his colleagues did include measures of personality in much of their research (see, for example, Cattell, 1967) and did find, for instance, connections or relationships between personality variables and cognitive measures.

Speed as a Factor

One issue that has not yet been discussed with reference to intelligence is that of speed. Eysenck (1967) from his own experimental findings and of those of Furneaux

(1960) concluded that the basic factor in measured intelligence was that of speed; or put more accurately, the effect of speed in finding a correct answer when considered in the context of individual test item difficulty was found to be the main difference between individuals on cognitive tasks. Thus, in his model of the structure of intellect Eysenck (1967) included the components of mental speed and power while at the same time retaining in his hierarchical structure a central concept of g.

With reference to the speed factor Eysenck (1967) emphasizes the findings of Furneaux where the rate of increase in time taken with increasingly difficult tasks (as measured by the slope of the plots between the log of time or latency and increase in item difficulty) was constant in all individuals tested. Since the findings of constants is not common in psychology, Eysenck suggests that further work along these lines would increase our scientific knowledge of intelligence. The following discussion will illustrate, however, that the significance of the role of speed in cognitive tasks, whether they be convergent or divergent in nature, is still controversial. Thus, as discussed in the section on methodology, speed as a dependent variable was eliminated from this investigation.

As children become mature they become quicker in their motor and intellectual responses until young adulthood, after which they remain capable of quick responses (given good health) until middle age after which there is a gradual slowing. The relevance to intellectual abilities of this developmental trend over the life span is by no means agreed upon. Some researchers, because of their conceptions of the nature of the intellect, have no room for a "speed factor" in explaining the organization or changes in thought processes.

Since one reason for being slow in response in early childhood is because of lack of familiarity, it is easy to believe that practice not only "makes perfect" but it also makes one quick. It is difficult to refute the argument that one becomes quick with practice whether it is a matter of doing arithmetic problems, reading or skipping rope. In this sense, the time taken in doing a task is a variable which is dependent upon the level of familiarity or skill. There is another sense in which time might be involved in which one would be more skillful if one were quicker. Here time or speed is cast as an independent variable. This paradigm is not commonly discussed in relation to children and as a result time is most often seen as a consequence of ability or skill. In the case of older adults it is less easy to slide over the issue since many older adults seem to slow in response processes, independent of the level of practice or use. If one wishes to be rigid and retain time or speed in the role of a dependent variable one may say that the additional time taken by the older adult is compensatory for a reduction in ability. Implicitly there is the belief that there is change in a "performance factor" in which the thought processes are unchanged and it is but a matter of the individual's not being able to register responses to the outside world as quickly.

One of the early workers in the field of intelligence, Thorndike (1926) considered the intellect as having the properties of power, altitude and speed. His followers did little to elaborate on the role of speed and it became essentially ignored, perhaps understandably so since its importance seems to diminish over the years of development toward adulthood. The issue was again raised, however, by data on the later years of life with the embarrassing findings that speed measures are intercorrelated and also correlate with untimed measurements of cognitive processes (Birren, Botwinick, Weiss & Morrison, 1963). It was suggested by Birren (1964, 1974), Botwinick (1973), and Botwinick and Storandt (1974) that perhaps timing and speed of response was a reflection of the integrity of the central nervous system. Elevated blood pressure and related diseases are also associated with slow responses (Birren & Spieth, 1962; Spieth, 1965; Hicks & Birren, 1970; Abrahams & Birren, 1973).

Time or speed, rather than being an artifact of measurement or a manifestation of a peripheral change, began to assume major significance in the differential assessment of the mature and older adult (Birren, 1965; Spieth, 1965; Birren & Renner, 1977; Botwinick, 1977). Thus, the conceptual turn-around consisted of not partialling out time and discarding it, but rather giving it major conceptual signifiance and emphasizing its measurement. If speed and timing present a good scientific vantage point for studying abilities and the state of the central nervous system of older adults, why should not speed have an important role in the thought processes of children and young adults as well?

The evidence on mature and older adults suggests that there is a general speed factor of considerable magnitude. While perhaps not explaining as much of the total common test variance as "g", speed may be in the next order of magnitude of signifiance when judged from the viewpoint of explained variance. The lack of a large or clear speed factor in children while in part explained by the predilection of investigators not to look for it may also result from the fact that it may be small as a limiting or determining factor for the quality of thought in children. Reaction time in children, for example, tends not to correlate highly with intellectual tests (Eysenck, 1967).

Perhaps one is dealing with a neural phenomenon in which, should slowing be excessive, the slow scanning of current or past information could result in a decreasing probability of eliciting appropriate material from long-term store. Also, slowness may limit the ability to combine the material into a new percept before it disappears from the short-term store. Slowness appears also to be related to a higher probability of distraction which is observed in very young children or in retarded older children. The essential point of relevance here is the fact that given slow response processes beyond some critical level, our memories may seem to diminish when in fact it is a matter of speed of retrieval rather than storage. Also, the probability of occurrence of new relevant thoughts may be affected by the slowness, and with slowness the probability of inappropriate responses may rise. This, of course, does not negate those findings in which an increase in speed of performance is achieved at the expense of accuracy, but the decline in actual performance with decreased rate of response has now been documented (Eysenck, 1967).

Eysenck (1967) as has already been stated, presented a model of the structure of the intellect in which he included a speed factor. Unfortunately, he did not attend to the empirical literature on intellectual functioning of older adults and he did not build a bridge between information on intellectual functioning in early life to middle age and then to the later years. This gap is to be regretted, for it would appear that the measurement of the speed of mental processes represents a fruitful avenue of study for convergent thought processes and perhaps more so for divergent thought processes. With respect to the latter, the present state of the organism may be more critical in forming a high quality response than for convergent processes in which retrieval from store may be less embarrassed by sluggish neural processes than would be the formation of a high quality original response.

From this present line of reasoning, one should expect to find a relationship between divergent thinking and some estimate of general response speed. Although speed of response was not investigated in this study, the author recognizes the significance attached to speed of response, particularly with respect to older adults and indeed one should, in the future, expect profitable research in this area.

Genetics, the Environment and Intelligence

Cattell in his definition of intelligence talks of two meanings of the term, one of them basically concerned with genetic potentiality (namely, that of fluid intelligence) and the other the result of an interaction with experience, learning and environmental factors (namely, that of crystallized intelligence). Hebb and Vernon have likewise talked in terms of these two components although using different terminology.

Butcher (1968) points out there is a certain appeal to concepts which are explained by reference to neurological constructs, particularly since we may be nearing the time when quantitative assessment of neural structures would permit us to calibrate the growth and decline of performance on cognitive tests.

Hebb (1949) was one of the pioneers in citing experimental evidence to indicate that neural pathways were involved in learning and that also environmental experience plays a significant role in establishing an individual's level of performance. Those advocating the use of intelligence tests frequently do so by arguing that such tests not only measure innate capacity but also learned experience which is relevant to the assessment of scholastic aptitude and prediction of scholastic success in the school curriculum. In addition, the literature on readiness and neural sets (Hebb, 1949, 1972; Vernon, 1958) suggests that there are neurological limits to what any individual is intellectually capable of. Early environmental stimulation certainly plays a significant role in determining these upper limits, first by dictating the optimal amount of stimulation the arousal system, such as the reticular formation and hypothalamus, can cope with per unit of time, and secondly, through largely unknown connections that are formed in the cerebral cortex. Such neural learning patterns and neural learning boundaries are set at quite an early age. It has been proposed that readiness to take on new experiences or the ability to take on new experiences is limited by the time we get to secondary school. At this point of time we are apt to build on old foundations.

This does not mean that intelligence should be regarded as imposing a fixed level of performance that an individual is capable of or as being the sole determinant in setting the rate of mental growth. Rather, as Vernon (1958) argues, intelligence test levels and educational attainments alike exhibit marked stability or constancy, and because of the limits set by neural structures, mental growth is essentially regular and predictably cumulative. This point of view is also proposed by Piaget (1950) although the biological explanations are given in a different terminology.

Vernon (1958) further argues that the interaction of potential capacity and environmental stimulation have consolidated in the child a certain level of functioning by the time he has reached the age of five or six. This attained level of ability will essentially determine his rate of progress in his education. Quality of teaching, environmental influences, personality problems, emotional adjustments, interests and motivation all play a role in cognitive growth, but the results from Vernon's work suggest such factors have only caused major alterations in a minority of students. In a series of follow-up studies with English secondary school children over a period of five years, correlations of intelligence tests with English and arithmetic attainment tests, with teachers' ratings, and with scholastic performance, were between 0.80 and 0.86.

Arguments in favor of a physiological or genetic basis of explaining individual differences in intelligence have been postulated for a long time. Some of the best known attempts to determine the relative importance of genetic and environmental factors contributing to intelligence were done by Burt (1955, 1958). Using the school population of London, Burt and his colleagues were able to test large samples of many families drawn from residential institutions and private case files of the London County Council. The Stanford-Binet Intelligence Test was used, and if after consulting the child's teacher a test score appeared to not be representative of the child's ability then such children were retested and an adjusted score was given.

The following table adapted from Burt (1958) demonstrates the percentage of variance found in the genetic and environmental components (see Table 1).

Random environmental effects were those sources of variance not related to genetic components. These include the cultural, social and educational influences on the individual as well as pre-natal and post-natal biological influences, such as illness and nutrition.

Systematic environmental effects were those environmental influences which were correlated with genetic effects and which Jensen (1969) refers to as the covariance of heredity and environment. This was the component, which in Burt's results was most affected by making an adjustment, reflecting the fact that the IQ test is not culture free and that a better assessment is made when a teacher's opinion is sought and a corrected score given. Unfortunately, as Burt (1958) himself points out, such intensive inquiries were too expensive to be carried out in all cases and he admits that reliance on tests alone is not the most satisfactory method of assessing an individual's innate ability.

The genetic factors are composed of what Burt called "fixable" and "non-fixable" components. He adapted these terms from the terminology used by Sir Ronald Fisher. Fisher named the fixable component as "the essential geno-

Table 1

Analysis of Variance for
Assessments of Intelligence*

Source of Variation	Unadjusted Test Scores (Percent)	Adjusted Assessments (Percent)
Genetic component:		
Fixable	40.51	47.92
Non-fixable	16.65	21.73
Assortative mating	19.90	17.91
Environmental components:		
Systematic effects	10.60	1.43
Random effects	5.91	5.77
Unreliability or test error	6.43	5.24
TOTAL	100.00	100.00

*After Burt, 1958, p. 9.

types" which are those genes which are passed unchanged from parents to offspring. Jensen (1972) refers to this as the "genic" or "additive" variance which is "attributable to gene effects which are additive; that is, each gene adds an equal increment to the metric value of the trait," (p. 106). Such effects are responsible for similarities between families.

The "non-fixable" components are considered to be the percentage of variance due to dominance and similar genetic influences. As described by Jensen (1972), it is that variance due to the observed "systematic discrepancy between the average value of the parents and the average value of their offspring on a given characteristic," (p. 108). Some genes are what are described as recessive, which means that the phenotype is not expressed unless one such gene is paired with another recessive gene for a particular characteristic. If a recessive gene is paired with a dominant gene then the phenotypic effects of the former is "dominated" by the latter. In other words, the phenotypic expression is always under the control of a dominant gene if the dominant gene is present in either the homogeneous or heterogeneous state.

The assortative mating component is that variance due to the recombination of genes for a specific characteristic, in this case those genes from each mate which are associated with intelligence. Jensen (1972) claims that IQ levels show "a higher degree of assortative mating in our society than any other measurable human characteristic," (pp. 106-107). After analyzing the literature concerned with this topic on studies done in Europe and North America he estimated that the average correlation between intelligence scores of spouses was 0.60, which incidentally is higher than that found between siblings, namely, a correlation of 0.50. Jensen cites Eckland (1967) to explain that this high correlation between marriage partners is not so surprising in our developed industrialized societies where selection of mates is heavily influenced by such factors as the contacts made within the educational system and occupation. Indeed, Jensen points out that it is the very structure of our social environment, educational and occupational selection as well as the occupational hierarchy, which serves to reinforce or consolidate the strong genetic influence on cognitive functioning.

Burt's results, as presented in Table 1, certainly lend strong support to these claims. On unadjusted test scores 22.94 percent of the total variance is due to environmental influences including that variance contributed to unreliability or test error, and 77.06 percent of the variance is due to genetic factors. When adjusted test scores are used, the amount of variance due to genetic factors increases to 87.56 percent and the variance attributed to the environmental components is reduced to 12.44 percent. As previously mentioned, the main reason for the reduction in percentage variance contributed to environmental influences was due to the decrease in the score of the covariance of heredity and environment.

Although there have been a great many studies done on the contributions made by both genetic and environmental influences on intelligence, the issue really came to a head with the work of Jensen (1969, 1972).

The topic is of pertinence in this investigation because of the influence the arguments from both the geneticists and the environmentalists had on the way this research conceptualized convergent and divergent thinking. One of Jensen's main contributions was to emphasize the importance of the inheritance of intelligence (Jensen, 1969, 1972). Referring to the general factor common to all intelligence tests he proceeded to show how 80 percent of the population variance in intelligence can be attributed to genetic factors and 20 percent to environmental influences (which included both social and biological, pre-natal and post-natal influences). These figures were arrived at as an average, after Jensen had made his very comprehensive review of all the evidence.

Jensen develops his arguments about the large genetic contributions to intelligence systematically. With regard to compensatory education, he believes that it had largely failed because of an over-dependence on two theories. The first of these was that most children were alike in their mental capacities and mental growth; that differences between individuals in scholastic performance were due to environmental influences such as family background, pre-school experiences, out-of-school influences, motivation, interests, and so on. The second theory, the social deprivation hypothesis, claims that some children lack the crucial pre-requisite experiences for successful academic performance which many educators had felt could be remedied with pre-school and compensatory education. Jensen is not arguing that there is no truth in either of these theories, but rather that there are other more fundamental differences between individual cognitive abilities which would require a rather different approach by educational authorities.

In relation to the concept of intelligence, he points out that measurement and operational definitions of intelligence originated in the school setting and were primarily intended for scholastic purposes. Such a practice, Jensen estimates, has performed its task adequately for traditional forms of schooling and for the majority of children. Such tests have also aided in selecting those children who were likely to fail in school so that they could receive remedial education. That there is a heavy bias in using such tests for academic purposes does not mean that such tests measure nothing of psychological importance. Indeed, Jensen claims that contrary to the beliefs of some we can measure intelligence.

What then does he mean by intelligence? After reviewing the literature, Jensen (1972) claims that while _g_ is a "hypothetical construct" designed "to explain covariation among tests" it should be regarded as the "nuclear operational definition of intelligence," (p. 77). The existence of _g_, while not an actual entity, should be regarded as "the factor common to all tests of complex problem solving," (p. 77).

Instead of viewing _g_ as some unitary ability or as a collection of different "sub-abilities" Jensen conceptualizes it as a source of variance, differences between individuals on measures from different tests which have a high degree of commonality. There are many components that make up general intelligence, but at the same time common to all the different components which make up intellectual abilities there is a general factor.

Jensen also points out that there is a close relationship between the occupational structure in society and its educational system, both of which are highly correlated with IQ tests. Occupational assortment of persons may not be fair but it is an objective reality which reasonably we can cope with by striving for equal opportunity processes and selection according to true merit. Although Jensen does not mention this, the present author would add that determining the best context for administering tests at different ages, as well as selection of appropriate test material, would also aid such a process.

After analyzing the correlations reported in the literature on intelligence scores taken on any one individual at two different points of time, Jensen found, while not constant, the IQ increased in stability throughout childhood.

Thus, Jensen (1972) confined the term intelligence to "the general factor common to standard tests of intelligence," (p. 88) particularly characterized by those abilities which require "abstract reasoning and problem solving," (p. 88). There are many mental faculties and it was stressed that intelligence does not cover all mental abilities, although, as previously pointed out, it does correlate highly with educational attainment and certain types of occupational success.

From here, Jensen (1972) proceeds to present his concept of heritability which he defines as "the proportion of phenotypic variance due to variance in genotypes,"

(p. 114). He emphasizes that it is a population statistic rather than a statistic derived from individual measurement, and that its use is to account for the proportion of variance in a population that can be accounted for by genetic and environmental components. The phenotypic variance depends on both genetic and environmental conditions. Each of these variance components Jensen carefully describes. The fact that this population statistic is empirically determined means that it is subject to such things as sampling error, the population samples, and so on, all of which Jensen points out. It might also be expected to vary from time to time in the population.

Jensen then proceeds to analyze the empirical findings associated with the heritability of intelligence concluding, as mentioned earlier, that on an average approximately 0.80 of the variance is due to genetic and 0.20 due to environmental components. Although Jensen feels that the bulk of evidence is more consistent with a genetic hypothesis rather than with an environmental hypothesis, he is quick to point out that such results do not exclude the influence of environmental factors or the interaction of such factors with genetic ones.

As part of his reports, Jensen (1969, 1972) describes some of his own empirical results indicating that there are differences in ability between lower and middle class children. Very briefly, the following are a description of these results.

Jensen maintains that his heritability index is perhaps the best objective criterion for reducing the amount of culture-load on tests. Assessing the rate of a child's learning ability in the context of a new learning situation was also felt to be a more culture-free way of assessing intelligence than assessing what a child had learned prior to the testing situation, as is the situation on traditional intelligence tests. He found, however, that the criterion of culture-free mode of assessment, such as the above, was not always a useful way of distinguishing psychological tests since some of the lower socio-economic children did well on high culturally loaded tests but did not do well on some of Jensen's learning tasks or on low culturally-loaded tests of g.

Jensen hypothesized that explaining these IQ differences between social classes could be done by using a two-dimensional concept of intelligence, each of which required different basic abilities.

Level I ability he called "associative" learning ability which he measured with the use of such tests as digit memory, simple serial learning and learning of paired associates. Such tests he reasoned showed how well a child can learn something entirely new in a testing situation where no prior experience was an advantage. Low socio-economic class children with IQs between 60 and 80 were found to do better on these tasks than middle-class children with the same IQ range. Above an IQ of 100 there was no difference in Level I learning ability between the two social classes.

Level II learning, called "conceptual" learning, involving such abilities as problem solving and formation of concepts was measured with such tests as the Raven's Progressive Matrices which have high g and low cultural loading. (Because of these characteristics the Raven's Progressive Matrices was one of the measures used in the present study.) Tests involving Level II learning did show marked social class differences in Jensen's samples, with middle-class children doing significantly better. Level II learning abilities have a high correlation with traditional intelligence test scores.

Making an overall analysis, Jensen found that Level I learning ability correlated well with IQs among middle-class children, but poorly amongst lower class children. Furthermore, the scattergram plot analysis of these results suggested that the arrangement of these abilities is hierarchical and that associative learning ability, although a necessary condition for conceptual learning ability, is not sufficient.

From his heritability studies on Level II ability tests Jensen (1972) concluded that,

> Level II processes are not just the result of interaction between Level I learning ability and experientially acquired strategies or learning sets. That learning is necessary for Level II no one doubts, but certain neural structures must also be available for Level II abilities to develop, and these are conceived of as being different from the neural structures underlying Level I. The genetic factors involved in each of these types of ability are presumed to have become differentially distributed in the population as a function of social class, since Level II has been most important for scholastic performance under the traditional methods of instruction, (pp. 199-200).

Unfortunately, Jensen's results have been too often linked only with differences between social classes and races. It is true that he did point to features of selection processes which he hypothesized have resulted in genotypic as well as phenotypic differences between social classes and within races. Furthermore, he discussed the subject of differences in gene pools of different races and he did collect and analyze data that indicated that environmental arguments purporting to explain differences in intelligence were inadequate. However, these scholarly considered topics were never interpreted by him as advocating social or class segregation educational policies. Indeed, as Jensen (1972) himself takes pains to point out, he never made the claim

> that the high heritability of intelligence _within_ either or both racial groups was sufficient to prove that mean differences _between_ the groups was attributable, in whole or in part, to genetic factors. It is axiomatic in quantitative genetics that _within_ group heritability cannot prove _between_ group heritability. The relationship is one of probability or likelihood, that is, the higher the heritability of a trait _within_ each of two groups, the greater is the likelihood that a mean difference _between_ the groups has a genetic component and the smaller is the likelihood that the group difference is attributable solely to environmental variation, (p. 29).

Another way of evaluating his interpretations of his results would be to initiate future enquiries that would investigate two groups of children, each of whom would learn different things and be exposed to different assessment or learning methods. Jensen is arguing against a uniformity of educational system which he feels is not in many children's best interests, and promoting a diversity of approach in assessment, teaching methodology and goals. Earlier in his arguments he had emphasized that there was a diversity of mental abilities of which general intelligence or _g_ was one large part and one which was highly correlated with our present scholastic and occupational systems. Even if one does not completely agree with

Jensen's conclusions, there is enough substance in his empirical findings and analyses to warrant further research.

To, in part, determine the inheritability of intellectual abilities one can look at whether environmental context can change test performance. Thus, one can determine whether different modes of administration will influence levels of performance on intelligence tests, and extend this beyond the ability factors traditional intelligence tests usually measure to include factors such as creativity or divergent thinking, abilities Jensen did not include in his deliberations. Expressed another way: the question can be asked as to whether the way in which we assess children will influence their performance on a variety of intellectual tasks. Although manipulation of administrative conditions has been attempted on divergent thinking tasks, it has not been attempted with convergent thinking tests and thus this became a central focus in the present research.

Jensen's contributions also influenced the present investigation in another vital way. When we examine the feasibility of the use of variables, as described above, on educational practices, the investigator needs to take into consideration possible sources of variation between social classes and race. Accordingly, such sources of population variance seriously influenced the sampling procedure utilized in the investigation described in Chapter IV of this thesis.

Arguments Against the Genetic Models of Intelligence

There have been many attempts to criticize the Jensenist position. Most of his basic tenets, however, seem to have remained intact, certainly the ones that have been described here. In order to make the story complete, however, some of the main arguments against Jensen should be mentioned. They are merely outlined here since, unlike the previous work described, they are less central to the thesis research problem. These opposing schools of thought can be grouped into four opposing viewpoints as presented and reviewed by Rowell (1975, personal communication).

First, there are those who argue from the environmentalist position claiming that physical and social experiences or circumstances can have a major modifying effect on intelligence. Hunt (1961, 1969) is a good representative of the environmentalist position. In particular Hunt objects to the apparent exclusion in Jensen's work of evidence from social psychology and from biological experiments on the influence of early experiences of animals or infants on intellectual processes.

For any concept of general intelligence, Hunt would prefer greater emphasis on the importance of early experience, stating that this is implied in the changes that occur in IQ values in the first four years of life. Hunt (1969) views intelligence as "a cumulative, dynamic product of the ongoing informational and intentional interaction of infants and young children with their physical and social circumstances. . . ." (p. 284). By this interaction, Hunt emphasizes, particularly, that between biological and social circumstances. Considerable evidence is cited to show actual changes in neural structures and improved performance with enriched environments. Central to the environmentalist's position is the degree of plasticity in behavioral development of the organism, particularly in the early years.

Unfortunately, many of these arguments, by resorting only to studies of the effect of early environmental experiences, deprivation experiments, or animal analogies appear to have left many of Jensen's conclusions unchanged, including the question of the variance contributed to heritability and his findings on Level I and Level II abilities.

A second group of theorists attempt to maintain an eclectic point of view, pointing out the problems associated with accepting the Level I/Level II approach to ability without attacking their possible "existence." Smoliez (1972), for instance, points out several difficulties associated with Jensen's findings without making judgments as to whether Jensen's conclusions are right or wrong.

Smoliez points out that if we are going to consider genetic differences we need to consider the causes of natural selection which operated on different groups. For the American Black, the forces of natural selection were historically toward physical strength rather than mental ability, in order that they could survive and breed. This came about because of many generations of enforced slavery. Formerly, the Jews in their struggle to survive may have found that intellectual pursuits were more necessary for their survival than physical strength. Smoliez cites the number of Nobel Prize winners who are of Jewish origin.

Smoliez, then says, however, that such speculations need to be supported by comprehensive empirical evidence which he feels is still lacking, but he also points out the other side of the coin that Jensen establishes in his study, that to account for differences in IQ between individuals in terms of environmental factors is erroneous. Not only that, but the genetic differences between social classes increases when social class barriers are removed from society for several generations, permitting social mobility and so-called equality of opportunity. Thus, "a liberal-democratic society . . . (quite paradoxically), tends to maximize genetic differences and minimize environmental differences as a basis of social and economic rewards," (Smoliez, 1972, p. 48).

Smoliez quotes studies which show that there are educationally or occupationally relevant abilities besides those measured by the IQ test to indicate that Jensen's Level I and Level II abilities have support elsewhere, although not in the same language. Such studies, including that of Jensen, do emphasize the importance of accounting for a broad spectrum of abilities in the school situation.

The issue is then taken up as to whether the diversification of aims and approaches that Jensen proposes could be achieved in a comprehensive school so that associative learners are rewarded under a different scheme than that for conceptual learners. Smoliez feels that it would be very difficult to instigate a system whereby each group of children received equivalent educational rewards and even greater difficulties would arise with the question of social rewards. However, the opposite point of view is also expressed; i.e., providing equal opportunity means that those who are associative learners are more likely to suffer frustration and feelings of defeat when their educational and occupational aspirations are dashed because of their failure to compete adequately with conceptual thinkers.

However, the main concern of Smoliez is not the empirical evidence, although this is important, but with what empirical evidence is selected and what values guide us in the making of judgments or in finding implications. The question is can we make choices of value and apply them in our scientific work or should the scientific problem and its solution be completely separate from whatever may be the social and ethical implications? For Jensen, this was not the issue. His primary concern was presumably as an empiricist.

For Smoliez, however, this is a very important issue and one on which he criticizes Jensen. He describes first how value directed opinions are made about Jensen's work. Both the progressives and conservatives can, for example, use Jensen's results to their advantage. The former group would use them to support an elitist position while the conservatives would use Jensen's findings to provide evidence for their proposals of maintaining stability of the social system and discouraging that social mobility which does not have a sound scientific basis.

In addition, Smoliez (1972) feels that while the extent to which the question of values permeates theoretical models is still a highly controversial issue, "there is general acceptance of the need for social scientists to clarify their theoretical presuppositions so that readers are aware of them before they begin to interpret experimental observations and findings," (p. 60). He feels that Jensen's findings have severe limitations because Jensen failed to do this and also because Jensen neglected to take into account such issues as the effects of the children's emotional needs during test performance, their aspirations or expectations of the tests and the effects of test climate on results when dealing with children from different social and racial backgrounds. Smoliez accounts for this neglect because he feels Jensen expected that the non-biological factors would be "normal."

A "pre-empirical" model of man can be made for Jensen's work, although this was not Jensen's concern. Smoliez (1972) concludes that, in addition to his previous comments, it is because of this failure to present such a model, to not take into account "a historical reality" (p. 67), that the theoretical and practical implications of Jensen's work will be limited.

The present author finds such a conclusion rather harsh. Where it is possible, a theoretical framework is desirable for any empirical study, yet hypotheses can be formulated and scientific observations made without necessarily having to making pre-empirical judgments of the individuals studied. Valid experimentation may vary from the unstructured descriptive investigation to the highly structured study with causal hypotheses and the use of multivariate methods of analyses. Further elaboration on this topic has recently been given by Birren and Renner (1977). It is in the interpretation of these results that caution must be taken. The empiricist should not go beyond what his results have told him, and the limitations of his conclusions should, if necessary, be stated. That is implicit in the empirical method, although it is often forgotten by experimenters. Jensen, however, did, throughout his text, state his awareness of the limitations of his work, but this feature is rarely admitted by his critics. Indeed, too many of the critics of Jensen's work have made their own interpretations of Jensen's results in light of their own biased views of man.

A third approach to evaluating Jensen's work, as represented by the Piagetian viewpoint, provides a possible reinterpretation of the position in terms of developmental stage as well as proficiency within any given stage (Rowell, 1975, personal communication). Piaget's theory has already been briefly discussed. The relevant points here are that Piaget conceives man within a biological evolutionary framework; and development, including intellectual development is an inherent, unalterable evolutionary process which contains an element of absolute continuity, an unfolding of progressively higher mental processes in a definite pattern. Each level of development has its roots in a previous phase. In this evolutionary, unfolding biological process cognitive structures are continuously created into a new and higher cognitive order or intellectual form (Piaget, 1950, 1973; Flavell, 1963; Maier, 1969). As Furth (1973) points out: "For Piaget intelligence is not a content but a mechanism of individual construction," (p. 72).

Piaget is thus opposed to considering intelligence in terms of mechanisms involving genetic, environmental or learning forces. A score on a cognitive test reflects the stage or phase of intellectual development the child has reached. For Piaget the important force in attaining a new stage of intellectual development is equilibration, by which process earlier contradictions are resolved by the individual actively moving toward a higher thought process in a self-regulatory manner.

Riegel (1973) criticizes Piaget on this very issue, claiming that such a view of cognitive development involves closure of thinking. He would prefer to conceptualize disequilibriums or contradictions as the bases for new ideas rather than as obstacles an individual has to overcome. While following Piaget in much of his theory, Riegel claims that there is a final stage of cognitive thought beyond that of formal operations which is the stage of dialectical operations. This involves that thinking which is involved in the mature, creative person who is able to resolve contradictions using this to go on to other contradictions. In addition, such a reorganization of thought Riegel claims runs through the other four stages of development during childhood and adolescence.

Such a conceptualization is an advance on the work of traditional psychometricians and Piaget, in one sense at least. It does take into account the growth of divergent thinking as part of mental development, a feature which the above schools largely neglect, and in so doing broadens the concepts of intelligence. Unfortunately, Riegel does not discuss the issue of divergent thinking beyond saying that such thinking is an example of dialectical thought processes.

Moessinger (1977) makes an answer to Riegel's criticism of Piaget by pointing out that whereas Riegel uses dialectics in terms of ideas, as well as in terms of development, Piaget is concerned only with cognitive development. For Piaget disequilibrium is the motivating force or triggering device to reach higher thought processes. Moessinger (1977), therefore, says that for Piaget "the real source of progress is to be found in re-equilibration in the sense of a better equilibrium, and not in the sense of a return to the previous equilibrium - the insufficiency of which has led to a conflict," (p. 182).

But for both Piaget and Riegel, the issue of the conceptualizing of intelligence is developmental and as Elkind (1969) points out when referring to Piaget, the question is not how much heredity or environment contributes to mental ability but to what extent cognitive processes are "autonomous" (p. 330) from these influences. Rational processes are, for Piaget, the most autonomous and thus Piaget calls such processes intelligence. Nature and nurture factors play a regulatory role but their contributions to intellectual ability are conceived of in a different manner from the other theories. For Elkind (1969) the psychometric and the Piagetian-type approaches should be seen as complementary concepts both providing "useful starting points for the assessment and interpretation of human mental abilities," (p. 330).

A fourth approach to criticizing Jensen actually attacks the basic assumptions of Jensen which give rise to Level I/Level II abilities, thereby removing the significance of some of Jensen's educational implications (Rowell, 1975, personal communication). Rohwer (1971), for instance, presents a case against Jensen from within the same conceptual model. The questions he asks are whether intelligence tests are an accurate measure of a child's learning ability, and whether Jensen's model sufficiently explains individual differences on intellectual tasks.

Some of Jensen's tests in his experiments used measures of learning ability, such as paired-associate learning for assessing Level I ability. Rohwer also used the paired-associate method. Although this method has been traditionally viewed as measuring only elemental processes in learning, Rohwer cites evidence to show that more complex mental activity is involved, such as "ingenuity" or cleverness, and that such tasks are also related to scholastic achievement among heterogeneous groups of children.

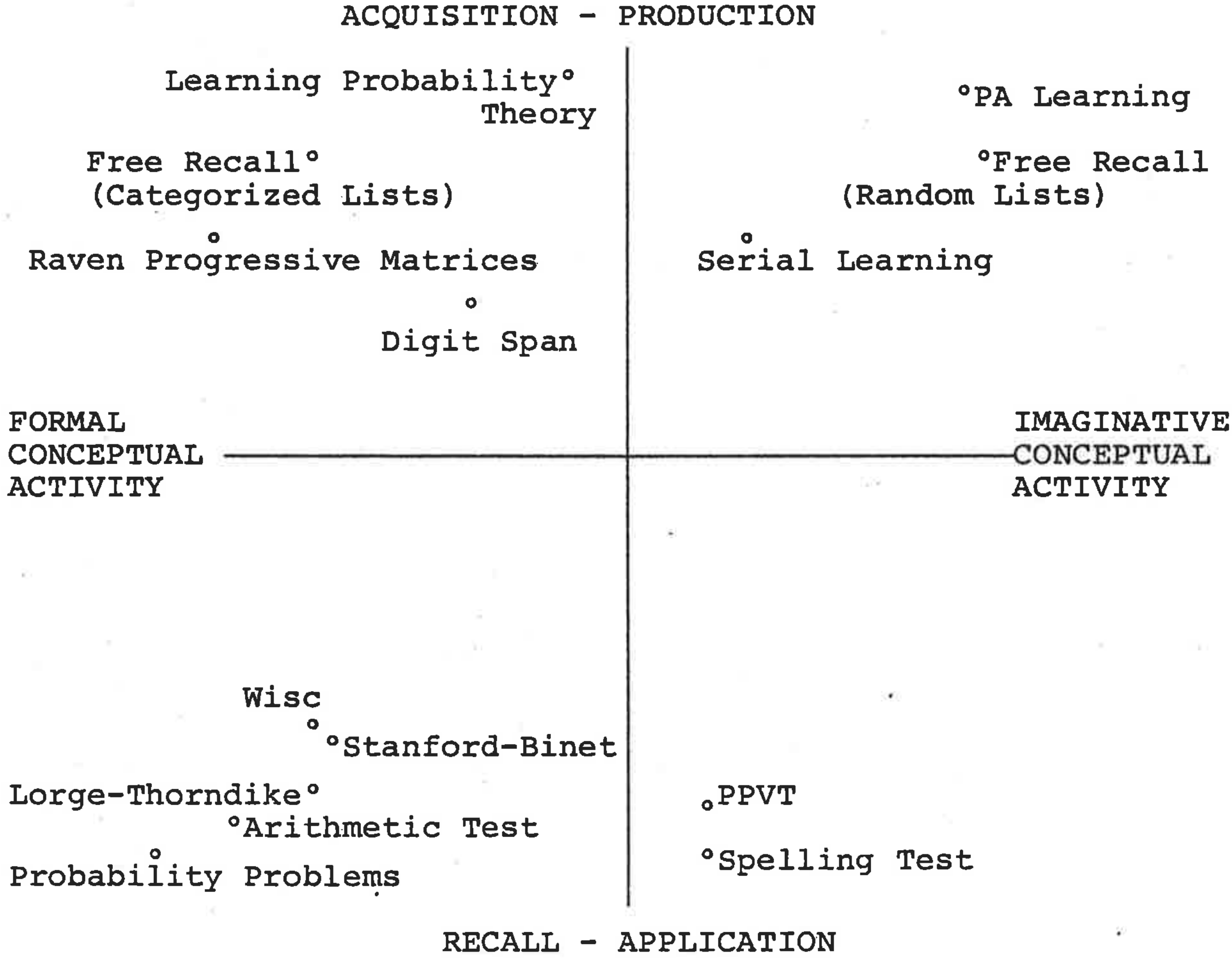
Using samples of high socio-economic strata (SES) white children and low (SES) Negro children from two different grades (kindergarten and grade 1) each child was given a picture vocabulary test, the children's form of the Raven's Progressive Matrices and two lists of paired associates to learn. On both the IQ tests there was a significant difference between the two groups, the white high SES sample doing very much better. The older white children also had higher results than the younger white children. In contrast, for the low SES Negro children there was "a cumulative deficit in intelligence test performance" with age (Rohwer, 1971, p. 198). However, between the white and Negro groups there was very little difference on their paired-associate test results.

Rohwer claims that paired-associate learning is an index of learning ability and that these results show that low SES Negro children are almost as accomplished in this ability as high SES white children. The difference in IQ scores means that "Negro children have not learned as much as white children. But these scores do not indicate that Negro children cannot learn as much as white children," (Rohwer, 1971, p. 198). Rohwer (1971) criticizes Jensen's explanation of differences between such white and Negro SES groups on Jensen's models of Level I and Level II abilities by claiming that paired associative learning requires not associative but conceptual ability. In analyzing the learning abilities required on paired-associate tasks and citing considerable evidence Rohwer (1971) concludes that "such processes involve transformation of input" which are "conceptual in character," (p. 201). Furthermore, Rohwer analyzes the digit-span task which Jensen found was an example of associative learning ability and demonstrated that contrary to Jensen's findings there were substantial SES differences on this task.

To account for this evidence Rohwer offers a modification of Jensen's model by postulating his own two-dimensional model to classify intellectual tasks. Rohwer's model is presented in Figure 1. It bears many similarities to the present investigation in the way it conceptualizes

Figure 1

A Two-dimensional Model for
Classifying Intellectual Tasks
From Rohwer (1971, p. 203)



Notes:

PA = Paired Associate

WISC = Wechsler Intelligence Scale for children

Lorge-Thorndike = Lorge-Thorndike Intelligence Test

On both the WISC and Lorge-Thorndike regular SES differences have been observed

PPVT = Peabody Picture Vocabulary Test

intellectual tasks.

In describing this model, Rohwer first makes the assumption that whatever the task may be, conceptual (not associative) abilities will be needed to master it successfully. From this, Rohwer (1971) makes two distinctions. In terms of the task, he distinguishes between those that "require the recall of information or the application of skills acquired previously," and those that "require the acquisition or production of new information or skills," i.e., the "recall-application" tasks versus the "acquisition-production" tasks, (p. 202). The second distinction was made between the type of conceptual activity involved, whether this be "formal conceptual activity" or "imaginative conceptual activity." Formal conceptual activity is "characterized in terms of a set of formal, explicit rules that . . . permit little deviation from their culturally agreed upon form," (Rohwer, 1971, p. 202). Such a definition of formal thinking is similar to the definition used in this thesis for convergent thinking. An example of such a task is the Raven's Progressive Matrices. Imaginative conceptual activity involves diverging or departing from the conventional, or as Rohwer (1971) puts it "departure from highly formalized conventions," (p. 202). This is the conceptual activity which is involved in divergent thinking tests. Its definition is similar to the definition of divergent thinking given in this thesis.

It is significant that Rohwer brings up this latter conceptual type since, although he has not measured SES differences in imaginative conceptual thinking it should, as he suggests, be the focus of other research.

With respect to his model, Rohwer predicts that lower SES children will have lower scores on the "recall-application" tasks than on the "acquisition-production" tasks because of their fewer learning opportunities. His results suggest this, but since his model was constructed to do just this, further evaluation with other empirical tests is required, as Rohwer, himself, admits. Lower SES children have also been found to do less well than higher SES children on tasks that require "formal conceptual activity." It should be noted that paired associative learning in this model, following Rohwer's analysis of this task, is placed in the acquisition-production/imaginative conceptual activity quadrant.

This critical attack of Jensen's model appears sound, and Rohwer's conceptualization of intellectual abilities is perhaps superior to that of Jensen. Rohwer's model, however, does not dispute Jensen's findings that lower SES children perform poorly on certain ability tasks, including those involving formal conceptual-type thinking where Rohwer, too, finds SES differences. Rohwer's quarrel is with Jensen's suggestion that certain children, namely those of low SES, do poorly on conceptual type tasks. He removes the necessity for two levels of thinking by postulating two types of conceptual thought, suggesting that paired-associate learning is not associative learning ability but imaginative conceptual ability. This hypothesis needs further investigation with the appropriate empirical tools such as divergent thinking tests. Rohwer's investigation also does not question Jensen's theories on the genetic basis of intelligence; i.e., the question of heritability.

With these restrictions in mind, Rohwer's conclusions are worthy of note. From his model, Rohwer (1971) predicts some interesting testable educational implications. With reference to his model, he states the following:

> It implies that any type of learning proceeds best when conditions are such that conceptual activity is elicited in the learner, whether the conceptual activity called for is formal or imaginative. It does not imply that some subject matters should be taught to some students by engaging them in rote activity and to other students by engaging them in conceptual activity. Instead, it implies that some students should be presented information for learning in such a way as to permit acquisition by means of imaginative conceptual activity, while for other students the subject matter should be presented so that it can be acquired by means of formal conceptual activity.

> The model also implies that, for low-SES students, care should be taken to provide ample opportunities for acquiring information and skills missed because of inadequate early environmental experience. Of equal importance, these opportunities should be tailored to the students' relative propensities for formal or imaginative conceptual activity. Simply, the argument is that a given subject matter can be mastered efficiently either by the route of formal or that of imaginative conceptual activity, depending on the propensities of the students being taught; the corollary argument is that the achievement of mastery by means of rote activity is probably inappropriate for all students, (p. 204).

Although Rohwer has emphasized differences between SES groups, it is obvious that these implications apply to high versus low IQ groups or low versus high achievers <u>within</u> SES groups also.

The model used by Rohwer fits quite well with the model used in the present research. Tasks used involved recall or application of previously acquired skills, such as vocabulary and sub-tests of the WISC; and those that involved producing new information or skills such as the divergent thinking tests and the Raven Progressive Matrices. In terms of conceptual activity the convergent tests involved formal type activities; the divergent involved imaginative-type activities. In addition, the giving of these tasks under different administrative conditions, tests Rohwer's hypothesis as to whether different children master a task more efficiently (as measured by a child's level of performance) by the formal route involving administration under test conditions or by the "imaginative" route simulated by providing a play-like atmosphere in which to do the task. Such tasks, as we have already discussed, should not only involve the convergent or formal conceptual abilities of intelligence but also those that involve divergent processes or creativity. If Rohwer's hypotheses are correct, the following observations would be expected. High level ability children on formal conceptual tasks should perform best under test conditions. Children who have better abilities on imaginative conceptual tasks should do better in an informal atmosphere. Extrapolating further, one can then predict that higher levels of performance on convergent tasks should be under test conditions, whereas higher levels of performance on divergent tasks should be under play conditions.

Divergent Thinking / Creativity

Statement of Problem

Traditional intelligence tests, as already stated, do not measure creativity or divergent thinking, yet the ability to thinking divergently is now generally regarded as an important area in the map of the intellect.

There is much we do not know about creativity. We know less about the relationships of heredity and the environment to divergent thinking than we do to traditional intelligence abilities. Divergent thinking test batteries are not typically used in scholastic assessments, yet it would seem that the service of the intellect in some creative process is of the greatest importance, and one with which educational systems should be rightly concerned. It is possible that one might take convergent tests under formal competitive conditions and perform quite representatively of one's abilities. Yet creative products or expressions may best take place under different conditions, arousal, competitiveness and mental set. Thus, it is possible that the circumstances for developing and assessing creative abilities should perhaps be done under different circumstances from those under which intelligence or convergent thought measures are usually given. Although divergent thinking and creativity are often used synonymously as found in the review of the literature to be reviewed here, they may be better differentiated as the following discusssion hopes to show. But first, the more fundamental problem of how divergent thinking or creativity should be conceptualized needs to be considered.

Definitions of Divergent Thinking / Creativity

Mention has already been made, and it should be emphasized here that creativity or divergent thinking is conceptualized in this thesis as part of the dimension of general intelligence, or put another way, creativity of divergent thinking is a related part of general intellectual abilities. The relationship of conventional intelligence measures to creativity will be discussed, but first a description of how others have traditionally described creativity must be given. Because these are so vast in number, and many of them not pertinent to this thesis, the descriptions given are selective rather than comprehensive.

There is no universally agreed upon definition of creativity and indeed, the term has been used in a myriad of ways. Creativity has become the generic term to cover such qualities as "spontaneous," "original," "original but not bizarre," "imaginative," "ingeneous," "inquisitive," "productive," "problem-solving," "problem finding," and "divergent." Rowell (personal communication, 1975) states that it is often said in the literature that one characteristic of a creative person is his tolerance for ambiguity and that such a quality is also one of the primary pre-requisites for looking at the literature on creativity.

Sprecher (1963) suggests that a universal meaning for the term creativity is not necessary as long as investigators are clear as to what they mean by the term and to what extent, and how they disagree with others. "If we can get people to be precise . . . we can learn to live with any discovered disagreement," (p. 77). Another point is that many formulations on creativity are too narrow and specific to be generally useful.

Alpaugh, Renner and Birren (1976) defined creativity

> as an intellectual process resulting in high-quality productivity such as scientific discovery, notable insight, or artistic composition. The ability to discover and define, as well as solve, problems is commonly used by most investigators as part of the concept of creativity. The word 'creative' is usually used in daily life to commend people for their performance. A creative person is one who has a high likelihood of showing _many_, _different_, _original_ (creative) productions. In other words, creativity is regarded as a complex trait that may be decomposed into component processes, (pp. 18-19).

As a general rule, however, omnibus definitions should be avoided because, although they may have the advantage of inclusiveness, like other omnibus conceptions they have the disadvantage of being an inventory without a unifying rationale. It may be better to restrict the creativity construct to specific operations of measuring it. Many theoretical and experimental investigations emphasize one of several interrelated positions, the product, the process, the motivational and personality characteristics of the creative individual, or the features of the environment most conducive to developing creative talent. Studies placing emphasis on the latter include those which investigate whether changes in the conditions of measurement or administration affect creative performance and will be dealt with in Chapter III of this thesis, since such an issue is also the subject matter of this investigation.

The Product-centered Approach to Creativity

In the product-centered approach, the creativity construct is restricted to recognized performance or achievement or to the qualities that are associated with a recognized creative product. The studies of Lehman (1953, 1960) and Dennis (1956, 1958, 1966), for example, are product-centered in that they look at the productivity of major creative persons over their life span.

Jackson and Messick (1967), on the other hand, looked at the creative product, suggesting four main criteria which should be used in evaluating creative performance, namely, unusualness, appropriateness, transformation and condensation.

Unusualness implies both novelty and statistical rareness of response. As Jackson and Messick (1967) point out, a response is unusual only with respect to a specific population so that in applying this criterion we do so with reference to an appropriate "baseline" group, which serve as the "judgmental standard." "In short, the infrequency of a response is relative to norms, which thus serve as a judgmental standard for evaluating unusualness," (p. 4).

The second criterion appropriateness means that a creative product "must fit its context" (Jackson & Messick, 1967, p. 4). There are both external and internal elements to the "context," respectively, the demands of the tasks or the situation and the motives of the person involved. Primarily, the purpose of this criterion is to eliminate those responses which may be unusual but are also absurd.

The third criterion of transformation involves "transcending traditional boundaries and limitations" (Jackson & Messick, 1967, p. 16). It involves creating new forms out of certain material or ideas in such a way that a new and different perspective is reached. The judgmental standard for this criteria is that of convention or "the constraints of reality," (Jackson & Messick, 1967, p. 6).

The final criterion that these authors propose is that of condensation which involves the endurance of a creative product. It looks at the product in terms of how long its meaning will have lasting value in society. Such a judgmental standard Jackson and Messick called "summary power."

Other writers, who have taken this product-centered approach, include Ghiselin (1963), Mackworth (1965) and Sprecher (1959, 1963). Ghiselin (1963) distinguished between primary and secondary creativity, the first involving a unique product which has not been precedented, the second involving a new application of old ideas such as is involved in many inventions and other problem-solving activities. Mackworth (1965) distinguishes between problem finding and problem-solving. Problem finding is similar to Ghiselin's primary creativity involving "the detection of the need for a new program based on a choice between existing and expected future programs," whereas problem-solving is similar to Ghiselin's secondary creativity and involves "the choice between existing programs or sets of mental rules," (Mackworth, 1965, p. 57). Both Ghiselin and Mackworth see a distinct dichotomy between these two products of human thinking.

Sprecher (1959, 1963), in contrast, emphasizes a continuity or degree of difference that exists between high quality and low quality creative products, the former being evident in a scientific theory, for example, the latter being evident by the fact that some immediate problem is solved.

Rowell (1975, personal communication) makes the following comments with respect to these approaches. Noting that Mackworth, in particular, and also Ghiselin, regard uniqueness to be either present in full force or absent, he concedes that their position may be defensible but asks appropriately whether a distinction should be made between, for example, the creation of some auxiliary hypothesis and the creation of a completely new viewpoint? For Rowell, Sprecher's view is too broad. It is meaningless to have the word creativity if there is nothing which can be classified as non-creative. On the other hand, Rowell feels that Ghiselin's view seems to lack any "sliding scale" of degrees of creativeness. He suggests that perhaps Piaget's terminology could be of use here at both the individual and the group level. The greater the accommodation to assimilation, the greater the degree of creativity.

Rowell (1975, personal communication) further points out that in all these considerations of creativity the term uniqueness (unusualness or originality), the production of something new, is involved. He suggests that the question might reasonably be asked as to whether originality is a synonym for creativity. It has been used in this way by several authors, such as Mackworth (1965) and Hudson (1967). But as Rowell points out there are other uses of the term originality other than using it synonymously with creativity. For example, Guilford (1950, 1967) uses originality in a much narrower sense. For him originality refers to a particular factor of the intellect - one of the 120 hypothesized factors in his model of the intellect. Originality as an ability implies that responses are produced which are statistically rare in a specific population, are remotely related to the stimulus (in contrast to those responses which are conventional) and in addition are, what Guilford calls "clever." Implicit in this is the ability to produce shifts or changes in the meaning of the stimulus so that the end product is novel, unusual or clever. The ability to produce as great a number as possible of such "clever" answers in giving names to stories or punch lines to cartoons illustrates Guilford's meaning.

Maltzmann, Bogartz and Breger (1958) preferred to use the term originality rather than creativity because they believed this could be defined in terms of unusual or uncommon verbal responses without entering the debate over whether these responses were "creative" or merely bizarre. This distinction is also accepted by Levy (1968).

Mednick (1962) makes a clear distinction between originality and creativity be requiring that the response be also useful.

> Creative thinking as defined here is distinguished from original thinking by the imposition of requirements on originality. Thus, 7,363,474 is quite an original answer to the problem 'How much is 12 + 12?' However, it is only when conditions are such that this answer is useful that we can also call it creative, (Mednick, 1962, p. 221).

As Mednick admits, reliable measurement of the usefulness of the product is very difficult in some areas of creative endeavour so that the criterion of usefulness may not always apply. Jackson and Messick (1967), as we have seen, preferred to use the criterion appropriateness.

Thus, although similar terminology is often used by different authors, it is not always used in the same way and indeed, the measurement of a creative product is still controversial. While agreement is more easily reached in the recognition of a product or performance as creative, what makes it creative is still in dispute.

<u>The Process-centered Approach to Creativity</u>

Controversy also surrounds the second approach to creativity, that which is process-centered. Kogan (1973) aptly describes this approach as one where "the principal question concerns the kinds of thinking processes presumed essential to current or future generations of creative products," (p. 147). He also makes the valid point that because a great deal of this research has employed children, rather than adults, as subjects, "the validity issue; i.e., the link between creativity test scores and creative performance in the 'real world' often has to be held in abeyance," (p. 147).

This approach to studying creativity is typical of many investigations where the subjects are given a number of appropriate tests and the "elements" presumed to be part of the creative process, such as fluency, flexibility, elaboration and originality are examined. These studies are usually preceded by a statement of what the authors mean when they use the term creativity. Examples from Torrance and Mackinnon will illustrate this.

Mackinnon (1968) viewed creativity as a process which was characterized by "originality, adaptiveness, and realization," (p. 124). In his study,

> Creativeness fulfills at least three conditions. It involves a response or an idea that is novel or at the very least statistically infrequent. But novelty or originality of thought or action, while a necessary aspect of creativity, is not sufficient. If a response is to lay claim to being a part of the creative process, it must to some extent be adaptive to, or of, reality. It must serve to solve a problem, fit a situation, or accomplish some recognizable goal. And, thirdly,

> true creativeness involves a sustaining of the original insight, an evaluation and elaboration of it, a developing of it to the full, (p. 124).

Torrance (1967) also conceptualized creativity as a process. For Torrance, creativity was

> the process of becoming sensitive to problems, deficiencies, gaps in knowledge, missing elements, disharmonies, and so on; identifying the difficulty; searching for solutions, making guesses, or formulating hypotheses about the deficiencies; testing and retesting these hypotheses and possibly modifying and retesting them; and finally communicating the results, (pp. 73-74).

Torrance (1967) favors this definition for several reasons.

> It enables us to begin defining operationally the kinds of abilities, mental functioning, and personality characteristics that facilitate or inhibit the process. It provides an approach for specifying the kinds of products that result from the process, the kinds of persons who can engage most successfully in the process, and the conditions that facilitate the process. The definition also seems to be in harmony with historical usage and equally applicable in scientific, literary, dramatic, and interpersonal creativity, (p. 74).

Torrance, in this article, points out some of the objections that have been made to his definition. He cites Ausubel (1963), for example, who objected to the definition because it did not distinguish between creativity as a general constellation of intellectual, personality and problem-solving factors and as a highly specific capacity. Ausubel claims that the essence of creativity does not involve general creative abilities, although these exist. Torrance's reply to this is to point out that the abilities one measures during a creative task are "a constellation of general abilities" rather than a particular or specific ability, (p. 74). These general abilities that one measures include fluency, flexibility, originality and what Torrance calls the "ability to sense deficiencies, elaborate, and redefine," (p. 74). While a high level of these abilities does not mean that the endowed individual will necessarily perform creatively, it does mean that there is a greater likelihood of finding creative behavior in such a person.

Follow-up studies of subjects measured by such tests have shown that these measures have a good predictive validity (Torrance, 1972a, 1972b). Scores on creativity tests administered at certain stages of a subject's educa-tion predicted creative behavior in adult life in a significant number of persons. Understandably, their predictive power was more reliable for males than females because of the lack of or variability in career opportunities for women. The subjects in these studies, however, were socioeconomically and educationally advantaged and as

Torrance (1972a, 1972b) admits, these same results would be unlikely to appear in a lower socioeconomic population where opportunities are limited. However, these longitudinal studies are significant, particularly as they are among the first to be conducted with respect to creative behaviour.

Torrance emphasizes the general ability approach in identifying the processes of creativity. This is in sharp contrast to the work of Guilford, who is one of the major figures to use the process-centered approach. Guilford (1950, 1967) and his colleagues worked on the development of what he called divergent thinking tests discovering a number of discrete divergent abilities rather than one general ability of divergency.

A critique has already been presented of Guilford's structure of intellect model. These points will not be reiterated here. In applying his structure of intellect model to creativity, Guilford defined several special abilities which were purported to be essential in creative performance. Of particular importance in this connection are the divergent production abilities which are responsible for the generation and development of ideas.

Divergent production is one of the five categories of "operation" included in the structure of intellect model, as previously discussed. The other "operations" are cognition, memory, convergent production and evaluation. Divergent production means the generation of a variety of responses or diverse ideas about a given stimulus. The response or the product is not bound by the given information. This compares with convergent production where there is only one right answer which is fully determined by the presented information. In the case of divergent productions, Guilford suggests we are looking for logical possibilities or alternatives. With respect to convergent productions he suggests the examinee is generating logical necessities or imperatives. Creative thinking, claims Guilford (1967, 1970, 1971) involves the combined use of both kinds of activities, a feature of his work which is often overlooked, but is one of his most significant contributions. It seems logical that producing a creative product involves both convergent and divergent processes.

Guilford (1967) sees a great deal of similarity between problem-solving and creativity, both involving the sequential use of convergent and divergent thinking. This is a very important suggestion, for as mentioned in the previous paragraph, it seems reasonable to believe that in order to produce a creative product, both open-ended and closure of thought must be involved. Such a conceptualization was, in fact, partly responsible for this author's preference of the terms <u>convergent</u> and <u>divergent</u> when describing the two modes of thinking tested in this thesis.

Guilford also stressed, following this, that there were a number of different types of creative abilities existing within the factorial framework. With respect to the divergent production abilities, Guilford (1950, 1967, 1970) identified theoretically 24 divergent production abilities or functions, 23 of which he claimed to have demonstrated with factor analysis (Guilford, 1970).

Keeping in mind that Guilford found this large number of divergent production factors (i.e., 24), the abilities that are often described, concern the factors word fluency, ideational fluency, associational fluency, expressional fluency, adaptive flexibility, spontaneous flexibility and originality. The following descriptions of these are according to those given by Guilford (1950, 1967).

Word fluency is the ability involved in the divergent production of symbolic units, while ideational fluency is that involved in the divergent production of semantic units. Guilford describes units as being composed of information such as verbal meanings, syllables or perceived objects. Associational fluency involves the divergent production of semantic relations, the production of a variety of responses which are related to the specified object. Expressional fluency is that ability pertinent to the divergent production of systems such as the rapid formation of sentences or phrases. Guilford describes systems as being organized constructs which have more than some minimal complexity. Thus, a sentence is a type of system. Adaptive flexibility involves the divergent production of figural transformations where a shift in meaning is required in a problem-solving task. Spontaneous flexibility on the other hand refers to the divergent production of classes of ideas where a shift in meaning is required from one category to another.

Guilford's notion of originality has already been discussed. Here it is important to restate his emphasis on the transformation abilities which involve a change, revision or a redefinition of the stimulus. These abilities are regarded to be essential to creativity and yet most of them lie outside the divergent production category (Guilford, 1971).

Although in this thesis, Guilford's model of the intellect has been criticized, it must be admitted that his work in addition to being voluminous, has brought into prominence certain intellectual characteristics which had previously received relatively minor recognition. His emphasis on the distinction between convergent and divergent thinking, and their contributions to creative behaviour, has been particularly valuable. Unfortunately, the relationship between "factors of the intellect" and what has been described as "true creativity" is currently very vague.

Furthermore, Guilford's approach may have brought about an excessive fragmentation of abilities or component processes neglecting the possible relationship between factors. Guilford, of course, denies this, claiming, for instance, that his 24 divergent production abilities are stable and "more or less independent," (Guilford, 1970, p. 157). It is the use of the expression "more or less" which suggests to the present researcher that these factors may be more intrinsically related than Guilford states. Future factor analytic work could perhaps ask the question to what extent are certain abilities related rather than to what extent are they independent? One such ability concerns the group of abilities contained within the category of transformations which Guilford (1970) stated had been neglected in relation to creativity. Guilford has identified no less than 20 abilities in this category, yet 16 of these are outside the divergent production category, meaning only four can be regarded as divergent production abilities. The question can reasonably be asked as to whether these four transformation factors in the divergent production category are completely independent stable entities from the other 16 transformation factors? These and other such questions need to be asked if the structure of intellect model is to contribute further to our knowledge of creative processes, keeping in mind that such answers must go beyond explaining statistical significance by providing psychological and functional meaningfulness too.

Vernon (1967) contends that from the experimental evidence, creativity cannot be viewed as a collection of stable abilities. He cites evidence to show that correlations among divergent tests are low and claims that until one can demonstrate consistently high correlations among several such tests, one cannot view divergent thinking as a separate entity, distinct from the well-established general factors such as g and V. Citing some of his previous research Vernon (1967) claims that "the question whether divergent thinking abilities are part of, or distinct from, what is measured by intelligence tests depends more on the form of the tests, the way they are given and scored, than on their apparent creative or non-creative content," (p. 157). In addition, Vernon stresses that the divergent tests used have very poor predictive power and that divergent tests do not necessarily measure creativity. The point is well taken and in this research the investigator does not claim to be measuring "true creativity" with the battery of divergent tests used.

Investigations mentioned so far have used open-ended tests when looking at the divergent process. One investigator, who has used a different approach under the name of "creative thinking" is Mednick (1962). He developed an associative rationale of the creative process and emphasized the essential similarities of all creative acts rather than their differences. From his perusal of the introspections of recognized creative artists and scientists he generated a definition of creative thinking as "the forming of associative elements into new combinations which either meet specified requirements or are in some way useful. The more mutually remote the elements of the new combination, the more creative the process or solution," (Mednick, 1962, p. 221).

In this regard it is worth pointing out the criticism of Jackson and Messick (1967) who, although applauding Mednick for going beyond the criteria of unusualness to include usefulness, contend that his criteria still fall short. A description of what these authors propose as other criteria to remedy this has already been given.

Mednick (1962) proposes three ways in which the required associations can be brought together. The first way is serendipitous, whereby the required associative elements are elicited by fortuitous environmental stimuli. The second method involves a similarity between the elements presented or between the associative elements required. And finally, the process of mediation of other commonly related elements may trigger the production of the requisited associative elements.

As an operational statement of his theory, Mednick devised a test of creativity which he called the Remote Associations Test (RAT). The test requires that the subject "form associative elements into new combinations by providing mediating connective links," (Mednick, 1962, p. 226). Operationally this means that the stimulus elements involving several words "from mutually distant associative clusters must be presented to the subject, his task must be to provide mediating links between them . . . the mediating link must be strictly associative rather than being a sort that follows elaborate rules of logic, concept formation, or problem-solving," (Mednick, 1962, p. 227). The example often quoted is that where the three stimulus words are: RAT BLUE COTTAGE. The answer which provides the requisite associative link is cheese.

The essential point is that to each question there is a right answer, and the subject's score is the total number right. Unlike most other "creativity" or divergent thinking tests, including those of Guilford, the subject, although he is required to carry out a mental search, is set much stricter limitations on acceptability of response. Since this is not the case with open-ended tests one must conclude that they are not measuring the same thing as the

RAT. As discussed, open-ended tests are aptly described as tests of divergent thinking. The RAT may require divergence, in the sense of utilizing divergent thinking processes in the early stages of attempted solution, but these are tempered by criteria, and in producing a simple solution, the later stages of thinking, at least, seem more properly described as convergent, (Rowell, 1975, personal communication).

Yet the theoretical basis of Mednick's definition of creativity has been used in important investigations such as those of Wallach and Kogan (1965). The RAT was used by Taft (1967) as his creativity test in his investigation of the relationship of creativity to academic achievement. His results were used by Ginsburg and Whittemore (1968), who also used the RAT as their creativity measure to collect their own independent data, to support their curvilinear theory of the relationship between IQ and creativity and to dispute the threshold hypothesis. In view of the convergent like characteristics of the RAT their conclusions are possibly not valid.

Before leaving the process-centered approach to creativity it should be recognized that there are those who have attempted to identify the processes involved by studying the introspective reports of men of recognized creative talent which, at the least, provide a degree of agreement on successive "stages" involved in the creative act. Wallas (1926), for example, suggested that these stages were:

(a) preparation;

(b) incubation;

(c) illumination; and,

(d) verification.

Despite variations and modifications, these four stages of creative thought have been accepted since this time. As Rowell (1975, personal communication) points out, we run into difficulty when we examine what we mean by "incubation" and find that it is only a name which, although it conjures up a tempting analogy, is really just a cover for a considerable and abysmal ignorance.

But the introspective approach should not be ignored and other contributions provided by this approach are neatly analogued in a book edited by Ghiselin (1952). Kneller (1966) points out that contributions to theories of creativity came from ancient and modern philosophy including logic, a range of psychological theories such as associationism, gestalt theory, psychoanalysis, neopsychoanalysis, neo-Freudianism, S-R learning theory, experimental psychology, factor analysis and so on.

An integrating attempt from this wide range of disciplines was attempted by Koestler (1964), who adopted the thesis that all creative processes share what he calls "bi-sociation" which means by thinking on more than one plane of experience (as compared to working by past association, typical of conventional thinking) man engages in a creative act which involves relating together previously unconnected pieces of thought or unrelated levels of experiences to higher mental order thought processes. Although such a definition is reminiscent of Piagetian theory, Koestler's statements still remain untestable hypotheses. Indeed, Rowell (1975, personal communication), commented that many of these contributions could be merely described as blind-fold subterranean excursions, or as Medawar (1967) so acidly described it, with reference to Koestler's work "slopping around in the amniotic fluid," (p. 87).

The Personality-centered Approach to Creativity

The personality-centered approach is concerned with the motivational and cognitive bases of the creative personality. It studies people judged to be creative in specific fields in an attempt to determine or measure, through the use of a variety of psychological assessments, what personality, motivational and cognitive bases are associated with their creative achievement. This was not the concern of this thesis so a discussion of these studies is not given here, except to say that little is known about this approach. The main contributions come from the efforts of American psychologists, particularly Barron (1955, 1963, 1969), Cattell and Butcher (1968), Cattell and Drevdahl (1955), Drevdahl (1956), Drevdahl and Cattell (1958), Mackinnon (1962a, 1962b), and Roe (1952).

There has been little work of a systematic nature done in this field in Britain, although one study by Cross, Cattell and Butcher (1967) examined some differences in personality between creative writers and a sample of the general population. Hudson (1967, 1970) did find that convergers and divergers used different types of ego-defense mechanisms, so that the former tends to deny emotions while the latter overstates or exploits them. These approaches to dealing with their emotional life are reflected in their styles of thinking. Little research has been done in this area examining the relationship of the extraversion-introversion dimension to different styles of thinking such as convergency and divergency. One such study by Hudson (1970) found that contrary to what might be expected there was no relationship between convergence and introversion or divergence and extraversion amongst secondary school boys. However, he did find that divergent boys had a higher neuroticism score on the Maudsley Personality Inventory than the convergers. Rowell and Renner (1975, unpublished study) in an investigation of postgraduate students of both sexes, found no evidence to suggest that extraversion, introversion or neuroticism, as measured by the Eysenck Personality Inventory, were related to scores on convergent and divergent tests. Despite these negative results, the personality approach to the study of creativity is an attractive one and warrants a great deal more research.

Cognitive Bias

Each of the three previously described approaches run into problems in measuring creativity with open-ended tests and validating such measures with "true creativity." Hudson (1967, 1970) provided an alternative. He suggested that another way of approaching the problem of creativity and intelligence may be more valid. Creativity, he claimed could be attained by excellence in either convergent subjects like science, or divergent subjects like arts. Much of the previous research, he contended, had thus confused the issue of intelligence and creativity by failing to recognize a more fundamental issue, that of differences in cognitive bias, state or mode. Thus, it was far more meaningful to discuss differences in performance on conventional convergent and divergent tests in terms of cognitive bias. The theoretical and empirical importance of this issue of cognitive bias is particularly demonstrated by examining the differences in scores between art and science students.

Hudson maintains that creativity and originality should not be identified exclusively with the diverger. He suggests that the convergence and divergence of an individual will not determine whether he is original. Rather, if he is original, the dichotomy of convergence and divergence will predict the type and style of the creative products.

Hudson further suggests that the roots of the individual's originality are not in his convergence or divergence but rather in other aspects of his personality. In addition, Hudson emphasizes that the convergence/divergence dimension is a measure of cognitive bias, not of level of ability. He does not deny that there is a relationship between IQ and creativity or IQ and achievement, but believes that after a certain level of IQ the relationship becomes less. Stating such a case Hudson (1967) claims:

> The relation of IQ to intellectual distinction seems, in fact, highly complex. As far as one can tell, the relation at low levels of IQ holds quite well. Higher up, however, it dwindles; and above a certain point, a high IQ is of little advantage. However, there are differences between one occupation and another, the relationship dwindling lower down the IQ scale in some subjects than in others. In the arts, for instance, it seems to peter out lower down than in science. For practical purposes, therefore, it might be fruitful to distinguish, for each occupation or

> subject, both the IQ levels above which a strength in IQ is not an advantage in real efficiency; and also lower limits, below which a weakness in IQ becomes incapacitating. Where one sets such limits depends, of course, on the criteria one has in mind. For argument's sake, we might define success academically: a good second-class degree at Oxford or Cambridge; or, in more worldly terms, a successful novel or a good piece of scientific research, (pp. 124-125).

A certain basic level of intelligence is essential for creative behaviour. Beyond that minimum, however, other variables such as personality and motivation become essential. A further discussion of the relationship between intelligence and creativity will be made in Chapter III.

Hudson (1967) tested groups of Public and Grammar school boys who were of above average ability. From these results he divided his sample into contrasting groups of divergers and convergers with the middle 30 percent being described as "all-rounders." The terms diverger or converger were used because he claimed that this distinction was more meaningful than that between creative and intelligent. Hudson then investigated the differing characteristics of the divergers and convergers and found that the differences involved not only cognitive style but also personality characteristics, particularly the type of ego-defence mechanisms adopted.

But perhaps the central finding of Hudson's work was the relationship he discovered between convergence/divergence bias and intellectual specialization. He found that specialists in Arts subjects (History, English and Modern Languages) tended to be relatively weak on IQ tests and relatively better on open-ended divergent tests; i.e., there was a marked relationship between divergent thinking and specialization in Art subjects. In contrast, Science specialists (Maths, Physics, Chemistry and Classics) included more "convergers" who did relatively better on IQ tests but less well on open-ended tests; i.e., there was a strong relationship between convergent thinking and science specialization.

In terms of academic achievement, Hudson (1970) reports the following differences between divergers and convergers.

> Among English fifteen-year-olds . . . the diverger is less academically successful than the converger in a ratio of more than two to one. At the level of university entrance, the balance has begun to redress itself. Of the fifteen-year-old convergers, divergers and all-rounders in my sample, nearly two hundred have now left school, and three-quarters of them have gone to university. The convergers and all-rounders were distinctly more likely than the divergers to reach university. Of those admitted to Oxford and Cambridge, twenty-seven won scholarships or exhibitions; again, these were more likely to be convergers than divergers, (p. 95).

Thus, the trend was for divergers to do increasingly better academically as they progressed through school although the convergers still had better performances at university entrance. Hudson suggested that this trend of increasingly better performance academically on the part of the diverger might continue throughout the university. There has been other research evidence to support this.

Cropley (1967) made a longitudinal study of male science students at the University of New England in New South Wales, Australia. On entering the university, a selection of these students completed a battery of convergent and divergent tests. There was a fairly even distribution of convergers and divergers amongst the able students selected and there was no difference between these two groups with respect to university entrance marks. A comparison of examination grades at the end of the first year showed no difference between convergers and divergers. However, there was a difference at the end of the third year with respect to excellence in academic performance. Divergers did significantly better than the convergers and formed the majority of those students selected into the fourth year honours programme. Those who received first class honours at the end of this fourth year were almost all divergers also. Cropley (1967) makes the comment that although his sample is small and, therefore, implications limited, his study does "suggest that later years of undergraduate study, and especially the honours year, may give students more opportunities to utilize divergent thinking skills. . . ." (p. 672).

In commenting about this study, Cropley (1969) further suggested that,

> the real difference between the men who later bloomed in science, and those from whom they did not differ in achievement at lower levels, lay not in the level of their intellects, but in the style in which they utilized whatever intellectual endowments they had. It then becomes relatively easy to understand why two people of apparently equal levels of ability may differ so markedly in the areas in which they achieve, and the levels of the educational process at which that achievement occurs, (p. 6).

A larger longitudinal study was carried out by Field and Poole (1970). There were 101 male Arts and Science students involved who, on entering the University of New England, New South Wales, Australia, were given a divergent test and two convergent tests. The investigators examined the relationship between cognitive bias and academic achievement by looking at the quality of their levels of performance on examinations. At the end of the first year the convergers from both the Arts and Sciences achieved better results, but at the end of the second year there was no difference in achievement between those who were convergently biased and those who were divergently biased. Field and Poole (1970) suggest that at the second year level the skills of the convergent student may be less appropriate, a trend which they predict will continue during the third and Honours year, while the diverger would be expected to excel. At the time of their report they had no results to confirm this prediction.

Hudson's findings on sixth form boys were confirmed by Cropley and Field (1968) who found in a sample of Australian male high school students that those specializing in science subjects were predominantly convergently biased. Only a partial confirmation of Hudson's results, however, came from a study by Mackay and Cameron (1968). They took a large sample of Scottish male undergraduates who were similar in age and ability to those boys in Hudson's sample but differed in the degree of specialization they had been subjected to during their secondary school years. For those students who had elected to specialize in their first year they found that those in Arts were predominantly divergently biased whereas those in Science were predominantly convergent, which confirms the association between cognitive bias and subject choice. However, this relationship was not found amongst those students who did not elect either Arts or Science as a speciality. The authors suggest that the relationship between cognitive style and subject specialization may be due to a greater degree of subject specialization in the English educational system and that "the educational experience of following an Arts or Science curriculum moulds or at least reinforces a particular style of thinking," (MacKay & Cameron, 1968, p. 316).

Renner and Rowell (1975, unpublished study) found a significant relationship between cognitive bias and subject specialization amongst 136 full-time postgraduate Diploma in Education students of both sexes at Adelaide University, South Australia. Those students with a B.A. were generally better on the divergent thinking tests. Conversely, those students with a B.Sc. performed better, generally, on convergent tests. In addition, when cognitive bias was examined it was found that the Science students tended to be convergently biased, and the Arts students tended to be divergently biased. In terms of academic achievement, Art students did better than the Science students in three out of four subjects.

Povey (1970) looked into the question as to whether academic bias affected test performances before formal specialist training in the Arts and Sciences began. They followed a group of English school boys from the fourth form level where there was no specialization of subjects, through to the sixth form where a choice of specialization into either Arts or Sciences was made. The performance of this sample was compared with two other groups of boys who were receiving an arts or science training. The main result showed that there was a definite cognitive bias or style before the boys were introduced into specialist curriculum courses. Environmental influences, including the home, genetics, particularly personality predisposition and educational experiences are involved as possible reasons. Cognitive bias was found to increase with age and to be reinforced by the length and type of specialist education. Povey's results also point out the importance of studying children as early as possible in their education before the cognitive bias reinforced by secondary school educational factors set in.

Another related issue concerns the topic of mental set. It is possible that although the basic abilities may be the same in convergent and divergent thinking, such as some underlying general factors, the "set" of the mind may determine the way the abilities are deployed or perhaps mixed. This would fit quite well with the model of Rohwer (1971) already discussed. Allied to this issue is the question concerning appropriate environmental or educational contexts, for example, increasing performance on divergent tests. Hudson (1970), for instance, found that under specific instructions convergers could diverge quite well. But in the absence of instructions the question arises as to whether convergent and divergent performance levels can increase in a specific environmental context.

The influence of specific environmental contexts on convergent and divergent thinking performance was the subject of the present study. The literature specific to it, and this fourth approach to divergency/creativity will be dealt with in the next chapter.

CHAPTER III

RESEARCH RELATED TO THE PRESENT STUDY

Statement of Problem

It is desirable at this point to repeat the purpose of this study, which was to investigate how varying the conditions of administration would affect the levels of performance on "tests" of convergent and divergent thinking. A secondary purpose was aimed at examining whether the relationship between convergent and divergent thinking changed under different assessment procedures. These purposes have both practical and scientific significance. The practical significance lies in the determination of what conditions best permit the expression of the abilities of children. The scientific importance lies in the responsiveness of convergent and divergent tests to different testing conditions; i.e., should convergent and divergent "tests" respond differently to the conditions of this research it provides further insight into the basic nature of convergent and divergent thinking.

Thus, the present study is concerned with the effects of the context of taking tests of mental abilities, specifically the effects on performance on tests of convergent and divergent thinking under test and play conditions or the effect of test order. The term context effects refers to changes in performance of these tests resulting from the conditions under which the tests are taken. The effects on mental tests can be of several kinds. Level of performance might be effected so that we might question under what conditions, either test or play, do children best perform mental tests or whether doing either convergent or divergent "tests" first affects the level of performance on the other type of test. A related question might also be raised as to what kinds of conditions result in the most valid measurements of mental abilities.

The current research did not begin in a vacuum. Previous studies bear on the present research question. Discussion in the last chapter was given to show the systematic background of previous enquiries into intelligence/convergent thinking and creativity/divergent thinking, and the origins of the author's personal views about the nature of convergent and divergent thinking. This provides the framework in which we can more meaningfully examine the possible effects of conditions on these aspects of intellectual function. The following discourse describes previous research in the area of the effects of different environmental contexts on convergent and divergent thinking, in this way providing the rationale for the author's hypotheses in this thesis. The question of environmental influences has, for the most part, only been investigated with respect to divergent thinking / creativity.

Levels of Performance on Convergent and Divergent Tests Under Different Administrative Conditions

Wallach and Kogan (1965) suggested that divergent thinking was facilitated under play conditions. However, in their study the entire sample of elementary school children had their creativity tasks under non-evaluative conditions. Since there was no control group, no comparisons can be made. The main purpose of their study was concerned with the dimensionality of creativity and intelligence, a topic which will be later discussed.

Part of the research design of Boersma and O'Bryan (1968) was to evaluate the performance of 46 fourth grade boys on divergent thinking tests in evaluative and non-evaluative atmospheres. All subjects were given an intelligence test under standard evaluative conditions. The subjects were then divided randomly into two groups for the divergent tests. Group A did their tests under similar evaluative conditions. Group B, however, were taken out of the traditional school setting to eliminate any kind of association with the classroom, time limits and the atmosphere of evaluation. At a nearby university gymnasium, Group B were allowed to engage in uninhibited play activities for 60 minutes, after which the investigator introduced the divergent tests as different kinds of games. Time conditions were imposed, but unobtrusively. Analysis of the results showed that there was no significant difference between the two groups on the verbal and non-verbal intelligence measures, both of which had been given under test conditions; but that the two groups did differ significantly on the verbal and non-verbal creativity measures. Group B did very much better. While this appears to lend support to the Wallach and Kogan hypothesis for enhancing creative behaviour, a few qualifying statements need to be made.

Wallach and Kogan (1965) tested their children individually. Boersma and O'Bryan (1968) tested their boys in a group setting. For educational purposes the study of Boersma and O'Bryan would appear more applicable, since evoking and assessing divergent thinking behaviour at school necessarily has to take place in a group context. However, in the study of Boersma and O'Bryan, those subjects who had their divergent "tests" in a play atmosphere were taken out of the school setting to a recreation center. Boersma and O'Bryan point out that such a change in normal school conditions could have produced the difference in results, and that this Hawthorne effect could be leveled as criticism against the design of Wallach and Kogan also. Certainly the removal of students from the school setting, and the applicability of individual administration, limits the applications that can be made from these studies.

There is some evidence that students from elementary schools which have an informal atmosphere, lack an authoritarian approach, and use informal teaching approaches, do better on divergent thinking "tests" than students from formal, authoritarian subject-centered schools. Haddon and Lytton (1968) tested two pairs of contrasting primary schools, each pair being matched for socioeconomic background. The evidence from this study gave a very clear indication of superiority of divergent abilities among pupils from the informal schools. However, this distinction between formal and informal schools was not evidenced when a similar procedure was used in secondary schools. Criteria between the two schools rested upon a permissive versus authoritarian approach in teaching methods but no differential results were observed, (Lytton & Cotton, 1969).

Lytton and Cotton (1969) explained their results with the following comments:

> The negative result of the comparison between the 'types' of schools may be due to the lack of contrast

> between the schools, but it may also be that the effect of a more flexible approach to learning are not reflected in the performance of 14-year-olds in these divergent tests, perhaps because of the limitations of the tests. . . . The result should therefore, not be interpreted as a negative verdict on 'informal' schools, but it illustrates the difficulty of this kind of investigation in the more complex secondary school organization, (p. 190).

Vernon (1971) took two sets of seven classes at the grade eight level. Such classes were made up of Canadian adolescents with a median age of 13 years and 11 months. To one set, seven group tests of divergent thinking were given under evaluative formal conditions, with each test being timed. To the other half of this sample, the divergent tasks were administered in a more game-like relaxed atmosphere, without the pressure of time limits, although the total sessions were time limited.

"Testers" in the game-like atmosphere emphasized the dissimilarity of the divergent tasks from intelligence and attainment tests. Under the informal conditions, students not only read their instructions, but organized their own time for completing each "test." Vernon (1971) pointed out that "the relatively permissive atmosphere had its disadvantages; there was a good deal of conversation and some copying from neighbouring students. And some students concentrated too long on a few tests that they liked the look of, leaving very little time for others," (p. 250).

Thus, although there was less extreme in administrative procedures, as in the Boersma and O'Bryan or the Wallach and Kogan studies, there was also less careful control.

Since the results were positively skewed, Vernon (1971) presented his data as percentile scores. Overall results for originality of response on the seven divergent tests showed that higher scores were obtained with the relaxed groups; but the results did differ from test to test which Vernon feels is due to the fact that students under relaxed conditions gave more time to the "tests" they liked. The greatest difference between the two groups occurred at the upper end of the distribution of scores. There were more people from the relaxed group who reached scores at the 90th percentile. Fluency scores did not show this difference between the relaxed and formal groups, indicating that the superiority of performance of those under informal conditions concerned the quality of the response; i.e., originality rather than total quantity.

Vernon (1971) cites a study done by Nicholls (1971a) whose results also substantiated the Wallach and Kogan hypothesis. One group of children were given four of the Wallach and Kogan tests individually under untimed game-like conditions. In a subsequent session they took two convergent tests and a parallel battery of divergent tests under timed formal conditions. A second group, the control took the test battery only. Both the fluency and originality scores were considerably higher under the game-like conditions. The children who did the divergent tests after having parallel tasks under a play context also had higher scores under the formal conditions than the control group did but no significant levels are reported. This trend in the results, however, does suggest that the game-like conditions administered first may have affected scores on the subsequent formally administered tests, a feature which Nicholls (1971b), in an earlier paper, admits. Indeed, administering even parallel tests under two sets of conditions runs the risk of one condition contaminating the other and in the investigation of the present thesis, this was avoided.

In both the study of Vernon and Nicholls there was a difference between the two groups with respect to time limits. This could also have affected the results. Wallach and Kogan (1965) and Ward (1969) showed that although response rate may decrease with time, unusualness of response or originality increases, a conclusion validated by Cropley (1972). This would appear to be a significant point in view of the fact that the greatest difference between these groups tested under informal and formal conditions was in the area of originality.

Hargreaves (1974) took two groups of schools of 10-11 year-old children of both sexes. One group had three tests of divergent thinking under test conditions. The other group had the same battery of "tests" under game-like conditions. All tests were administered in the one session in group form and a reference intelligence measure was given at the end of the session under standardized test conditions. There were no time limits for the divergent tests under either the formal or the informal conditions.

Higher scores under relaxed conditions were found for fluency on both the non-verbal tests, but only on one non-verbal test was there a significant difference for the originality score which Hargreaves feels reflects the differential popularity effect for certain divergent tests. There was no significant difference between the scores of the two groups for the verbal divergent test. There was only one sub-score where the group under test conditions achieved a higher score (which was non-significant), so that Hargreaves (1974) concludes that "the general superiority of game-like conditions has been demonstrated over <u>untimed</u> test-like conditions," (p. 87).

Surprisingly, Hargreaves (1974) claims that the depression of divergent test scores among those doing it under formal conditions was due to the removal of time limits. To have no time limits under test conditions, he claims introduces ambiguity so that children under these conditions attempt to behave as if they were under timed conditions and attempt to complete the tests as fast as possible. Such a claim would have to be validated.

However, the Wallach and Kogan hypothesis claiming better performance on divergent thinking tests has not been confirmed in all studies.

Williams and Fleming (1969) gave three Wallach and Kogan tasks and the Peabody Picture Vocabulary test (convergent) under play and test-like conditions to 36 four-year-olds. This study is of note since it is one of the rare studies to consider whether performance on a convergent test may alter with change in conditions.

Very little difference in the convergent score is found between the play and evaluative conditions, the mean scores being 113.9 and 114.9, respectively, and Williams and Fleming (1969) fail to make any comment. In addition, the Peabody Picture Vocabulary test, unlike paper and pencil tests is a game-like task by its nature. Mean scores on the divergent tasks were all higher under the evaluative conditions, but the difference was significant only on the non-verbal tasks.

However, the study has limited application because of its various design weaknesses. The most obvious one was that each child experienced both conditions, and as with the study of Nicholls (1971a, as cited by Vernon, 1971; 1971b) there was the possibility that one condition might influence the other, despite the fact that the children were randomly assigned to four groups, and order of tests and order of conditions were varied. In addition, as Williams and Fleming (1969) admit, their subjects were pre-school children who had not been exposed to the formal atmosphere of a school system and would be unaware of an evaluative set. They would be, therefore, less influenced by efforts to induce such a test atmosphere. Finally, their sample is small and findings could only be suggestive rather than conclusive.

Much more substantial evidence was found by Kogan and Morgan (1969). They took 104 Grade 5 children of both sexes and divided them into two groups. Two Wallach and Kogan tests, a verbal and non-verbal divergent task, were given under test-like conditions to one group and under play-like conditions to the other. The "tests" were given in group form rather than being administered individually, and the group under test conditions was timed; while the group under game conditions was unobtrusively timed. In this case the children under evaluative conditions produced better performances both in terms of the total number of responses and the number of unusual or original responses.

Leith (1972) in a study which examined whether responses on divergent thinking tests were influenced by the amount of stress imposed by different testing procedures (and the subject's personality) amongst 106, 9, 11 and 13-year-olds confirmed the above findings. On three verbal divergent tasks the total number and originality of responses was greater in those who completed the tasks in a moderately stressful atmosphere, rather than in a relaxed, informal atmosphere. Introverts were found to cope better with the more structured stressful condition than extraverts.

Channon (1974) also found evidence that mild stressful conditions increased divergent thinking scores. Two hundred and four third form children of both sexes, from an English comprehensive school were used. The sample was divided into two groups and each group was tested on two occasions with parallel forms of six measures of verbal and non-verbal divergent thinking tests. The Raven's Progressive Matrices and the Mill Hill Vocabulary Scale were used as the convergent thinking measures. One group had both sessions under relaxed, informal conditions. The second group had the first session under relaxed conditions and the second session under evaluative formal conditions. The complete divergent and convergent battery were given in one session. The relaxed regime was established by the tester spending some time with the students sharing examples of humour and poetry from other pupils in order to establish a friendly atmosphere. Participation under the game conditions was voluntary, but not so under test conditions.

Overall, the results showed that performance, particularly on the divergent tests, is increased under formal conditions. Channon (1974) presents his results for boys and girls separately, without stating whether there is a sex interaction. With three exceptions, both fluency and originality were higher under the test regime for boys, compared with those under relaxed conditions. Boys also attained a higher score for the non-verbal convergent test under test conditions. Both groups of boys, under the two sets of conditions, showed a very slight decrease on the second administration of the vocabulary "test."

Amongst the girls Channon (1974) found this same trend, for an increase in scores under test conditions, with only one instance of a decrease in an originality score on a drawing completion test. Generally, however, higher scores are recorded on the divergent thinking tasks under test conditions. There is little difference in scores between the two sessions in either group of girls on the non-verbal convergent measure, but the increase in vocabulary scores among those girls who had both sessions in a relaxed atmosphere is significant at the 0.5 level.

Unfortunately, Channon does not present other results which would show which differences are significant between the two regimes. However, there is a less obvious difference between the two regimes among the girls, a feature which Channon contributes to the greater motivation on their part to do well, whatever the conditions. He concludes that moderate stress improves scores, particularly divergent thinking ones, and that the evidence that originality as well as fluency increases under test conditions, as in his study, fails to support those arguments which claim the necessity of an informal atmosphere for producing unique ideas.

This study by Channon in addition to examining the effect of different conditions on divergent scores has also tentatively examined the effect of different regimes on convergent thinking. However, to do this, he tested two groups twice, thereby risking contamination of conditions in that group which experienced two different regimes.

The significance of his results lie in the implication that when two samples are re-tested under different conditions those tested in a formal atmosphere have higher divergent thinking scores. The question then arises as to what would happen to the level of divergent and convergent scores if different but parallel samples of children were given these tests under different regimes, with some groups assessed in a formal atmosphere and others assessed in an informal atmosphere?

To include tests of convergent and divergent thinking in the same testing session fails to isolate the problem of the effect of administration on each battery sufficiently. Divergent thinking tests, being open-ended, have a content set which is very different from that implicit in convergent thinking tests. What is the effect of regime on convergent thinking tests if this contamination is removed? No study has given emphasis to the situational influences on convergent thinking. Thus, such a question, in addition to evaluating the effect of different conditions on divergent tests, where the issue is still controversial, comprised the main purposes of this study.

Further contamination may arise if order of presentation of the convergent and divergent tests is not controlled. In none of the studies cited is this issue discussed, and the effect of the order of presentation of these two groups of tests may vary. Control of this problem was incorporated into the design of the present thesis.

The Relationship of Convergent and Divergent Thinking Under Contrasting Administrative Conditions

With the exception of the study by Leith (1972) and Channon (1974), the main emphasis of the series of investigations described in the last section was on the relationship between convergent thinking/intelligence and divergent thinking/creativity. The present investigation did look at the correlation coefficients between average intelligence and creativity scores, but it was a secondary purpose. Thus, it is the intent of this literature review to indicate some of the major trends in this area only, rather than to undertake a full description of the very extensive research that has been done concerning the relationship between convergent and divergent thinking.

The study by Wallach and Kogan (1965) has been mentioned several times. It was the first of a long series of studies to investigate whether a change from the usual formal conditions of administration to a play-like atmosphere would clarify the relationship between intelligence and creativity. In their study of 151 children, creativity was found to be a separate dimension of the intellect from intelligence. This was based on the finding that the correlation coefficients between intelligence and creativity were low, no correlation exceeding .23. In addition, they found high reliability for their intelligence and creativity measures, and found that indices of creativity and indices of intelligence were highly related among themselves. The low relationship between intelligence and creativity was also found when boys and girls were examined separately. However, this result is severely criticized by Cronbach (1968) on the grounds that Wallach and Kogan failed to first demonstrate an interaction effect between the sexes.

Guilford (1971) also criticized Wallach and Kogan's results on the grounds of changing the test conditions. Guilford felt that the removal of time limits diminished the control needed in experimentation, and that by allowing an unlimited amount of time could change the nature of the variables measured. Guilford (1971) also voices doubts as to a playful atmosphere inducing creativity, claiming that "a test is a test, even when it is called a 'game'," (p. 81).

However, despite these criticisms, the Wallach and Kogan study is widely cited. It is a remarkable piece of work, for it goes beyond an analysis of creativity and intelligence to examine other problems such as conceptual style, temperament and personality.

Several attempts have been made to confirm or disprove the findings of Wallach and Kogan (1965) with respect to the separate dimensionality of creativity. One major difficulty with many of these studies is that by not adding a control group working under test conditions so that comparison of a play-like regime can be made with other procedures, conclusions are limited.

Williams and Fleming (1969) in a study previously described with pre-school children did find a low non-significant correlation between the associative fluency score measured under play conditions and the intelligence score measured under test conditions. However, the correlations between the divergent thinking score and the vocabulary were also low under evaluative conditions, and the former scores were significantly correlated under both an evaluative and play set so that Williams and Fleming (1969) concluded that the Wallach and Kogan hypothesis had not been supported, and that a game atmosphere was not necessary for "a valid assessment of associative fluency," (p. 161).

Strong confirmation of the Wallach and Kogan findings was found in a study by Wallach and Wing (1969) on 500 college students. Students completed the battery of "tests" in group form, written but with no time limits. There were high correlations amongst the different divergent measures, and low correlations between the divergent and convergent tests. However, since there were no control groups, no comparisons with other procedures can be made.

A similar criticism can be made against the studies of Cropley (1968) and Cropley and Maslany (1969). Cropley (1968) gave the Wallach and Kogan tests under informal conditions to 124 university men, in group form, in addition to an intelligence test under formal conditions. Principal axis factor analysis yielded a large general factor which accounted for 41.2 percent of the variance common to both variables. A large second factor was also obtained. Loadings on this fell into two groups and accounted for

29.2 percent of the common variance.

Similar results were obtained by Cropley and Maslany (1969) who gave, in group form, the Wallach and Kogan tests to university men and women under informal conditions. Intercorrelations between the intelligence and the creativity tests were very low so that they concluded that intelligence and creativity formed a separate cluster of cognitive abilities. But factor analysis of the combined matrix of intercorrelations between these two abilities yielded two oblique factors. The first factor was a general factor, although 76.3 percent of its variance was derived from the creativity tests. The second factor was bipolar yielding poles of both intelligence and creativity. Cropley and Maslany (1969) conclude that creativity tests do "measure a stable and internally consistent intellective mode albeit one which is substantially related to general intelligence," (p. 398). Kogan (1971) in using a different factor-analytic method on the same data, namely a Promax rotation of Cropley and Maslany's principal components, found a distinct separation of creativity and intelligence.

Investigations by Boersma and O'Bryan (1968) and Nicholls (1971a, 1971b) used a control group which completed a test-like battery only. In both studies correlations of the divergent thinking tests with convergent tests were lower under the play conditions, which these investigators claim support the hypothesis that these two groups of abilities form separate clusters of the intellect.

A follow-up study to test the stability of creative ability under informal conditions was made by Kogan and Pankove (1972). Divergent thinking scores remained relatively stable, although for boys this applied to group testing and for girls applied in the context of individual administration. There was a sex difference, however, in the way intelligence and creativity were related over a five-year period. At the fifth grade level intelligence and creativity were unrelated for both sexes; i.e., both abilities were separate dimensions of the intellect. But five years later, although the relationship between intelligence and creativity remained unrelated in the case of the girls, there was a positive statistically significant correlation in the case of the boys. Kogan and Pankove attribute their results to a difference in personality and motivational factors which they suggest are more important in the case of girls.

Other investigations have produced results which have provided even more direct proof against the Wallach and Kogan hypothesis. Ward, Kogan and Pankove (1972) point out that the role of testing conditions in isolating creativity from intelligence is still questionable. Indeed, the results from studies which have contrasted formal and informal administrative regimes are not only complex but contradictory. Some, as has been noted in this text, support the Wallach and Kogan hypothesis, but others, such as the study of Kogan and Morgan (1969), do not.

Kogan and Morgan (1969) gave some of the Wallach and Kogan tests under two types of conditions to 104 fifth grade boys and girls. Half the sample received their divergent "tests" under formal conditions, the other half had a more game-like regime. Correlations between the convergent and divergent measures were generally insignificant or negative so that there was no evidence to indicate that greater separation of the two abilities occurred under game or test conditions.

Vernon (1971) in his study of Canadian adolescents, where half his group did the divergent battery under relaxed conditions and the other half in a formal setting, found that the correlations between intelligence and creativity were higher in those scores from the relaxed group. A similar type of design with 10-11-year-old children was utilized by Hargreaves (1974) and yielded the same result. Expressed another way, in these two experiments by Vernon (1971) and Hargreaves (1974), the correlation between intelligence and creativity was lower under test conditions. This is an unexpected result because the main body of evidence which looks at the relationship between creativity and intelligence under test conditions does not find any real separation of these two dimensions.

Getzels and Jackson (1962) claimed that their study did show that creativity was a separate dimension from intelligence, but reanalysis of their data by Marsh (1964) showed that when corrections were made for their biased sample there were very significant correlation coefficients between intelligence and creativity. The children tested by Getzels and Jackson were highly atypical, coming from a high socioeconomic strata and having a mean IQ of 132.

In a less biased sample, however, among Irish public grammar school pupils, Dacey, Madaus and Allen (1969) did find that their divergent thinking measures were relatively independent of each other. These investigators also found that verbal and non-verbal divergent thinking tests could be divided into two different factors.

Other studies using test conditions for both the assessment of intelligence and creativity, among representative samples, have found high correlations between these two abilities (Edwards & Tyler, 1965; Hasan & Butcher, 1966; Cropley, 1966; Lovell & Shields, 1967). Hasan and Butcher (1966), for example, after testing a group of Scottish adolescents, found that intelligence and creativity overlapped to such an extent as to be hardly distinguishable.

The Threshold Theory

To explain these apparent divergent results when examining the relationship between creativity and intelligence, several investigators have proposed the threshold hypothesis. The threshold hypothesis is based on the ability gradient theory of Anderson (1960) which suggests that intelligence exerts an effect upon creative performance up to a certain threshold level, beyond which increments in IQ do not effect such performance, and factors of creativity function independently. According to this theory, a certain level of intelligence is required for creative thinking but is not a sufficient prerequisite. Below this minimum threshold, creative functioning is largely limited by the level of IQ. Above this threshold level, if creative abilities are present, creative functioning begins to depend on factors other than merely IQ. The theory suggests that substantial correlations will exist between IQ and creativity at the lower intelligence levels, but after some critical IQ the relationship will diminish so that IQ and creativity become independent. Torrance (1962) proposed an IQ of about 120 as being the IQ threshold beyond which intelligence and creativity become independent.

The threshold theory has found support in such correlational studies as Torrance (1962), Yamamato (1964, 1965); and, Haddon and Lytton (1968). Cicerelli (1965) found very weak evidence for the hypothesis, finding instead that, in general, the relationship of intelligence and creativity was of a linear nature and additive in its effect upon achievement.

Evidence disputing this hypothesis was found by Ginsburg and Whittemore (1968) who found that the relationship between intelligence and creativity was curvilinear with a dramatic change in the shape of the curve after an IQ of 120. In addition, these investigators did not find the large variances in creativity scores at the upper IQ range that the threshold theory predicts.

Lytton and Cotton (1969) also found evidence which did not support the threshold theory, the correlation coefficient between their convergent and divergent measures decreasing and becoming negative for their lower IQ group.

The results reported by Guilford and Christensen (1973) also did not support the threshold hypothesis, showing instead a continuous relationship between the two measurements so that the higher the IQ the more likely the appearance of creative potential. Scatter-plots on the visual-figural tests tended to be elliptical, whereas those from the semantic measurements tended to have a triangular relationship with IQ, meaning that at low IQ levels there were no high divergency scores but at high IQ levels divergency scores could be high or low.

Thus, the relationship between intelligence and creativity is still controversial, as is the question of what type of conditions or regime best enhances the performance on divergent tests. These questions could well be asked with respect to convergent tests. Wallach and Kogan (1965) posed such a question postulating that changing the atmosphere in which intelligence tests are administered may change its relationship with creativity. They express doubt as to the effect of play conditions on influencing children's performances on intelligence tests, however, because of their convergent nature. They believe that children are sufficiently sensitized to intelligence tests so that the cues of having to provide only one right answer would set them into a test-taking mood.

The present investigation did not take this view, but rather proposed that test and play-like atmospheres would produce different results on convergent and divergent thinking tests. Lundsteen (1966) further suggested that if intelligence tests are given in a game-like context the correlation between convergent and divergent thinking may be substantial, thus paralleling the correlation found between the two constructs in many studies using test conditions only. Boersma and O'Bryan (1968) echoed this prediction. If this were so, there would be a strong argument to negate the claims that intelligence and creativity are separate dimensions of the intellect. It would support the notion that convergent and divergent thinking are complementary components of the human intellect and that such abilities are not independent of, or exclusive from, the general factor of intelligence. This is the view favoured in the present thesis.

Furthermore, the review of the literature on convergent tests has convinced the writer that a large, if not the major, part of individual differences in convergent thinking rests upon a heritable basis. The author would also adopt the position for a similar view for divergent thinking, although the evidence here is not substantial.

The stated purpose of the present thesis was to examine the conditions of test administration upon the level of performance on convergent and divergent tests. To a considerable extent, therefore, in view of the author's stance stated above, the importance of this effort is justified on the grounds of defining the conditions under which children express, to an optimal degree, their intellectual endowment. This means that the position is taken that whether potential ability is low or high there are certain optimal conditions for assessment of a child's capacities, irrespective of his level of ability. Furthermore, the outcome of this study should reveal more about the nature of convergent and divergent abilities. The conceptual position taken by the writer throughout the review of the literature is that convergent and divergent abilities are two different functional entities sharing a common attribute. Thus, although there is some general intellectual factor, such as g, common to both, convergent and divergent thinking have also discrete ability components similar to the second order factor abilities proposed by investigators such as Burt and Vernon.

Since prior research suggests that divergent thought productivity is susceptible to environmental contextual changes, it may be a more labile ability. The lability of convergent thinking under different administrative regimes has not yet been investigated prior to this research, although it is likely to be a more stable mental function in view of its demonstrated relative stability over the early part of the life span at least, a characteristic described in Chapter II. In view of this it would be expected that the relationship between convergent and divergent abilities may vary according to the administrative context. Where convergent and divergent tests are taken under similar conditions the degree of correlation should be higher than when the two types of tests are administered under different regimes. If divergent thinking is the more labile of the two abilities then it should be more influenced by a change in the traditional test assessment procedure so that separation of the two abilities should be greatest when the convergent and divergent tasks are administered under formal and informal conditions, respectively.

This predicted change in relationships with different administrative conditions does not necessarily mean that the basic nature of the two abilities has also changed. Unless the correlation coefficients between convergent and divergent thinking are extremely low, separate dimensionality of the abilities cannot be claimed. Rather, the expected change in relationship with varying conditions may be a reflection of the possibility that convergent and divergent abilities are discrete second-order components of the intellect, a feature which is enhanced when each is taken in a different environmental context.

It was within this theoretical framework, therefore, that the present investigation was conceived. The fundamental question that this study proposed to answer was whether levels of performance on a battery of convergent and divergent thinking tests changed under different administrative regimes. Related to this question was whether the relationship between convergent and divergent thinking changed significantly under test or play conditions. With respect to these questions, a number of hypotheses and predictions were made.

Hypotheses and Predictions

1. Order of giving the "tests" will not affect the level of performance on the convergent tasks.
2. Order of giving the "tests" will not affect the level of performance on the divergent tasks.
3. Being subjected to two types of conditions will not affect the level of performance on convergent "tests."
4. Being subjected to two types of conditions will not affect the level of performance on divergent "tests."
5. Schools subjected to test conditions for the convergent tasks will perform at a higher level than schools subjected to the game conditions.
6. Schools subjected to game conditions for the divergent tasks will perform at a higher level than schools subjected to the test conditions.
7. The relationship between convergent and divergent "tests" will vary according to the conditions under which tests are taken, specifically,

(a) The correlation will be highest when both "tests" are taken under formal conditions.

(b) The next highest correlation will be obtained when both "tests" are taken under game conditions.

(c) The next highest correlation will be obtained when convergent "tests" are taken under game conditions and divergent tests under formal conditions.

(d) The lowest correlation will be obtained when convergent tests are taken under formal conditions and divergent "tests" under game conditions. This hypothesis is proposed since it is predicted that divergent tasks will be more affected by the change in conditions.

8. Order of giving "tests" will not affect the relationship between convergent and divergent thinking.

CHAPTER IV

METHODOLOGY

Subjects

The subjects for the present study were pupils in the fifth grade of elementary State Metropolitan schools in Adelaide. The mean age was 120.22 months with a standard deviation (S.D.) of 5.75. A total of 493 subjects, were used in the main study, completed all "tests" and were included in the analyses. Of these, 253 were boys and 240 were girls. The mean age of the boys in the sample was 120.19 months (S.D. = 5.72) and that of the girls was 120.25 months (S.D. = 5.78).

These children were studied in nine schools in the City of Adelaide. Two classes of the fifth grade were studied in each school, except for two schools in which only one class participated. Each of these single classes were subjected to the same conditions and for experimental purposes were treated as the same school. Thus, although there were nine schools involved, there were only eight different conditions. The number of children from each school ranged from a low of 47 to a high of 69, with a median of 63.

Table 2 shows the distribution of children by school, age and mean Otis IQ. The number of children from each school included in the final analysis are also tabulated.

Testing of these children from each of these schools occupied most of the academic year from March through to December. The difference in ages, generally speaking, reflects the difference in time between each school when testing took place. Data were collected from each of the schools in the following order:

1. School 1
2. School 5
3. Schools 4a and 4b
4. School 7
5. School 8
6. School 3
7. School 6
8. School 2

As would be expected, School 2, which was tested during December, has the highest mean age.

The average mean raw Otis score for the whole sample is 36.86 with a standard deviation of 12.10. This gives a mean level of IQ of 115 and a mean IQ range of 109-119. (Schools 4a and 4b are treated as one school.)

Table 2

Average Ages and Otis Scores for All Schools (Both Sexes Combined)

School	N	Mean Age (in months)	S.D.	Mean Otis Raw Score	Mean Otis IQ	S.D.
1	65	117.55	5.03	30.82	109	12.25
2	61	125.09	4.65	38.72	117	10.71
3	60	121.03	4.04	36.95	115	10.62
4a	33	119.73	4.52	40.27	118	10.44
4b	30	121.67	7.58	30.43	108	11.39
4(a&b)	63	120.70	6.05	35.35	113	10.92
5	65	116.81	4.81	33.83	112	12.11
6	63	123.63	5.06	40.46	118	13.53
7	69	118.41	4.61	38.80	117	9.77
8	47	118.25	6.49	40.87	119	12.99

Tables 3 and 4 show the distribution of boys and girls, respectively, by school, number in sample, age, mean Otis score and Mean Otis IQ. The average mean Otis score for the 253 boys is 36.70 with a standard deviation of 12.13. The mean level Otis IQ for boys is thus 115 and the mean IQ range is between 108 and 120. A maximum mean level IQ of 120 is high and was not expected in this sample. The high mean IQ level among the boys in School 4a is balanced by the relatively low mean IQ level of the boys in School 4b to give an Otis score which is nearer to the general mean.

The average mean Otis score for the girls was 37.02 with a standard deviation of 12.09. This again yields a mean IQ score of 115. Thus, both boys and girls, in the sample overall, have an equivalent mean IQ. The mean level IQ range for girls is from 110 to 118.

There are no extreme observable differences in mean Otis scores between boys and girls for any one school, the greatest difference being 4.05 raw score points or 3 IQ points between boys and girls in School 7. Similarly, there is little difference in ages between the sexes within any one school, the maximum being approximately two months in Schools 3, 6 and 8 where the girls have a slightly higher mean age.

The children were white, English speaking and generally

Table 3

Average Ages and Otis Scores
for All Schools (Boys)

School	N	Mean Age (in months)	S.D.	Mean Otis Raw Score	Mean Otis IQ	S.D.
1	31	117.55	5.03	30.03	108	12.40
2	35	125.09	4.65	38.49	117	11.60
3	32	121.03	4.04	36.31	115	11.99
4a	15	119.73	4.53	41.67	120	9.98
4b	12	121.67	7.57	27.08	105	9.28
4 (a&b)	27	120.70	6.05	34.38	113	9.63
5	32	116.81	4.81	33.63	112	12.80
6	30	123.63	5.06	41.47	119	12.23
7	42	118.41	4.61	37.21	116	9.69
8	24	118.25	6.49	42.21	120	11.60

Table 4

Average Ages and Otis Scores
for All Schools (Girls)

School	N	Mean Age (in months	S.D.	Mean Otis Raw Scores	Mean Otis IQ	S.D.
1	34	117.44	4.60	31.53	110	12.25
2	36	124.08	5.51	39.04	118	9.58
3	28	123.46	4.96	37.68	116	8.96
4a	18	118.72	4.79	39.11	118	10.95
4b	18	118.44	3.50	32.67	111	12.34
4 (a&b)	36	118.58	4.15	35.89	115	11.65
5	33	116.03	3.99	34.03	113	11.60
6	33	125.18	5.82	39.55	118	14.75
7	27	117.59	4.10	41.26	119	9.57
8	23	120.87	5.32	39.48	118	14.42

middle-class. Pupils of Grade 5 were chosen, because in Adelaide a standardized group intelligence test is given to these classes as part of the state assessment system. Results of this test (the Otis) were used as a control variable. The Otis is also given in Grade 7 before the students leave for secondary school. However, it was felt that the younger children would be more suitable for this study, particularly with reference to successfully affecting game-like conditions for the convergent tests. In addition, it was assumed that minimal subject specialization would have influenced the subjects' cognitive styles.

Sampling of Schools

The primary schools were selected from the total number of state schools in Adelaide from considerations based on the ecological structure of residential areas such as indices of socioeconomic status, familial status (household composition), ethnicity, numbers of migrants, age structure, and growth areas of new developments. Specifically, the schools were chosen on the basis of findings from Stimson's (1971) study of social differentiation in residential areas in Adelaide.

The selection process was designed to avoid areas of high ethnic concentration and recent migrants. It was desired to achieve selection of schools from middle-class areas which were not recent land development areas or which were unusually old in population structure. Thus, the selection of schools was done to produce groups of children for study who represented middle-class, white collar families who had been relatively long term residents.

The sampling of schools was done from the Stimson (1971) maps in which he showed divisible regions of Adelaide according to various characteristics. Variables considered were the following factors isolated by Stimson:

(a) Socioeconomic status which was defined by occupational level;

(b) Familism, which was defined by household composition, such as the percent of the non-single or widowed population aged 15 years and over; the percent of the population aged 60 years and over; the percent of males not at work; pensioners; and the percent of private dwellings and flats, all factors found to be indicative of low familism;

(c) High occupancy by United Kingdom migrants meaning those areas recently occupied by migrants from that country;

(d) High ethnicity which refers to a factor which has a high association with European, as distinct from English migrants;

(e) An aged low familism factor which refers to those communities with a large number of retirement settlements.

Each of these factors influenced the sampling in the present study and the relevant ecological maps on which this selection of schools was based are found in Appendix A.

The first map, in Figure 1, represented by factor I represents socioeconomic status. Stimson found that high socioeconomic status was found to be associated with those males who had employer status, were in businesses associated with finance and property, professional and business services, professional and technical occupations and those males who had university qualifications. Low socioeconomic status included those males who had employee status, were in manufacturing, craftsmen-production or labouring occupations and those females working in manufacturing type jobs.

Adelaide can be divided into certain socioeconomic areas. A loading of +2.00 and over represents the highest status suburbs, while loadings of -1.50 and below represent the lowest status suburbs. The latter areas were excluded from this study. Most schools were chosen in the areas with factor loadings of +1.50 to +1.99. Inevitably there were some children in the sample who lived in areas with +2.00 factor loadings.

Factor II, in Figure 2, represents the household composition. Areas with very high positive scores and high negative scores on this dimension were avoided. High scores are associated with low familism, which includes those areas where single, widowed, divorced or separated people live. The aged, low familism factor is represented by factor V in Figure 5. This dimension is identified with an ageing population. Such areas were also excluded. High negative scores on factor II represent high familism and are associated with recently developed suburbs.

Factor III, in Figure 3, representing the percentage of recent United Kingdom migrants was only taken into consideration with respect to those areas which have a high loading on this factor, and this again tended to be associated with recently developed areas, so was avoided in the sampling. The reason for taking this factor into consideration was to ensure that the sample had children of long-standing Australian residency.

Australia is an example of a country where economic growth and demographic growth has depended a great deal on immigration, particularly for its work or labour force. Thus, there is a close relationship between social stratification and ethnic affiliation. With respect to factor I, for instance, high socioeconomic status is associated with low ethnicity. Factor IV, in Figure 4, identifies this dimension as such. High ethnicity is associated with

European as distinct from United Kingdom migrants. In particular, it refers to those migrants from countries such as Italy, Greece, Yugoslavia, Malta and Hungary. There is a tendency for particular ethnic groups to live in the same locality. These areas were also excluded from the sample since the introduction of large numbers of migrant children in the sample could have contaminated the results. Each school selected did have a few children who had Greek or Italian parents; however, their numbers were negligible. Factor VI, in Figure 6, represents high Northwestern and Eastern European ethnicity involving migrants from Poland, U.S.S.R., Yugoslavia, Austria and Germany. These areas were also excluded.

Figure 7 in Appendix A shows a map of Adelaide. Numbers refer to particular suburbs which are identified by name in the key.

The following schools were finally selected:

School 1	Paringa Park (area 114)
School 2	Mitcham (area 63)
School 3	Warradale (area 75)
School 4a	Belair (area 65)
School 4b	Lockleys (area 86)
School 5	West Beach (area 109)
School 6	Burnside (area 55-56)
School 7	Clapham (area 62)
School 8	Glen Osmond-Glenside (area 53)

School 6 is located on the border of areas 55 and 56, and draws many of its pupils from area 56, which has a +1.50 to +1.99 SES classification and which was the reason it was included in the study. Thus, it was felt that the sample was a good representation of the particular "class" of Australian child that this research intended to study, a description of which has already been stated.

Tests and Measurements

A. Selection of Tests

There are a large number of convergent and divergent tests which have been used in previous investigations, so that it was not felt that special tests had to be constructed for this research. In addition, the use of already constructed "tests" eliminated the necessity of having to do one's own reliability and validity studies. It was decided to use a battery of both verbal and non-verbal tests for each cognitive mode. This comprised two verbal convergent tests, two non-verbal convergent tests, two verbal divergent tests and two non-verbal divergent tests. Selection of actual tests was influenced by several factors. They had to be tests which could be administered in group form, since this was the intention of the study. They had to have similarity with test instruments that had been used in other studies so that the results of this research were comparable. The nature of the tests had to be clearly either convergent or divergent so that there was no risk of contamination. Thus, measures such as the Remote Association test were excluded because, although previously used as a divergent measure it has convergent components to it. A very serious consideration was selection of measures which could convincingly be administered in group form, in the classroom, in a game-like context. Thus, several weeks were spent in areas not selected in the research, interviewing and "testing" children and adolescents with possible test instruments. These students were asked which tests appeared more like games, and how they would give them to other children as games. There was no lack of volunteer subjects. Students already tested brought their friends, and so on. Many of them had constructive suggestions, particularly the adolescents. This information was pooled and from it the convergent and divergent test batteries were chosen, and a decision on both the test and game administrative contexts was made.

B. <u>Convergent Tests</u>

A sample of these tests are given in Appendix B. These tests are available through published sources (Raven, 1961, 1965; Wechsler, 1952). There were no time limits on any of the tests.

1. Similarities: This is one of the verbal tests from the Wechsler Intelligence scale for children (WISC). Normally, this test is administered in individual form to subjects as part of the WISC battery. In this study, it was administered as a group test. All 16 items were given. Each question was given verbally by the investigator and then written on the blackboard. Children wrote down their answers. Each child's answer was checked for both questions 1 and 2 and the answers told to the group to ensure that every subject understood what was required of him or her. Such aid to the child is advised in the manual. The manual gives instructions to discontinue the test if there are three consecutive failures. In this research, the subjects were encouraged to attempt an answer to every question.

Similarities is a verbal reasoning test where the subject is required to analyze the properties of two objects to determine in what way they are alike; for example, in what way are a plum and a peach alike? The test has a high correlation with the total verbal WISC score (.72) and with the full-scale WISC Score (.65) (Wechsler, 1952). This sub-test has also a fairly high correlation with the Wechsler performance or non-verbal scale of .48.

2. Coloured Progressive Matrices (CPM). This is a perceptual non-verbal reasoning test which Raven (1965)

claims has a high relationship with _g_. Raven regards the test as measuring an individual's present capacity for intellectual activity. By itself it is not regarded as a test of general intelligence. To get such an estimation it is used in conjunction with a vocabulary test. In the CPM there are three sets of 12 problems, totaling 36 problems in all. The test was administered in group form. Children were given the test booklet and appropriate answer sheet. No time limits were imposed. On the first pattern, the problem to be solved was discussed and the subjects informed what they were to do. Following this, the children attempted the first three test items which were then checked with each group of children to ensure that they understood the nature of the task. Subjects were warned against turning over more than one page at a time and the administrator checked throughout the "test" to ensure this was not happening.

3. Picture Completion Test. This is one of the nonverbal tests from the WISC. This is usually administered in individual form by presenting the child with a series of pictures, one at a time, which the child has to respond to by pointing out what is the most important thing missing. For example, on one card there is a coat with some buttonholes missing. It was stressed that only one response could be given. The cards used for individual administra-

tion are too small for group administration. A professional artist was employed to redraw enlarged pictures of each item on 20 x 12 (inches) white cards. One of the cards from the WISC was eliminated on the grounds that it was old-fashioned and inappropriate. This was the item of a hat from which a hatband was missing. This item is near the end of the test and is presumed to have a certain degree of difficulty. It was replaced with a drawing of an automobile in motion, from which the license plate was missing. During the pilot studies, when this item was used initially, it was found that the children with the highest Otis IQ scores were the only ones to give the correct answer. From this it was judged that the item was of the right level of difficulty and suitable for inclusion in the test.

For actual administration, each group of children was put into a fairly compact group and the investigator presented the cards one by one by holding them in a suitable position so that each person could see. For each card only 15 seconds of exposure time was allowed as per the instructions in the manual (Wechsler, 1952). Although the manual instructions are to discontinue the test after four consecutive failures, all 20 items were given to all children. Answers were checked with each group for the first two questions as the manual suggests.

The correlation coefficients of the Picture Completion test with other WISC scores are lower than those correlations reported for Similarities. The intercorrelations of Picture Completion with the total non-verbal score is .48, with the verbal score .45, and with the full scale score .51 (Wechsler, 1952).

4. The Crichton Vocabulary Test. It was felt essential to include a comprehensive vocabulary test in the convergent battery. The Crichton Vocabulary test was chosen over the WISC vocabulary test, because there are two parallel sets of the former. By giving children the two sets a reliability coefficient could be estimated between Set 1 and Set 2, to check on the validity of the test when administered in written group form. Normally, the test is given in individual oral form. In this research, each set of words, with plenty of space left for answers, was typed out. The whole vocabulary scale consists of 80 words with 40 words in each set. Order of the words in each set is graded according to difficulty and was based on the frequency with which the sampled 11-year-old children of Raven (1961) were able to correctly explain their meaning. Raven (1961) claims that this vocabulary test provides "an index" of a person's "general cultural attainments" and his "acquired fund of verbal information," (p. 1). Used in conjunction with the CPM it is regarded as a measure of general intelligence.

C. Scoring of Convergent Tests

Manuals are available for the four convergent tests described. Scoring was, therefore, done according to the instructions in the appropriate manual.

1. Similarities. Each of the items 1-4 was scored one point if it was answered correctly. Each of the items 5-16 was scored two, one or zero points, depending on the degree and quality of the generalizations made in the answer. Judgment of this was done according to the criteria set up for each of these questions by Wechsler (1952) and Massey (1969). Thus, on the last item, which asks in what way are the numbers 49 and 121 alike, the answer that they are perfect squares or odd number square roots receives two points; a response which states they both can't be divided by two (three, four, five, etc.) or that they are both odd numbers receives one point; and all other responses, including that which restates they're both numbers receives zero points. The maximum score possible was 28. Raw scores ranged from three to 21.

2. Coloured Progressive Matrices. Each of the 36 items was marked either right or wrong. Each correct answer was awarded one point. The maximum score possible was 36. Raw scores ranged from 12 to 36.

3. Picture Completion. Each correct response was scored one point. No points were awarded for partially

correct responses. Children had been told to write only one answer, namely the most important thing missing. If a pupil did give more than one answer, only the first response was considered. The maximum score possible was 20. Raw scores ranged from four to 18.

4. Vocabulary. Each response is marked right or wrong and each right answer is awarded one point. Raven (1961) decided on this rather than qualitative marking. He feels that "since it is impossible to define any exact principle according to which an explanation of a word's meaning can be shown to be theoretically correct, partially correct, or wrong, any attempt to award double marks for accurate definitions, or half marks for vague explanations, seems unsatisfactory," (p. 3).

This contrasts with scoring in the Vocabulary Test of the WISC where responses are scored two, one or zero depending on the quality and degree of accuracy in the answer. The criteria for scoring the Crichton Vocabulary Scale are set out in the manual (Raven, 1961) and were followed. The maximum score possible was 80. Raw scores obtained from the whole sample ranged from nine to 72. The correlation coefficient between set 1 and set 2 of the vocabulary test, when estimated for the total sample of 493 children, was .85 which is significant at the .001 level. Such a high correlation suggests that the nature of the test was not affected by the use of group administration, the written form or the introduction of game conditions.

D. The Divergent Tests

The specific content of these tests is given in Appendix C. No time limits were imposed on any of the tests.

1. Just Suppose. Three of the questions from this test by Torrance (1962) were chosen. The "test" was presented in written form with each question written at the top of a page with plenty of space for responses below. The nature of the test was orally explained to each group; namely, that each question suggested an unlikely situation and that subjects were required to write down all the different things that could happen if, for example, someone got caught in a big soap bubble and couldn't get out.

One example which was not part of the "test" was orally given, namely: "just suppose you had two left feet." The investigator suggested four possibilities and then suggested to the class that they try the first question of the "test," and to go on to the second and third questions when they'd finished. Pupils were told their answers did not have to be "true" in the sense of being presently possible.

The example given, to illustrate the nature of the "test" was relatively non-stimulating in order that the subjects would not become preoccupied with it rather than with the questions set. During one of the pilot studies the example given was: "Just suppose you could look into the future." Because several students rushed through the test problems and then asked if they could do the example question this problem was excluded and another substituted which did not provoke the same enthusiastic response.

2. Patterns. Four of the patterns were selected from this test of Wallach and Kogan (1965) and presented to the children in group form. The instructions informed the children that on each page of their booklets there were some drawings or patterns, that they were to look at each one and write down all the different things they thought each drawing could be. One example was given and the administrator suggested four possible responses. Each group was told that they could turn the patterns any way they liked but to write down their answers on the paper from the angle they were viewing it. This was demonstrated.

3. Uses of Objects. Three of the objects used by Hudson (1967) were selected. The children were told to write down all the different uses they could think of for each object. They were told that their answers did not have to be "true" in the sense of being possible at the present time. One example was given, namely, the uses of a paper clip for which the investigator suggested four answers.

4. Inkblots. Four of the dark "nonsense" shapes used by McHenry and Shouksmith (1970) were selected. One other dark shape from their battery was used as an example for which four answers were suggested. McHenry and Shouksmith had used 12 of these shapes as a visual imagination test. They had presented their stimuli to their sample of 10-year-old children by the use of slides and a projector. In this investigation the inkblots were presented in written form.

A professional engineer enlarged to exact scale the appropriate stimuli taken from the inkblots presented in the paper of McHenry and Shouksmith (1970). Initially, the Rorschach inkblots as used by Vernon (1971) were planned as the test stimuli; but because of the importance of colour in these inkblots in the evoking of responses, and the desire to use a written form of this type of test, it was decided to use the dark shapes as described above.

As in the Patterns, the children were told that on each page of their booklets there were some inkblots, that they were to look at each one and write down all the different things they thought each inkblot could be. The example was then shown and the investigator suggested four possible answers.

E. Scoring of the Divergent Tests

All of the divergent tests were scored in the same way so that a generalized scoring system will be presented.

The explanation of divergency and the parameters of the scoring system have already been explained in Chapter I of this thesis. Three basic parameters were used, namely, that of fluency, flexibility and originality. Each of these scores was further differentiated to give seven variables for each test involving 28 variables for the divergent battery.

Fluency 1 was defined after Ward (1969; 1973, personal communication). This score involved the total number of relevant responses given, excluding repetitions, and excluding the superordinates of a class of objects if subordinates of that same class of object had already been given.

Ward (1973, personal communication) explained that the way he scored for fluency meant excluding responses with exact duplication of words, or the use of a concept and a superordinate to that concept. Thus, if a subject named the following uses of a barrel, as a storage place, to store things in, to store a toy in, to store clothes in, to store toys in, his fluency score would be two. Changes such as singular to plural are not counted, as in a toy to toys. The exceptions would be in the case of the "tests," Patterns or Inkblots, where a singular could imply one interpretation of the units in the figure and a plural could imply another. In naming the uses of a barrel, a storage place is implied to mean the same as to store things in, and the phrase to store things in is a superordinate concept. These rules were applied in calculating the score of Fluency 1.

Fluency 2 was defined as the total number of relevant responses given excluding repetitions which meant the exact duplication of words or duplication of a concept. Unlike Fluency 1, all subordinates and their superordinates were counted in the score.

Flexibility 1 and Flexibility 2 were the two flexibility scores. They involved a shift of meaning from one concept to another which, in operational terms, involved a shift in thinking from one category to another.

Flexibility 1 was defined by a shift from one major category to another. For the non-verbal stimuli this meant, for example, shifting from the category of animal to animal detail to animals expressing emotion, to monsters, to machines, to humans and so on. For uses of a barrel, as an example of a verbal stimulus, a major shift in categories involved, for example, shifting from using a barrel to store things in, to use as a cupboard, or as a wheelbarrow or as a shower or as a hospital trolley.

Flexibility 2 was defined by minor cognitive shifts within any one major category. In other words, Flexibility 2 is the shift or change in attitude or focus occurring within any one major category. Thus, for example, in the categories for the non-verbal stimuli, the following shifts were awarded a point within the category of "Weapon."

(a) Weapon;

(b) Weapon with elaboration;

(c) Weapon with movement;

(d) Weapon with movement and elaboration;

(e) Weapon being used in an aggressive act or in an aggressive situation;

(f) Weapon being used in an aggressive act or in an aggressive situation with elaboration.

An example of scoring for Flexibility 2 from the uses of a barrel are the shifts that may occur within the major category of "use as an animal building or shelter (excluding fish)," namely:

(a) Shelter for four-legged mammals;

(b) Shelter for birds; e.g., a bird's cage;

(c) Shelter for insects and reptiles;

(d) Shelter for invertebrates; e.g., worms;

(e) Shelter for water animals (excluding fish), such as frogs.

A catalogue of the flexibility scores are given in Appendix D. Because of the similarity in response categories found among the non-verbal tests of Patterns and

Inkblots, the same extensive category list was used for the four stimuli on both tests. Separate category sheets were composed for each stimulus of Just Suppose and Uses of Objects. The numbers in these category lists refer to major categories. Minor categories are indicated by letters of the alphabet.

Flexibility categories were established in the following way. All the responses for each test were read through for the entire sample. This included not only the 493 subjects in the main study, but all subjects included in the pilot studies. The pupils involved in the pilot studies who had completed all convergent and divergent tests was 204. This gave a total of 697 who completed the entire test battery. From these 697 children, two samples of 100 each were randomly chosen. (This involved putting the 697 names in a sack and, after mixing them, drawing out two samples of 100.) These two sets of 100 children (Set A and Set B) were used as the <u>standards</u> for establishing flexibility categories and establishing statistical norms of originality.

From the information gleaned from reading through the responses of the entire sample of 697, a rough outline was established for the flexibility categories of the non-verbal tests and of each stimulus of the verbal tests. Guidance for establishing categories was gained through the practice used in characterizing content scores in the Rorschach (Beck, 1949). Then each stimulus for each test was systematically gone through in the sample of 200 in order to finalize categories. Next, this standard sample of 200 was again systematically marked to categorize each response of each stimulus. Frequencies of each minor and major category were recorded for each set of 100. Finally, the remaining protocols were scored for. If new categories appeared they were listed as miscellaneous.

These are duly recorded in Appendix D as miscellaneous. The noteable point is that due to the thoroughness involved in establishing categories there are very few miscellaneous answers.

<u>Originality</u> was defined in statistical terms of rareness of response. There were three originality scores. Originality 1 refers to originality of flexibility 1 categories. Originality 2 refers to originality of flexibility 2 categories. Thus, there are two originality scores in terms of frequency of usage of categories; namely, the use of major and minor categories. Originality 3 refers to originality of content. As mentioned above, frequency of use of major and minor flexibility categories were recorded for two sets of 100 pupils. The actual frequency scores for both sets of 100 for both major and minor categories were very similar, the correlation coefficients between the two samples being .95 and .94 for the respective

groups of categories. Therefore, the two samples were combined for the purposes of establishing originality scores.

For originality of content the same standard of 200 protocols were used. Every response for each stimulus on each test was recorded on index cards and the frequency of response recorded. As Vernon (1971) noted, this often becomes difficult as to whether certain responses differ in content. Such problems were encountered in all the divergent tests in the present study. One such example occurred in a protocol of the Uses of Objects questionnaire. One girl listed 86 different things that could be stored in a barrel. From this and other experiences it was decided that in circumstances such as this if the response was "common" in the sense of being not unusual, it should not be regarded as unique. Although such decisions involve some subjective evaluation, this line of reasoning does seem pragmatic and follows the same guidelines used by Vernon (1971).

On this basis originality points were established on the standard sample. The scheme of Vernon (1971) was followed, although it should be noted that Vernon used a sample of 100 as his standard and awarded a maximum originality score of two. Because of the greater number in the sample of the present investigation relative, but comparable, originality scores were established with the maximum number of points awarded being three. The marking scheme for originality of content is shown in Table 5. Since all three originality scores were based on frequency scoring, this scheme was used for all originality scores. Examples of some of the childrens' responses on each of the tests is given in Appendix E.

F. Reliability of Divergent Test Scores

Each score for each stimulus of every divergent "test" used in this study involved to a greater or lesser extent some subjective evaluation. Therefore, it was felt that some reliability of "blind" marking should be established with three independent raters. A sample of 100 children was chosen, namely Set A of the standard sample described in the previous sub-section. The first stimulus on each of the four tests was chosen.

The raters were one post-doctoral psychology student, one graduate psychology student and the present investigator. Several sessions were involved in discussing the category schemes for the stimuli involved. The first task involved the marking of another independent sample of tests according to category, for practice. These evaluations were then discussed and appropriate adjustments made to the flexibility categories. Then the raters independently marked the standard sample of Set A on all four tests for Fluency 1,

Table 5

Originality Marking Scheme
for the Divergent Tests

Based on a Standard Sample of 200

Frequency	Score
4 or less	3
5-9	2
10-29	1
30 or more	0

Fluency 2, Flexibility 1 and Flexibility 2. Each rater also recorded the frequencies of each major and minor category in the cases of flexibility.

The reliability of this marking scheme was evaluated by calculating Spearman correlation coefficients for Fluency 1, Fluency 2, Flexibility 1 and Flexibility 2 on each of the four tests. The coefficients calculated reflect the reliability of scores awarded to each subject on each parameter for each test. All reliability coefficients are high, with an approximate .9 as the average. Tables 6-21 reveal these results.

Rater 1 is the present investigator; rater 2 is the psychology graduate student, and rater 3 is the post-doctoral student. A significance level of .02 was chosen. All results are significant.

Where N, the number in the sample is less than 100, as in Patterns where N = 98, and Uses of Objects where N = 99, this is because particular subjects failed to respond on these tests. Such incidences in the total sample were rare.

The results of these tables show that scores awarded by different raters on these parameters were reliable. In the analyses in this investigation, only scores established by the present investigator were used for reasons of consistency.

Table 6

Just Suppose Marker Reliability

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
Flu 1 0.89	Flu 1 0.94	Flu 1 0.87
N (100)	N (100)	N (100)
Sig .001	Sig .001	Sig .001

Key:

Flu 1 = Fluency 1

Sig = Level of Significance

Table 7

Just Suppose Marker Reliability

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
Flu 2 0.89	Flu 2 0.93	Flu 2 0.88
N (100)	N (100)	N (100)
Sig .001	Sig .001	Sig .001

Key:

Flu 2 = Fluency 2

Sig = Level of Significance

Table 8

Just Suppose Marker Reliability

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
Flx 1 0.92	Flx 1 0.90	Flx 1 0.91
N (100)	N (100)	N (100)
Sig .001	Sig .001	Sig .001

Key:

Flx 1 = Flexibility 1

Sig = Level of Significance

Table 9

Just Suppose Marker Reliability

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
Flx 2 0.93	Flx 2 0.91	Flex 2 0.91
N (100)	N (100)	N (100)
Sig .001	Sig .001	Sig .001

Key:

Flx 2 = Flexibility 2

Sig = Level of Significance

Table 10

Patterns Marker Reliability

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
Flu 1 0.94	Flu 1 0.98	Flu 1 0.93
N (98)	N (98)	N (98)
Sig .001	Sig .001	Sig .001

Key:

Flu 1 = Fluency 1

Sig = Level of Significance

Table 11

Patterns Marker Reliability

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
Flu 2 0.95	Flu 2 0.97	Flu 2 0.95
N (98)	N (98)	N (98)
Sig .001	Sig .001	Sig .001

Key:

Flu 2 = Fluency 2

Sig = Level of Significance

Table 12

Patterns Marker Reliability

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
Flx 1 0.96	Flx 1 0.96	Flx 1 0.96
N (98)	N (98)	N (98)
Sig .001	Sig .001	Sig .001

Key:

Flx 1 = Flexibility 1

Sig = Level of Significance

Table 13

Patterns Marker Reliability

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
Flx 2 0.97	Flx 2 0.96	Flx 2 0.96
N (98)	N (98)	N (98)
Sig .001	Sig .001	Sig .001

Key:

Flx 2 = Flexibility 2

Sig = Level of Significance

Table 14

Uses of Objects Marker Reliability

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
Flu 1 0.99	Flu 1 0.99	Flu 1 0.99
N (99)	N (99)	N (99)
Sig .001	Sig .001	Sig .001

Key:

Flu 1 = Fluency 1

Sig = Level of Significance

Table 15

Uses of Objects Marker Reliability

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
Flu 2 0.98	Flu 2 0.99	Flu 2 0.99
N (99)	N (99)	N (99)
Sig .001	Sig .001	Sig .001

Key:

Flu 2 = Fluency 2

Sig = Level of Significance

Table 16

Uses of Objects Marker Reliability

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
Flx 1 0.94	Flx 1 0.95	Flx 1 0.97
N (99)	N (99)	N (99)
Sig .001	Sig .001	Sig .001

Key:

Flx 1 = Flexibility 1

Sig = Level of Significance

Table 17

Uses of Objects Marker Reliability

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
Flx 2 0.96	Flx 2 0.96	Flx 2 0.98
N (99)	N (99)	N (99)
Sig .001	Sig .001	Sig .001

Key:

Flx 2 = Flexibility 2

Sig = Level of Significance

Table 18

Inkblots Marker Reliability

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
Flu 1 0.95	Flu 1 0.94	Flu 1 0.98
N (100)	N (100)	N (100)
Sig .001	Sig .001	Sig .001

Key:

Flu 1 = Fluency 1

Sig = Level of Significance

Table 19

Inkblots Marker Reliability

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
Flu 2 0.94	Flu 2 0.94	Flu 2 0.97
N (100)	N (100)	N (100)
Sig .001	Sig .001	Sig .001

Key:

Flu 2 = Fluency 2

Sig = Level of Significance

Table 20

Inkblots Marker Reliability

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
Flx 1 0.85	Flx 1 0.92	Flx 1 0.86
N (100)	N (100)	N (100)
Sig .001	Sig .001	Sig .001

Key:

Flx 1 = Flexibility 1

Sig = Level of Significance

Table 21

Inkblots Marker Reliability

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
Flx 2 0.91	Flx 2 0.90	Flx 2 0.93
N (100)	N (100)	N (100)
Sig .001	Sig .001	Sig .001

Key:

Flx 2 = Flexibility 2

Sig = Level of Significance

Since originality was also a parameter, the frequencies of each used category (major and minor) which were recorded by each rater were compared. Spearman correlation coefficients between each rater were calculated. The probability value decided upon for significance was .02.

Reliability coefficients for the number of times each category is used are all significant at the .001 level but are lower for the non-verbal tests (Patterns and Inkblots) as compared to the verbal tests (Just Suppose and Uses of Objects). Coefficients particularly drop when the minor categories are used. For Inkblots the correlation between Rater 1 and Rater 2 is .48 for minor category usage. This compares with the correlation for Uses of Objects between Rater 1 and Rater 2 where values of .83 and .84 are attained for major and minor categories, respectively. On Patterns the correlation for major categories is .83 but drops to .60 for minor categories in the case of Rater 1 and Rater 2. The relatively high correlations for both major and minor categories of Just Suppose are surprising since the test was regarded by all raters to be the most difficult. The correlations between raters on Just Suppose range from .77 to .84 for the frequencies of major categories and from .62 to .71 in the case of minor categories.

For both verbal divergent tests there were less minor categories than there were for the non-verbal tests, and this is reflected in the correlation coefficients. Reliability decreases and there is less agreement among raters when minor categories are taken into consideration. For reasons of consistency in the analyses of the research only those scores and frequences established by the present investigator were used.

Tables 22 to 29 illustrate these reliability results for usage of major and minor categories.

Procedure

For each school selected in the study, a similar type of initial procedure was used. The investigator first went to the school and spoke at length with the headmaster explaining the purpose and the design of the experiment. Emphasis was placed on the word test, if tests were to be used and games if a play atmosphere was to be used. The headmaster would then introduce the investigator to each of the classes as a research worker at the university interested in children. The investigator would then explain the purpose of the visit. To those classes only having tests it was explained that the children would be doing a series of tests. The children allocated games were told that they were going to be doing a number of familiar and new educational games. Those classes who were having both sets of conditions were told that they were to have some

Table 22

Just Suppose Major Category Usage

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
0.81	0.84	0.77
N (58)	N (58)	N (58)
Sig .001	Sig .001	Sig .001

Key:

N = Number of different major categories used

Sig = Level of Significance

Table 23

Just Suppose Minor Category Usage

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
0.70	0.71	0.62
N (81)	N (81)	N (81)
Sig .001	Sig .001	Sig .001

Key:

N = Number of different minor categories used

Sig = Level of Significance

Table 24

Patterns Major Category Usage

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
0.83	0.83	0.82
N (62)	N (62)	N (62)
Sig .001	Sig .001	Sig .001

Key:

N = Number of different major categories used

Sig = Level of Significance

Table 25

Patterns Minor Category Usage

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
0.60	0.62	0.60
N (59)	N (59)	N (59)
Sig .001	Sig .001	Sig .001

Key:

N = Number of different minor categories used

Sig = Level of Significance

Table 26

Uses of Objects Major Category Usage

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
0.83	0.90	0.86
N (73)	N (73)	N (73)
Sig .001	Sig .001	Sig .001

Key:

N = Number of different major categories used

Sig = Level of Significance

Table 27

Uses of Objects Minor Category Usage

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
0.84	0.82	0.85
N (79)	N (79)	N (79)
Sig .001	Sig .001	Sig .001

Key:

N = Number of different minor categories used

Sig = Level of Significance

Table 28

Inkblots Major Category Usage

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
0.75	0.78	0.67
N (76)	N (76)	N (76)
Sig .001	Sig .001	Sig .001

Key:

N = Number of different major categories used

Sig = Level of Significance

Table 29

Inkblots Minor Category Usage

Variable Pair	Variable Pair	Variable Pair
Rater 1 with Rater 2	Rater 1 with Rater 3	Rater 2 with Rater 3
0.48	0.63	0.54
N (190)	N (190)	N (190)
Sig .001	Sig .001	Sig .001

Key:

N = Number of different minor categories used

Sig = Level of Significance

educational games and some tests. The major purpose and design was also explained to the teachers so that their cooperation was ensured.

Following this, the researcher then spent a period of two weeks in the school under study. This was to establish rapport with the pupils and to emphasize what particular aspect of their schooling the investigator was interested in. Thus, if a test atmosphere was to be used only the formal school classes were observed. The researcher took an active part in their lessons, test, and expressed interest in their school progress. If a play atmosphere was to be used, the investigator only attended those classes where there was a more play-like atmosphere, such as art, drama and physical education classes. For those who were going to have both conditions, a balance of both kinds of activities were attended. However, in the last few days of the observation period, only those classes most pertaining to the first condition to be presented were visited.

Next, each child was interviewed. This again allowed the researcher to establish a better rapport with the subjects, to gather more information about them and to explain further the nature of the "testing." During this interview a projective personality test was administered, the "Hand Test." This short "test" designed by Bricklin, Piotiowski and Wagner (1970) presents 10 cards, nine of which have hands in different positions. The subject is to verbalize what he thinks the hand might be doing. On the last card which is blank the child is to imagine a hand and state what it might be doing. The test elicits very little anxiety and the children appeared to enjoy the task. All the responses that each child made were recorded and scored for the purposes of a future research project.

Children came to the interview room two at a time. This allowed the investigator to set up a runner scheme to get the next pupil and save time. The child waiting outside the room was given a drawing to do. Hudson (1967) had suggested that a drawing could be a suitable divergent test. Half the schools were given the drawing topic "Cat Walk," and the other half "Zebra Crossing." Both topics were thought to be capable of producing a wide range of responses. However, in the present task, Cat Walk produced a very small range of different responses and could not, therefore, be considered a good task if the intent is to elicit a wide range of divergent or creative responses. On the other hand, the task to draw a Zebra crossing was much more evocative and showed more embellishment as well as variation in the theme. If such tasks are to be scored the investigator must be aware of and explore the extent to which a subject's responses are inhibited by the particular task or stimulus. Also, it would be a major task to develop a system for evaluating the subject's responses with regard to quality of theme, integration of detail, originality and the weight to be placed upon embellishment of details. This is a formidable task well beyond the intent and scope of this investigation. A great deal of effort needs to be expended in this area which still may be looked upon as one with limited validity and reliability.

The use of the personality test and the drawing also served another very major purpose. Many of the tests to be used in the convergent and divergent battery were new to the subjects, and it was essential to try to overcome a possible Hawthorne or novelty effect (Cook, 1968). By having all children exposed to two novel situations prior to testing there was a better chance that the level of performance achieved during the experiment was due to the administrative conditions. All children who were included in the formal analyses had to have had this interview, personality "test" and drawing. This number, as mentioned before, totalled 493. The number who took part in some of the "testing" procedures, excluding the pilot studies, totalled 547. This gave a mortality rate of 54 or 10.95 percent, which was quite low when one takes into consideration the amount of "testing" involved. The number of children involved in the pilot studies totalled 204. This gave a grand total of children tested at some time of 751.

For the experiment proper, each class was divided into two, with roughly equal proportions of boys and girls. The average size of a class was 30. Most commonly, therefore, there were 15 members in each experimental group. This was felt to be a manageable size, particularly for the game conditions. Group membership stayed constant throughout the experiment.

Each of these groups was "tested" by the investigator over four sessions with two "tests" per session. The order of presentation for the convergent tests was:

Session 1:

(a) Similarities

(b) Coloured Progressive Matrices

Session 2:

(a) Picture Completion

(b) Vocabulary

The order of presentation for the divergent tests was:

Session 1:

(a) Just Suppose

(b) Patterns

Session 2:

(a) Uses of Objects

(b) Inkblots

At no time were there any test manuals in sight. The investigator had all instructions memorized before testing the first school. There were no time limits for any of the conditions, and no unobtrusive timing was applied. It was the intention to have this variable constant over all conditions and the introduction of time limits was felt to be detrimental to establishing a play atmosphere. In addition, the intent of the researcher was to facilitate the production of unusual or original responses on the divergent tests. Previous studies such as Wallach and Kogan (1965), Ward (1969), and Cropley (1972) have found that there is a significant tendency for original responses to appear more commonly in late responses.

The question of length of time did not become a problem in any of the schools. Test sessions took 30-45 minutes. Play sessions usually took an hour, except for the second convergent session which took 1-1/2 hours.

The collection of the data of the Otis was done by the school, namely, the teachers of the particular class. This was the normal procedure in the state school system. The researcher merely collected the scores and IQs from the school at a convenient time.

Conditions

Test atmosphere. Due to the long observation period in each of the selected "test" schools, there was no doubt in the children's minds that they were getting a number of different kinds of tests. The children also knew that their teachers and headmaster were going to see the results, so that the motivation was fairly high. The atmosphere was friendly but firm so that conditions could be described as moderately stressful. If some children finished before others they had to sit at their desks and read a book quietly.

Game Atmosphere. After the last interview each class of children had 1-1/2 hours of play time in the classroom, with a variety of indoor games with which most of them were familiar. These included card games, drawing with special instruments, and word games including Junior Scrabble. The session was organized in the sense that children were placed in small groups and moved at certain time intervals to another game. At the end of the session the children were told that they would be having more such games and other games with which they were not familiar. This included word games. They were also told that in one session they would be having a Scrabble competition with Senior Scrabble. This intrigued them because there is quite a qualitative difference between the Junior and Senior Scrabble sets and most of them at home were still made to play with the former set.

Convergent Games

Session 1. The session began with games of cards. Children sat in small groups and played "Donkey" or "Old Maid," and finished up with a game of "Snap." The administrator then suggested they play another type of game called "Word Snap" (the game name for Similarities).

Small pieces of paper had been placed on desks which had been arranged so that the group was divided in two sub-groups. Each child then selected from the administrator an index card on which there was a number. The child placed his name and number on a slip of paper and gave it back to the administrator. For the remainder of the game, for each answer, the child put only his number on a paper, as well as the appropriate answer. This procedure was explained to the children. For each problem the child was to put his number on the paper, "Snap" the answer out by writing it on the paper and then fold the paper into four. These papers were collected by two of the subjects, one from each sub-group, in a cardboard box which had a split in the lid, so that the children "posted" their answers in. Each sub-group started collecting papers at the same time and points were awarded for the first box on the administrator's desk. The rules of the game were that everyone had to have a fair chance of snapping out the answer without any peer pressure. This procedure went on for all 16 questions. The team fastest with the collection of boxes was announced, although on most occasions the teams were equal, and the person with the administrator's selected number was asked to announce himself or herself. This was one of the most popular games despite the fact that there were 16 questions to answer. It was one of the most unpopular tests! In addition, the children during the game kept to all the rules, particularly in letting everyone have a fair chance of giving an answer to each question. There was only one case in the whole sample where a child forgot to put her number on several answers and she was easy to trace.

"Word Snap" was followed by the "Mystery Gap Game," the name given to the CPM. The administrator had placed brown covers on the CPM booklets and printed the heading "Mystery Gap Game." Their answer sheets too had the heading "Mystery Gap Game" instead of the CPM. Children were given the booklets and answer sheet and normal administration followed.

Session 2. The session began with the game of "Pin the Tail on the Donkey," with each member of the group getting one turn. Since most children miss the target when they are blindfolded, the administrator made the comment that the game could be called "what is always missing from the donkey," and then suggested another game where players had to select what was missing from a series of pictures. Their answer sheet had the heading "The What's Missing Game." Administration then followed the normal procedure.

This was followed by a rapid series of word games, a crossword puzzle, the making of words and anagrams. Children were put through these series of games at a very rapid pace, and the level of difficulty was sufficiently high that no one could get them all finished correctly. They were then reminded of the word games in the Adelaide children's newspapers, and of their up and coming Scrabble game. It was explained to them that in order to do these games well one had to be able to know words and their meanings. They were told that there were 80 words in front of them to see if they could put down their meanings. When they were finished they were given another word scrabble sheet game, namely anagrams, to take home to see if they had improved. Prior to the vocabulary game the children were told of their take-home anagram game and that the word meaning game would help them to do it. Several of the vocabulary "test" words were on this anagram list.

The day following the second session of the convergent games each class had a Scrabble competition with Senior Scrabble. The winner was announced but no prizes were given out. The children had been informed of this during the first session. It had been emphasized that games were for fun as well as being educational.

Divergent Games

Session 1. The session began with each group experimenting with new instruments and coloured pens which allowed them to create a variety of shapes and drawings according to their own whims. Ideas were suggested to each of them according to their skills as to the possible direction their art efforts might take. For example, if the subject drew four dots in a circle and then used his colours and instruments he was challenged with the problem of possible outcomes. The type of approach used was: "Just suppose I drew such and such what would your drawing produce?" After this, Just Suppose was introduced as another game of "Just Suppose," whereby the subjects had to write down the certain consequences which might happen as a result of certain situations.

The game "Patterns," in the same session, was introduced in much the same way. The stimuli were presented as patterns, similar to those that they had been previously drawing. In this game, however, the children were to write down all the different things these patterns might represent.

Session 2. This session began with a game called "Twister" which involved two children at a time competing against each other to place their hands or feet on certain dots according to the call of the referee. Everybody had two turns after which the children sat relaxed on the floor and the experimenter explained to them how expensive games like the one they had just played could be constructed out of everyday materials such as paint and an old piece of canvas. (Many of the children actually did this and kept the game at the school or in their home.) Following this the experimenter introduced "Uses of Objects" as a different kind of game whereby the subjects were to think of as many different uses for each of the stimuli as they could.

This "game" was followed by Inkblots. The "Inkblots" game was introduced to them by reminding them of something that had been discussed with them in the interview. Every child who was to have divergent games had been asked in the interview as to whether he or she ever had trouble falling off to sleep, especially during the summer when daylight saving was in effect. Nearly every child had answered in the affirmative. The investigator told them they could now play a game of "Shadows" on those nights they couldn't fall asleep by making different shapes on the wall with their hands and thinking of all the different things the shape could be. The game of Inkblots was also called "Shadows," and played in much the same way. The children were to write down all the different things the "Shadows" (Inkblots) could be.

In the case of all the divergent games if any child finished before the others he was given other examples of <u>the same type of task</u> to keep him or her occupied until the others had finished. During the convergent and divergent games the children could sit where they chose, on the floor, at a desk, etc. The children were well-behaved during the "test" games having been told in the first introductory game session that there were some games such as Scrabble where it was not only important to keep your ideas to yourself but to be relatively quiet while you were thinking. The aim of these game sessions was to create a balance between excitement, enthusiasm, and order and, fortunately, the children cooperated while at the same time appearing to be relaxed. Only isolated cases of cheating appeared. The investigator was able to catch these immediately, pointing out to the child that he or she was capable of producing good ideas alone, and that he or she didn't have to resort to cheating in games. Peer pressure was also strong and there was no case of a child trying to cheat twice.

The investigator feels that the game context was successfully executed. Ratings of the games were taken at the end of the convergent and divergent sessions, and although the game "Twister" was usually the favourite, the other games rated very highly. Beyond this, there were numerous requests from both parents and children as to where they could purchase these so-called games, having failed in local toy stores.

Order

It is possible that the order in which the tests were done could affect the results. Thus, to do the convergent battery first may affect the divergent battery; and similarly if the divergent battery was administered first it could affect the convergent tests. Such an effect was not predicted, but in order to control for it the order of giving the tests was varied within the same set of conditions by applying any set of conditions to two schools.

Design

Eight schools were used to effect the desired administrative conditions and to control for order. Four sets of conditions were applied:

1. Test conditions for both convergent and divergent tests;
2. Game conditions for both convergent and divergent tests;
3. Test conditions for the convergent tests and game conditions for the divergent tests.

4. Game conditions for the convergent tests and test conditions for the divergent tests.

Within any one of these conditions the order of taking either the convergent or divergent tests first varied. Thus, the following experimental design was used:

School 1 - Convergent tests; Divergent tests;
School 2 - Divergent tests; Convergent tests;
School 3 - Convergent games; Divergent games;
School 4 - Divergent games; Convergent games;
School 5 - Convergent tests; Divergent games;
School 6 - Divergent games; Convergent tests;
School 7 - Convergent games; Divergent tests;
School 8 - Divergent tests; Convergent games.

There was a two-day rest period between the convergent and divergent "tests" for all schools. This was particularly important in the case of the last four schools who had a mixture of conditions. On finishing one battery the children in these schools were told that the investigator was returning in two days to give them either tests or games, whichever was appropriate. In the case of tests, the change in atmosphere was fairly easily accomplished; by emphasizing the conditions on returning two days later. In the case of the games, the transition was accomplished by the introductory informal 1-1/2 hour play period which occupied the first session when the games to be given were either convergent or divergent.

Pilot Study

Two hundred and four children from five schools were initially investigated in order to test the methodology of the research. For the most part little had to be changed from the initial design. The number of divergent stimuli were reduced to three from five in the case of the verbal tests and to four from five in the case of the non-verbal tests since it was felt that the original number of stimuli made the "testing" session too long. One test, "Circles" (Torrance, 1962), was added to the battery in one trial, but was abandoned because it proved to be a less reliable score than those on the other four "tests." Unobtrusive timing for the divergent tests was also attempted in one trial, the children being asked to put a mark on their sheets at certain time points. This, in the opinion of the children involved, was distracting. One time check could have been applied, but it was finally decided to have no time limit for reasons already given.

Two of the pilot schools had a large number of migrants as students. The application of the game atmosphere to a class where 80 percent were of mixed ethnic background proved to be of very limited success; particularly since the children found the divergent material difficult to respond to. Under test conditions control was easier to exercise but the migrant children still found the open-ended "tests" difficult to handle. The experience confirmed the decision made by the investigator not to include such migrant areas in the main study. A different administrative approach, and different "test" stimuli would appear to be more applicable to those children from a lower SES background and with problems in the English language, which many of these children had.

Data Analysis

The general approach was to use parametric methods of analysis, with a particular reliance on the univariate method of the analysis of variance (ANOVA). In every analysis the level of significance chosen was .02, i.e., referral to a significant result means at or beyond the level of .02. The major issue in this analysis was the adjustment of scores to eliminate the effects of age, sex, and the association with the IQ (as measured by the Otis). Although an attempt was made to control for these variables in the experimental design, they were found to affect at least some of the results. Thus, scores were adjusted to take into account the variance contributed by these parameters. Once this was achieved, univariate analyses were carried out on an average convergent score and an average divergent score computed from these two types of tests.

Further univariate analyses were then carried out for each convergent and divergent variable using raw or standard scores. To have carried out multivariate analyses might have yielded further results but these would have been beyond the questions raised by the research. Indeed, restricting the ANOVA to a univariate analysis does not disguise any major issues and the questions that would be raised by doing a multivariate analysis were not absolutely relevant to the thesis. Such an analysis would have revealed whether a relationship still existed between convergent and divergent scores after adjustment or whether they were independent. This issue was later examined through inter-correlation procedures. Furthermore, zero order correlations between the two dimensions could not be assumed, neither, as shown in the next chapter, are such correlations found under all conditions. Primarily, however, univariate analyses were preferred since it was felt to yield greater classification of any effects without intervening inter-relationships which could have masked the more important results to be sought. A balance has to be reached between extracting from the data a minimum number of results which are by their statistical "purity" meaningful, and the maximum number of results obtained from "milking" the data by the use of additional statistical methods. One may often lose psychological meaningfulness, as a consequence, because of the possible contamination of effects which may result. Thus, in view of the great number of variables already involved in this research, the major emphasis was placed on adjusting and normalizing the data so that two comparable scores could be used. In addition, a univariate ANOVA was felt to be a more pragmatic statistical tool to use.

In computing average convergent and divergent scores weighted means were not used since there was no evidence to suggest that any one variable contributed more to the overall result than another. Adjustment of scores for age, sex and Otis IQ was done systematically. Details of this procedure are given in the next chapter. Here, however, it is appropriate to explain the basic principles which guided these analyses. To get an average convergent or an average divergent score one needs to standardize. But in order to standardize such scores, a normal distribution is required to attain normal standard deviation and variance. Once the requirements of a normal distribution are met, scores can be standardized and then an average score can be computed. Examination of the raw scores for each convergent test revealed that the criteria of a normal distribution were met so that standardization of these scores was justifiable and a mean convergent score was computed. This was not the case for the raw scores on each of the 28 variables of the divergent tests and an intervening procedure had to be made. When the distribution of the divergent data was examined it was found to conform to a log distribution which justified converting these scores to log transform values. This statistical manipulation allowed the divergent data to meet the criteria of a normal distribution so that standardization of the scores could then proceed and a mean divergent value computed. In applying correction factors for age, sex and Otis IQ, the details of which are in the next chapter, certain rationale, such as condition or order were used to compute the correction value for each convergent and divergent variable. Correction factors were not applied to each school. If the investigator applies a factor over a defined population, such as each particular school, then theoretically one could confirm any hypothesis. Such potential confounding of the results was therefore avoided. When a correction was made, the distribution of the values of the corrected variable was plotted in order to ensure that the procedure had not de-normalized the data. Since, in fact, this procedure improved the normality of the distribution of each variable, standardization of the scores was justified. The reason for computing corrected scores for Otis, age and sex is to prevent the results from becoming contaminated with these independent variables of the study. Only in this way can one determine whether the results are due to the effects of the dependent variables, such as administrative conditions or order of the tests, and are not due to some other variable.

Due to the magnitude of this study and to the fact that sex differences in scores were found (and thus had to be corrected for in the combined sex sample), separate analysis of the data by sex was not done in relation to the independent variables. It is intended to make this topic the subject of a future study since sex differences in convergent and divergent thinking performances under different administrative conditions would appear to be a challenging and important research problem.

A presentation of the main results will now be made in the next chapter.

CHAPTER V

RESULTS

Main Effects

A. Average Convergent Score

It was found that there was a significant main effect of condition on the average convergent score. This means that the test atmosphere affected the level of performance on the convergent tasks. This is shown in Table 30. No other main effects are significant.

A further analysis of the convergent condition effect revealed that those schools who took the convergent battery under _test_ conditions did better than those who did them under _play_ conditions. The standardized grand mean was 0.04 with schools under test conditions deviating from the mean by +0.16 and those under game conditions deviating by -0.17. The conditions under which the divergent tests are taken and the order in which the convergent "tests" are administered do _not_ exert any main effects on the results.

Table 30

Results of Analysis of Variance
for Average Convergent Score

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	14.907	3	4.969	11.235	0.001
CONVTP	14.228	1	14.228	32.171	0.001
DIVRTP	0.355	1	0.355	0.803	0.999
ORDER	0.344	1	0.344	0.779	0.999
2-way Interactions	10.660	3	3.553	8.034	0.001
CONVTP DIVRTP	7.210	1	7.210	16.302	0.001
CONVTP ORDER	0.543	1	0.543	1.228	0.267
DIVRTP ORDER	3.462	1	3.462	7.827	0.006
3-way Interactions	8.853	1	8.853	20.017	0.001
CONVTP DIVRTP ORDER	8.853	1	8.853	20.017	0.001
Explained	34.420	7	4.917	11.118	0.001
Residual	214.503	485	0.442		
Total	248.922	492	0.506		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Thus, Hypothesis 1, which states that the order of giving the "tests" will not effect the level of performance on the convergent tasks is confirmed. Hypothesis 5, which states that schools subjected to test conditions for the convergent tasks will perform at a higher level than schools subjected to game conditions, is also confirmed. The average convergent score was obtained by calculating a mean from the sub-scores of the four convergent tests. Such a score had to be corrected for differences in age and Otis IQ scores since these two variables differed between the schools. Scores were standardized for each sex separately since there was a difference in scores between males and females. This is further described in Other Effects.

B. Average Divergent Score

In the case of the average divergent score there are two main effects. These are the conditions under which the divergent tests are administered and the order in which they appear in the test battery. These results are shown in Table 31.

Further analysis of the condition effect revealed that those children who were given the divergent battery under test conditions did significantly better than those who were administered the "tests" under play conditions. The standardized grand mean was -0.01. Schools under test conditions

Table 31

Results of Analysis of Variance
for Average Divergent Score

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	50.811	3	16.937	22.870	0.001
CONVTP	1.308	1	1.308	1.766	0.181
DIVRTP	27.719	1	27.719	37.429	0.001
ORDER	24.045	1	24.045	32.468	0.001
2-way Interactions	19.969	3	6.656	8.988	0.001
CONVTP DIVRTP	1.615	1	1.615	2.181	0.136
CONVTP ORDER	18.519	1	18.519	25.006	0.001
DIVRTP ORDER	0.286	1	0.286	0.386	0.999
3-way Interactions	0.703	1	0.703	0.949	0.999
CONVTP DIVRTP ORDER	0.703	1	0.703	0.949	0.999
Explained	71.484	7	10.212	13.789	0.001
Residual	359.179	485	0.741		
Total	430.663	492	0.875		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

deviated from the mean by +0.23; those under game conditions deviated by -0.22. This result means that Hypothesis 6, which states that schools subjected to game conditions for the divergent tasks will perform at a higher level than schools subjected to the test conditions, is rejected. Higher divergent scores are achieved under formal conditions.

Analysis of the order effect shows that schools which had the divergent tasks first did significantly better than those who did the convergent tasks first. Schools which had the divergent "tests" first deviated from the grand mean by +0.22, while those who had the convergent "tests" first deviated from the mean by -0.20. This result means that Hypothesis 2 is rejected. Hypothesis 2 stated that order of giving the "tests" will not effect the level of performance on the divergent tasks. Order did have an effect on these variables in this research. Doing the divergent battery first enhanced performance levels.

The average divergent score was obtained from a mean of the 28 divergent variables. For each divergent test there were seven variables, and all variables from the four tests were used in compiling this score. Since each of the sub-scores on each variable had such a wide range, these were standardized into log scores. Since there was a difference in scores between males and females, scores were then standardized for each sex separately. Finally, since the Otis IQ scores also affected the divergent results, the average divergent score was corrected for this. No age correction was made. This is further described in Other Effects.

C. <u>The Convergent Sub-Scores: Similarities, Coloured Progressive Matrices, Picture Completion and Crichton Vocabulary Scale</u>

Separate analyses of variance were done for both sexes on each convergent variable. Because there was a significant difference between the schools on the Otis test an analysis of covariance was also done. On all four variables the Otis IQ score had a statistically significant effect on the level of performance. When the Otis was held as a covariate the main effects of condition and order seen in the analysis of variance were sometimes altered.

The analysis of variance table for Similarities indicated a main effect of order. This is shown in Table 32. Further analysis of this result revealed that those schools who had divergent "tests" preceding this convergent task had the higher performance level.

However, this order effect was not apparent when the Otis was held as a covariate. No significant main effects with respect to condition or order are found. This is shown in Table 33.

Table 32

Results of Analysis of Variance for Similarities

Source Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	8.991	3	2.997	3.077	0.027
ORDER	8.799	1	8.799	9.033	0.003
CONVTP	0.006	1	0.006	0.006	0.999
DIVRTP	0.074	1	0.074	0.076	0.999
2-way Interactions	10.435	3	3.478	3.571	0.014
ORDER CONVTP	7.841	1	7.841	8.051	0.005
ORDER DIVRTP	2.772	1	2.772	2.846	0.088
CONVTP DIVRTP	0.050	1	0.050	0.051	0.999
3-way Interactions	0.169	1	0.169	0.173	0.999
ORDER CONVTP DIVRTP	0.169	1	0.169	0.173	0.999
Explained	19.595	7	2.799	2.874	0.006
Residual	472.395	485	0.974		
Total	491.990	492	1.000		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 33

Results of Analysis of Covariance for Similarities

Source of Variation			Sum of Squares	DF	Mean Square	F	Significance of F
Covariates			200.179	1	200.179	346.712	0.001
OT			200.179	1	200.179	346.712	0.001
Main Effects			2.444	3	0.815	1.411	0.238
ORDER			0.594	1	0.594	1.028	0.312
CONVTP			1.353	1	1.353	2.343	0.122
DIVRTP			0.377	1	0.377	0.654	0.999
2-way Interactions			9.387	3	3.129	5.420	0.001
ORDER	CONVTP		0.542	1	0.542	0.939	0.999
ORDER	DIVRTP		5.511	1	5.511	9.544	0.003
CONVTP	DIVRTP		3.810	1	3.810	6.598	0.010
3-way Interactions			0.534	1	0.534	0.925	0.999
ORDER	CONVTP	DIVRTP	0.534	1	0.534	0.925	0.999
Explained			212.545	8	26.568	46.016	0.001
Residual			279.445	484	0.577		
Total			491.990	492	1.000		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

In the case of the Coloured Progressive Matrices (CPM) there was a significant main effect of condition; i.e., the type of condition affected the level of performance on the CPM. This result is shown in Table 34.

Although the Otis score was shown to significantly effect the level of performance on the CPM, the analysis of covariance did not alter this experimental main effect. Further analysis of this main effect of condition on the CPM showed that those groups of children who had been administered this task under test conditions did significantly better than those who had had the task administered under play conditions. The normalized standard mean was 50. Those under test conditions deviated from this mean by +0.12 and those under play conditions by -0.12.

The analysis of variance table for the Picture Completion Test shown no main effects. This is shown in Table 35.

However, the analysis of covariance showed a significant order effect. This is shown in Table 36.

Further analysis of this main effect of order revealed that those groups of children who did their convergent tasks first did better than those groups who had the divergent "tests" prior to the convergent "tests." Those who had the convergent battery first deviated from the standard grand mean by +0.02 (a value which increased to 0.09 when independents and covariates were adjusted for). Those that had

Table 34

Results of Analysis of Variance for CPM

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	11.933	3	3.978	4.104	0.007
ORDER	4.676	1	4.676	4.825	0.027
CONVTP	6.777	1	6.777	6.993	0.008
DIVRTP	0.085	1	0.085	0.087	0.999
2-way Interactions	6.213	3	2.071	2.137	0.093
ORDER CONVTP	1.560	1	1.560	1.609	0.202
ORDER DIVRTP	2.231	1	2.231	2.301	0.126
CONVTP DIVRTP	2.118	1	2.118	2.185	0.136
3-way Interactions	3.801	1	3.801	3.922	0.045
ORDER CONVTP DIVRTP	3.801	1	3.801	3.922	0.045
Explained	21.948	7	3.135	3.235	0.003
Residual	470.056	485	0.969		
Total	492.003	492	1.000		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 35

Results of Analysis of Variance for Picture Completion

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	2.509	3	0.836	0.854	0.999
ORDER	0.199	1	0.199	0.203	0.999
CONVTP	0.871	1	0.871	0.889	0.999
DIVRTP	1.374	1	1.374	1.403	0.235
2-way Interactions	6.290	3	2.097	2.140	0.093
ORDER CONVTP	3.362	1	3.362	3.341	0.061
ORDER DIVRTP	0.125	1	0.125	0.128	0.999
CONVTP DIVRTP	3.232	1	3.232	3.299	0.066
3-way Interactions	7.963	1	7.963	8.129	0.005
ORDER CONVTP DIVRTP	7.963	1	7.963	8.129	0.005
Explained	16.762	7	2.395	2.444	0.018
Residual	475.124	485	0.980		
Total	491.885	492	1.000		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 36

Results of Analysis of Covariance for Picture Completion

Source of Variation			Sum of Squares	DF	Mean Square	F	Significance of F
Covariates			104.849	1	104.849	140.131	0.001
OT			104.849	1	104.849	140.131	0.001
Main Effects			5.294	3	1.765	2.358	0.070
ORDER			4.207	1	4.207	5.623	0.017
CONVTP			0.000	1	0.000	0.000	0.999
DIVRTP			0.839	1	0.839	1.122	0.290
2-way Interactions			10.204	3	3.401	4.546	0.004
ORDER	CONVTP		0.071	1	0.071	0.094	0.999
ORDER	DIVRTP		0.764	1	0.764	1.021	0.314
CONVTP	DIVRTP		9.613	1	9.613	12.848	0.001
3-way Interactions			9.402	1	9.402	12.566	0.001
ORDER	CONVTP	DIVRTP	9.402	1	9.402	12.566	0.001
Explained			129.749	8	16.219		
Residual			362.137	484	0.748		
Total			491.885	492	1.000		

Key:
CONVTP = Conditions of administration of convergent tests
DIVRTP = Conditions of administration of divergent tests
ORDER = Order of test administration

the divergent battery first had a deviation score of -0.03 (which became -0.10 when the above adjustments were made).

Conditions exerted a significant main effect on the level of performance on the Crichton Vocabulary scale; i.e., the type of administrative regime used for this task significantly affected the scores. This result from the analysis of variance is shown in Table 37.

It should be noted from this table that the conditions under which the divergent tests were done exert a near significant effect on the scores of the vocabulary test. When the results of the analysis of covariance are examined this latter effect is significant; i.e., the test atmosphere in which the divergent tasks are administered exerts an influence on the performance level of this convergent test. In addition, the conditions under which the vocabulary test itself is administered remains a significant main effect when the Otis is held as a covariate. These results are presented in Table 38.

Further analysis of these covariance results showed that with respect to the administrative atmosphere for the vocabulary "test," those children who did it under test conditions did significantly better than those who did it under game conditions. The unadjusted deviation scores from the standard grand mean of 50 were +.10 for the formal conditions and -.011 for the informal conditions.

Table 37

Results of Analysis of Variance for
the Crichton Vocabulary Scale

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	11.861	3	3.954	4.151	0.007
ORDER	1.620	1	1.620	1.701	0.190
CONVTP	5.349	1	5.349	5.617	0.017
DIVRTP	4.905	1	4.905	5.150	0.022
2-way Interactions	11.742	3	3.914	4.110	0.007
ORDER CONVTP	6.285	1	6.285	6.600	0.010
ORDER DIVRTP	0.013	1	0.013	0.013	0.999
CONVTP DIVRTP	6.104	1	6.104	6.410	0.011
3-way Interactions	6.234	1	6.234	6.546	0.011
ORDER CONVTP DIVRTP	6.234	1	6.234	6.546	0.011
Explained	29.837	7	4.262	4.475	0.001
Residual	461.917	485	0.952		
Total	491.754	492	0.999		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of administration

Table 38

Results of Analysis of Covariance for the Crichton Vocabulary Scale

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Covariates	227.379	1	227.379	503.528	0.001
OT	227.379	1	227.379	503.528	0.001
Main Effects	18.158	3	6.053	13.404	0.001
ORDER	1.244	1	1.244	2.755	0.094
CONVTP	13.429	1	13.429	29.737	0.001
DIVRTP	3.376	1	3.376	7.476	0.007
2-way Interactions	19.500	3	6.500	14.394	0.001
ORDER CONVTP	0.041	1	0.041	0.091	0.999
ORDER DIVRTP	0.427	1	0.427	0.945	0.999
CONVTP DIVRTP	19.262	1	19.262	42.655	0.001
3-way Interactions	8.156	1	8.156	18.060	0.001
ORDER CONVTP DIVRTP	8.156	1	8.156	18.060	0.001
Explained	273.193	8	34.149	75.623	0.001
Residual	218.561	484	0.452		
Total	491.754	492	0.999		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

When adjustments were made for independents and the co-variate the values became +0.16 and -0.17, respectively. With respect to the type of conditions under which the divergent tasks were taken and their influence on vocabulary scores, the analysis indicated that those schools who did the divergent battery under test conditions had better vocabulary scores. The unadjusted deviation score for the divergent formal atmosphere was +0.10 compared to -0.10 for the divergent informal atmosphere. When these scores were adjusted they became +0.08 and -0.08, respectively.

As a summary of this subsection it can be stated that when the convergent tests are looked at separately a variety of test and order effects appear. On the CPM and the vocabulary tests formal conditions produce a higher level of performance. Order exerts a significant event in the case of the Picture Completion test, namely the administration of convergent tasks first. The conditions under which the divergent tests are administered, namely test conditions produces an additional main effect on vocabulary scores. Order effects, namely the presentation of the divergent battery first were found to influence performance levels on Similarities but most of this effect was related to the level of Otis IQ scores since it was eliminated under an analysis of covariance. In terms of the hypotheses that were predicted, Hypothesis 5, which relates to the efficacy of test conditions applies to two of the convergent tests. Hypothesis 2 which states that order will not effect the level of performance on convergent tests remains confirmed with one exception. In the case of Picture Completion, the hypothesis is rejected.

D. The Divergent Sub-scores for Just Suppose, Patterns, Uses of Objects and Inkblots

Separate analyses of variance were also done for both sexes on each of the 28 divergent variables. In addition, analyses of covariance were performed. Although the Otis IQ test scores did have a statistically significant effect on the level of performance on every variable, there was no case where the main effects which were demonstrated on an analysis of variance were changed by using the Otis as a covariate. The main effects apparent over most of the divergent variables were those conditions under which the divergent battery was taken and the effect of order. The two exceptions are the parameters of Inkblots, Originality 1, and Just Suppose, Originality 1 where there is a main effect of order only. These results are shown in Tables 39-66.

The main effects of conditions under which the divergent "tests" are taken, and the effect of order are remarkably consistent over the 28 variables. There were only two exceptions, as previously mentioned, where order only was

Table 39

Results of Analysis of Variance for
Uses of Objects, Fluency 1

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	3.364	3	1.121	20.174	0.001
ORDER	2.011	1	2.011	36.173	0.001
CONVTP	0.000	1	0.000	0.008	0.999
DIVRTP	1.539	1	1.539	27.690	0.001
2-way Interactions	2.207	3	0.736	13.237	0.001
ORDER CONVTP	1.317	1	1.317	23.690	0.001
ORDER DIVRTP	0.491	1	0.491	8.837	0.003
CONVTP DIVRTP	0.424	1	0.424	7.632	0.006
3-way Interactions	0.000	1	0.000	0.006	0.999
ORDER CONVTP DIVRTP	0.000	1	0.000	0.006	0.999
Explained	5.571	7	0.796	14.320	0.001
Residual	26.957	485	0.056		
Total	32.528	492	0.066		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 40

Results of Analysis of Variance for
Uses of Objects, Fluency 2

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	3.373	3	1.124	20.017	0.001
ORDER	1.984	1	1.984	35.311	0.001
CONVTP	0.000	1	0.000	0.006	0.999
DIVRTP	1.576	1	1.576	28.063	0.001
2-way Interactions	2.224	3	0.741	13.197	0.001
ORDER CONVTP	1.321	1	1.321	23.508	0.001
ORDER DIVRTP	0.486	1	0.486	8.657	0.004
CONVTP DIVRTP	0.444	1	0.444	7.899	0.005
3-way Interactions	0.001	1	0.001	0.019	0.999
ORDER CONVTP DIVRTP	0.001	1	0.001	0.019	0.999
Explained	5.599	7	0.800	14.237	0.001
Residual	27.245	485	0.056		
Total	32.844	492	0.067		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 41

Results of Analysis of Variance for
Uses of Objects, Flexibility 1

Source of Variation			Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects			3.930	3	1.310	23.885	0.001
ORDER			2.484	1	2.484	45.290	0.001
CONVTP			0.023	1	0.023	0.419	0.999
DIVRTP			1.615	1	1.615	29.438	0.001
2-way Interactions			2.532	3	0.844	15.386	0.001
ORDER	CONVTP		1.956	1	1.956	35.666	0.001
ORDER	DIVRTP		0.023	1	0.023	0.420	0.999
CONVTP	DIVRTP		0.663	1	0.663	12.083	0.001
3-way Interactions			0.008	1	0.008	0.141	0.999
ORDER	CONVTP	DIVRTP	0.008	1	0.008	0.141	0.999
Explained			6.470	7	0.924	16.851	0.001
Residual			26.603	485	0.055		
Total			33.073	492	0.067		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 42

Results of Analysis of Variance for
Uses of Objects, Flexibility 2

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	3.409	3	1.136	21.782	0.001
ORDER	1.952	1	1.952	37.412	0.001
CONVTP	0.019	1	0.019	0.356	0.999
DIVRTP	1.610	1	1.610	30.858	0.001
2-way Interactions	2.122	3	0.707	13.556	0.001
ORDER CONVTP	1.511	1	1.511	28.965	0.001
ORDER DIVRTP	0.179	1	0.179	3.437	0.061
CONVTP DIVRTP	0.489	1	0.489	9.370	0.003
3-way Interactions	0.015	1	0.015	0.285	0.999
ORDER CONVTP DIVRTP	0.015	1	0.015	0.285	0.999
Explained	5.546	7	0.792	15.185	0.001
Residual	25.303	485	0.052		
Total	30.849	492	0.063		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 43

Results of Analysis of Variance for
Uses of Objects, Originality 1

Source of Variation	Sum of Squares	DF	Mean Squares	F	Significance of F
Main Effects	8.459	3	2.820	18.849	0.001
ORDER	4.946	1	4.946	33.060	0.001
CONVTP	0.233	1	0.233	1.559	0.210
DIVRTP	3.644	1	3.644	24.359	0.001
2-way Interactions	5.593	3	1.864	12.462	0.001
ORDER CONVTP	4.234	1	4.234	28.303	0.001
ORDER DIVRTP	0.433	1	0.433	2.894	0.086
CONVTP DIVRTP	1.068	1	1.068	7.138	0.008
3-way Interactions	0.153	1	0.153	1.023	0.313
ORDER CONVTP DIVRTP	0.153	1	0.153	1.023	0.313
Explained	14.205	7	2.029	13.565	0.001
Residual	72.556	485	0.150		
Total	86.761	492	0.176		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 44

Results of Analysis of Variance for
Uses of Objects, Originality 2

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	8.768	3	2.923	21.102	0.001
ORDER	5.858	1	5.858	42.292	0.001
CONVTP	0.012	1	0.012	0.089	0.999
DIVRTP	3.340	1	3.340	24.114	0.001
2-way Interactions	5.810	3	1.937	13.982	0.001
ORDER CONVTP	4.872	1	4.872	35.176	0.001
ORDER DIVRTP	0.065	1	0.065	0.467	0.999
CONVTP DIVRTP	1.088	1	1.088	7.854	0.005
3-way Interactions	0.001	1	0.001	0.004	0.999
ORDER CONVTP DIVRTP	0.001	1	0.001	0.004	0.999
Explained	14.579	7	2.083	15.037	0.001
Residual	67.176	485	0.139		
Total	81.755	492	0.166		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of administration

Table 45

Results of Analysis of Variance for
Uses of Objects, Originality 3

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	13.135	3	4.378	23.755	0.001
ORDER	7.585	1	7.585	41.154	0.001
CONVTP	0.002	1	0.002	0.013	0.999
DIVRTP	6.258	1	6.258	33.952	0.001
2-way Interactions	6.602	3	2.201	11.939	0.001
ORDER CONVTP	5.761	1	5.761	31.256	0.001
ORDER DIVRTP	0.096	1	0.096	0.520	0.999
CONVTP DIVRTP	0.953	1	0.953	5.171	0.022
3-way Interactions	0.002	1	0.002	0.013	0.999
ORDER CONVTP DIVRTP	0.002	1	0.002	0.013	0.999
Explained	19.740	7	2.820	15.299	0.001
Residual	89.395	485	0.184		
Total	109.134	492	0.222		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 46

Results of Analysis of Variance for
Patterns, Fluency 1

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	1.610	3	0.537	11.152	0.001
ORDER	0.524	1	0.524	10.881	0.001
CONVTP	0.047	1	0.047	0.986	0.999
DIVRTP	1.105	1	1.105	22.955	0.001
2-way Interactions	1.385	3	0.462	9.588	0.001
ORDER CONVTP	1.275	1	1.275	26.493	0.001
ORDER DIVRTP	0.092	1	0.092	1.907	0.164
CONVTP DIVRTP	0.000	1	0.000	0.000	0.999
3-way Interactions	0.004	1	0.004	0.082	0.999
ORDER CONVTP DIVRTP	0.004	1	0.004	0.082	0.999
Explained	2.999	7	0.028	8.900	0.001
Residual	23.346	485	0.048		
Total	26.345	492	0.054		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 47

Results of Analysis of Variance for Patterns, Fluency 2

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	1.595	3	0.532	10.915	0.001
ORDER	0.525	1	0.525	10.777	0.001
CONVTP	0.045	1	0.045	0.923	0.999
DIVRTP	1.091	1	1.091	22.392	0.001
2-way Interactions	1.429	3	0.476	9.779	0.001
ORDER CONVTP	1.318	1	1.318	27.053	0.001
ORDER DIVRTP	0.093	1	0.093	1.899	0.165
CONVTP DIVRTP	0.000	1	0.000	0.001	0.999
3-way Interactions	0.007	1	0.007	0.135	0.999
ORDER CONVTP DIVRTP	0.007	1	0.007	0.135	0.999
Explained	3.030	7	0.433	8.888	0.001
Residual	23.620	485	0.049		
Total	26.650	492	0.054		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 48

Results of Analysis of Variance for Patterns, Flexibility 1

Source of Variation			Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects			1.723	3	0.574	14.081	0.001
ORDER			0.811	1	0.811	19.891	0.001
CONVTP			0.027	1	0.027	0.656	0.999
DIVRTP			0.966	1	0.966	23.683	0.001
2-way Interactions			0.931	3	0.310	7.603	0.001
ORDER	CONVTP		0.833	1	0.833	20.425	0.001
ORDER	DIVRTP		0.056	1	0.056	1.385	0.238
CONVTP	DIVRTP		0.017	1	0.017	0.423	0.999
3-way Interactions			0.029	1	0.029	0.723	0.999
ORDER	CONVTP	DIVRTP	0.029	1	0.029	0.723	0.999
Explained			2.683	7	0.383	9.397	0.001
Residual			19.786	485	0.041		
Total			22.469	492	0.046		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 49

Results of Analysis of Variance for Patterns, Flexibility 2

Source of Variation			Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects			1.643	3	0.548	12.702	0.001
ORDER			0.684	1	0.684	15.857	0.001
CONVTP			0.014	1	0.014	0.324	0.999
DIVRTP			1.025	1	1.025	23.771	0.001
2-way Interactions			1.122	3	0.374	8.671	0.001
ORDER	CONVTP		1.018	1	1.018	23.622	0.001
ORDER	DIVRTP		0.083	1	0.083	1.936	0.161
CONVTP	DIVRTP		0.001	1	0.001	0.033	0.999
3-way Interactions			0.054	1	0.054	1.249	0.263
ORDER	CONVTP	DIVRTP	0.054	1	0.054	1.249	0.263
Explained			2.818	7	0.403	9.338	0.001
Residual			20.911	485	0.043		
Total			23.729	492	0.048		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of administration

Table 50

Results of Analysis of Variance for Patterns, Originality 1

Source of Variation			Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects			6.837	3	2.279	15.337	0.001
ORDER			4.864	1	4.864	32.736	0.001
CONVTP			0.001	1	0.001	0.010	0.999
DIVRTP			2.327	1	2.327	15.661	0.001
2-way Interactions			3.601	3	1.200	8.078	0.001
ORDER	CONVTP		3.442	1	3.442	23.160	0.001
ORDER	DIVRTP		0.119	1	0.119	0.802	0.999
CONVTP	DIVRTP		0.000	1	0.000	0.002	0.999
3-way Interactions			0.059	1	0.059	0.395	0.999
ORDER	CONVTP	DIVRTP	0.059	1	0.059	0.395	0.999
Explained			10.497	7	1.500	10.092	0.001
Residual			72.071	485	0.149		
Total			82.568	492	0.168		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 51

Results of Analysis of Variance for
Patterns, Originality 2

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	5.563	3	1.854	16.288	0.001
ORDER	3.195	1	3.195	28.066	0.001
CONVTP	0.095	1	0.095	0.830	0.999
DIVRTP	2.529	1	2.529	22.217	0.001
2-way Interactions	2.228	3	0.743	6.524	0.001
ORDER CONVTP	2.127	1	2.127	18.681	0.001
ORDER DIVRTP	0.076	1	0.076	0.669	0.999
CONVTP DIVRTP	0.029	1	0.029	0.257	0.999
3-way Interactions	0.003	1	0.003	0.031	0.999
ORDER CONVTP DIVRTP	0.003	1	0.003	0.031	0.999
Explained	7.794	7	1.113	9.781	0.001
Residual	55.214	485	0.114		
Total	63.008	492	0.128		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 52

Results of Analysis of Variance for
Patterns, Originality 3

Source of Variation			Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects			3.992	3	1.331	12.168	0.001
ORDER			2.081	1	2.081	19.025	0.001
CONVTP			0.028	1	0.028	0.253	0.999
DIVRTP			2.083	1	2.083	19.049	0.001
2-way Interactions			2.270	3	0.757	6.919	0.001
ORDER	CONVTP		2.111	1	2.111	19.303	0.001
ORDER	DIVRTP		0.131	1	0.131	1.202	0.273
CONVTP	DIVRTP		0.000	1	0.000	0.000	0.999
3-way Interactions			0.009	1	0.009	0.085	0.999
ORDER	CONVTP	DIVRTP	0.009	1	0.009	0.085	0.999
Explained			6.272	7	0.896	8.192	0.001
Residual			53.044	485	0.109		
Total			59.316	492	0.121		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 53

Results of Analysis of Variance
for Inkblots, Fluency 1

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	0.999	3	0.333	6.440	0.001
ORDER	0.605	1	0.605	11.706	0.001
CONVTP	0.075	1	0.075	1.448	0.227
DIVRTP	0.351	1	0.351	6.795	0.009
2-way Interactions	0.660	3	0.220	4.257	0.006
ORDER CONVTP	0.573	1	0.573	11.078	0.001
ORDER DIVRTP	0.008	1	0.008	0.148	0.999
CONVTP DIVRTP	0.057	1	0.057	1.098	0.295
3-way Interactions	0.026	1	0.026	0.498	0.999
ORDER CONVTP DIVRTP	0.026	1	0.026	0.498	0.999
Explained	1.685	7	0.241	4.655	0.001
Residual	25.082	485	0.052		
Total	26.767	492	0.054		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 54

Results of Analysis of Variance
for Inkblots, Fluency 2

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	1.008	3	0.336	6.483	0.001
ORDER	0.609	1	0.609	11.749	0.001
CONVTP	0.071	1	0.071	1.368	0.241
DIVRTP	0.362	1	0.362	6.986	0.008
2-way Interactions	0.639	3	0.213	4.112	0.007
ORDER CONVTP	0.553	1	0.553	10.673	0.002
ORDER DIVRTP	0.007	1	0.007	0.136	0.999
CONVTP DIVRTP	0.056	1	0.056	1.089	0.297
3-way Interactions	0.029	1	0.029	0.560	0.999
ORDER CONVTP DIVRTP	0.029	1	0.029	0.560	0.999
Explained	1.676	7	0.239	4.621	0.001
Residual	25.137	485	0.052		
Total	26.814	492	0.054		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 55

Results of Analysis of Variance
for Inkblots, Flexibility 1

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	0.929	3	0.310	8.127	0.001
ORDER	0.621	1	0.621	16.296	0.001
CONVTP	0.071	1	0.071	1.875	0.168
DIVRTP	0.264	1	0.264	6.923	0.009
2-way Interactions	0.681	3	0.227	5.956	0.001
ORDER CONVTP	0.595	1	0.595	15.625	0.001
ORDER DIVRTP	0.025	1	0.025	0.645	0.999
CONVTP DIVRTP	0.042	1	0.042	1.095	0.296
3-way Interactions	0.054	1	0.054	1.417	0.233
ORDER CONVTP DIVRTP	0.054	1	0.054	1.417	0.233
Explained	1.664	7	0.238	6.238	0.001
Residual	18.484	485	0.038		
Total	20.148	492	0.041		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 56

Results of Analysis of Variance
for Inkblots, Flexibility 2

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	1.020	3	0.340	8.316	0.001
ORDER	0.684	1	0.684	16.738	0.001
CONVTP	0.063	1	0.063	1.550	0.211
DIVRTP	0.305	1	0.305	7.461	0.007
2-way Interactions	0.718	3	0.239	5.850	0.001
ORDER CONVTP	0.604	1	0.604	14.772	0.001
ORDER DIVRTP	0.010	1	0.010	0.238	0.999
CONVTP DIVRTP	0.076	1	0.076	1.869	0.169
3-way Interactions	0.051	1	0.051	1.247	0.264
ORDER CONVTP DIVRTP	0.051	1	0.051	1.247	0.264
Explained	1.789	7	0.256	6.249	0.001
Residual	19.831	485	0.041		
Total	21.620	492	0.044		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 57

Results of Analysis of Variance
for Inkblots, Originality 1

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	3.224	3	1.075	7.469	0.001
ORDER	2.745	1	2.745	19.083	0.001
CONVTP	0.011	1	0.011	0.078	0.999
DIVRTP	0.585	1	0.585	4.063	0.042
2-way Interactions	1.477	3	0.492	3.422	0.017
ORDER CONVTP	1.437	1	1.437	9.985	0.002
ORDER DIVRTP	0.010	1	0.010	0.070	0.999
CONVTP DIVRTP	0.008	1	0.008	0.056	0.999
3-way Interactions	0.236	1	0.236	1.642	0.198
ORDER CONVTP DIVRTP	0.236	1	0.236	1.642	0.198
Explained	4.937	7	0.705	4.902	0.001
Residual	69.774	485	0.144		
Total	74.711	492	0.152		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test conditions

Table 58

Results of Analysis of Variance
for Inkblots, Originality 2

Source of Variation			Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects			3.409	3	1.136	12.309	0.001
ORDER			2.782	1	2.782	30.134	0.001
CONVTP			0.040	1	0.040	0.435	0.999
DIVRTP			0.708	1	0.708	7.674	0.006
2-way Interactions			0.953	3	0.318	3.440	0.017
ORDER	CONVTP		0.762	1	0.762	8.252	0.005
ORDER	DIVRTP		0.029	1	0.029	0.313	0.999
CONVTP	DIVRTP		0.123	1	0.123	1.334	0.247
3-way Interactions			0.091	1	0.091	0.987	0.999
ORDER	CONVTP	DIVRTP	0.091	1	0.091	0.987	0.999
Explained			4.453	7	0.636	6.891	0.001
Residual			44.772	485	0.092		
Total			49.224	492	0.100		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 59

Results of Analysis of Variance for Inkblots, Originality 3

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	2.628	3	0.876	11.953	0.001
ORDER	1.747	1	1.747	23.836	0.001
CONVTP	0.337	1	0.337	4.601	0.030
DIVRTP	0.596	1	0.596	8.138	0.005
2-way Interactions	0.672	3	0.224	3.058	0.027
ORDER CONVTP	0.600	1	0.600	8.192	0.005
ORDER DIVRTP	0.000	1	0.000	0.000	0.999
CONVTP DIVRTP	0.049	1	0.049	0.666	0.999
3-way Interactions	0.005	1	0.005	0.066	0.999
ORDER CONVTP DIVRTP	0.005	1	0.005	0.066	0.999
Explained	3.306	7	0.472	6.443	0.001
Residual	35.551	485	0.073		
Total	38.857	492	0.079		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 60

Results of Analysis of Variance of
Just Suppose, Fluency 1

Source of Variation			Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects			2.055	3	0.685	14.731	0.001
ORDER			0.727	1	0.727	15.628	0.001
CONVTP			0.018	1	0.018	0.384	0.999
DIVRTP			1.406	1	1.406	30.245	0.001
2-way Interactions			0.259	3	0.086	1.853	0.135
ORDER	CONVTP		0.174	1	0.174	3.745	0.050
ORDER	DIVRTP		0.055	1	0.055	1.186	0.276
CONVTP	DIVRTP		0.020	1	0.020	0.428	0.999
3-way Interactions			0.022	1	0.022	0.470	0.999
ORDER	CONVTP	DIVRTP	0.022	1	0.022	0.470	0.999
Explained			2.335	7	0.334	7.175	0.001
Residual			22.553	485	0.047		
Total			24.888	492	0.051		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 61

Results of Analysis of Variance of
Just Suppose, Fluence 2

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	2.063	3	0.688	14.861	0.001
ORDER	0.760	1	0.760	16.415	0.001
CONVTP	0.017	1	0.017	0.361	0.999
DIVRTP	1.385	1	1.385	29.917	0.001
2-way Interactions	0.241	3	0.080	1.738	0.157
ORDER CONVTP	0.163	1	0.163	3.514	0.058
ORDER DIVRTP	0.053	1	0.053	1.138	0.286
CONVTP DIVRTP	0.017	1	0.017	0.377	0.999
3-way Interactions	0.034	1	0.034	0.726	0.999
ORDER CONVTP DIVRTIP	0.034	1	0.034	0.726	0.999
Explained	2.338	7	0.334	7.217	0.001
Residual	22.447	485	0.046		
Total	24.78	492	0.050		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of convergent tests

ORDER = Order of test administration

Table 62

Results of Analysis of Variance of
Just Suppose, Flexibility 1

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	1.310	3	0.437	11.823	0.001
ORDER	0.574	1	0.574	15.551	0.001
CONVTP	0.009	1	0.009	0.244	0.999
DIVRTP	0.791	1	0.791	21.428	0.001
2-way Interactions	0.199	3	0.066	1.794	0.146
ORDER CONVTP	0.195	1	0.195	5.268	0.021
ORDER DIVRTP	0.000	1	0.000	0.004	0.999
CONVTP DIVRTP	0.001	1	0.001	0.034	0.999
3-way Interactions	0.029	1	0.029	0.773	0.999
ORDER CONVTP DIVRTP	0.029	1	0.029	0.773	0.999
Explained	1.537	7	0.220	5.947	0.001
Residual	17.912	485	0.037		
Total	19.449	492	0.040		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 63

Results of Analysis of Variance of
Just Suppose, Flexibility 2

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	1.626	3	0.542	13.598	0.001
ORDER	0.606	1	0.606	15.202	0.001
CONVTP	0.005	1	0.005	0.135	0.999
DIVRTP	1.095	1	1.095	27.475	0.001
2-way Interactions	0.252	3	0.084	2.104	0.097
ORDER CONVTP	0.180	1	0.180	4.505	0.032
ORDER DIVRTP	0.041	1	0.041	1.037	0.310
CONVTP DIVRTP	0.021	1	0.021	0.533	0.999
3-way Interactions	0.019	1	0.019	0.468	0.999
ORDER CONVTP DIVRTP	0.019	1	0.019	0.468	0.999
Explained	1.897	7	0.271	6.797	0.001
Residual	19.335	485	0.040		
Total	21.232	492	0.043		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test conditions

Table 64

Results of Analysis of Variance of
Just Suppose, Originality 1

Source of Variation			Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects			2.439	3	0.813	6.942	0.001
ORDER			1.686	1	1.686	14.398	0.001
CONVTP			0.276	1	0.276	2.360	0.121
DIVRTP			0.527	1	0.527	4.498	0.032
2-way Interactions			0.266	3	0.089	0.756	0.999
ORDER	CONVTP		0.181	1	0.181	1.546	0.212
ORDER	DIVRTP		0.002	1	0.002	0.015	0.999
CONVTP	DIVRTP		0.096	1	0.096	0.820	0.999
3-way interactions			0.084	1	0.084	0.719	0.999
ORDER	CONVTP	DIVRTP	0.084	1	0.084	0.719	0.999
Explained			2.789	7	0.398	3.402	0.002
Residual			56.798	485	0.117		
Total			59.587	492	0.121		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 65

Results of Analysis of Variance of
Just Suppose, Originality 2

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	3.988	3	1.329	12.331	0.001
ORDER	2.632	1	2.632	24.412	0.001
CONVTP	0.208	1	0.208	1.933	0.161
DIVRTP	1.287	1	1.287	11.933	0.001
2-way Interactions	0.359	3	0.120	1.110	0.345
ORDER CONVTP	0.170	1	0.170	1.573	0.208
ORDER DIVRTP	0.035	1	0.035	0.322	0.999
CONVTP DIVRTP	0.182	1	0.182	1.687	0.192
3-way Interactions	0.000	1	0.000	0.002	0.999
ORDER CONVTP DIVRTP	0.000	1	0.000	0.002	0.999
Explained	4.348	7	0.621	5.761	0.001
Residual	52.290	485	0.108		
Total	56.637	492	0.115		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test administration

Table 66

Results of Analysis of Variance of
Just Suppose, Originality 3

Source of Variation			Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects			4.807	3	1.602	13.602	0.001
ORDER			2.314	1	2.314	19.645	0.001
CONVTP			0.018	1	0.018	0.149	0.999
DIVRTP			2.722	1	2.722	23.108	0.001
2-way Interactions			1.500	3	0.500	4.244	0.006
ORDER	CONVTP		1.321	1	1.321	11.213	0.001
ORDER	DIVRTP		0.198	1	0.198	1.678	0.193
CONVTP	DIVRTP		0.011	1	0.011	0.091	0.999
3-way Interactions			0.123	1	0.123	1.042	0.309
ORDER	CONVTP	DIVRTP	0.123	1	0.123	1.042	0.309
Explained			6.429	7	0.918	7.797	0.001
Residual			57.134	485	0.118		
Total			63.564	492	0.129		

Key:

CONVTP = Conditions of administration of convergent tests

DIVRTP = Conditions of administration of divergent tests

ORDER = Order of test conditions

the main effect. Further analysis of these analyses of variance tables produced even more consistent results over all 28 variables.

Those who were given the four divergent tests under formal conditions had higher performance levels on the fluency, flexibility and originality scores than those who did the "tests" under game conditions. The two exceptions were on Inkblots, Originality 1 and Just Suppose, Originality 1, where there was no significant difference between the two conditions.

The effect of order applied to all 28 variables in the same way. Those groups of children who had the divergent battery first did significantly better than those children who had the convergent battery preceding their divergent "tests."

Thus, Hypothesis 6, which states that schools subjected to game conditions will perform at a higher level than schools subjected to the test conditions, is rejected. Likewise, Hypothesis 2, which states that order of giving the "tests" will not effect the level of performance on the divergent tasks, is rejected.

Levels of Means and Scheffé Tests

A further attempt was made to examine the differences in means between the schools to determine which combination of effects produced the best level of performance. Each of the four convergent "tests" and the seven variables on each divergent test was looked at.

A. <u>The Convergent Tests</u>

Scheffé tests were performed on each of the convergent tests. For Similarities, Coloured Progressive Matrices and the Crichton Vocabulary Scale, an overall significant difference between the schools was recorded at the required level of significance of .02, but this was not so for the Picture Completion test, which was non-significant at this level with a P level of .031.

When the Scheffé procedure was done to compare each of the schools, two at a time, and to form homogeneous subsets no such significantly different subsets were found on Similarities, Coloured Raven's Progressive Matrices and Picture Completion. Two subsets were found in the Crichton Vocabulary Scale. Thus, except for Vocabulary the Scheffé test did not find that any pair of schools had means which differed significantly for a subset of a size in this sample.

These results with levels of means are presented in Tables 67, 68, 69 and 70. Standardized scores for each school were used and it should be noted that the Scheffé procedure required a range of 6.06 for a significant difference to be found. Schools are ranked for each subset. The following abbreviations are used:

Table 67

Similarities Scheffé Test Results

	School	Condition	Standardized Mean Scores
	6	DG, CT	50.33
	2	DT, CT	50.19
Subset 1	4	DG, CG	50.12
	7	CG, DT	50.05
	3	CG, DG	49.93
	8	DT, CG	49.86
	1	CT, DT	49.80
	5	CT, DG	49.71

Note: The Scheffé test did not detect significant differences between any pairs of these schools on this variable.

Table 68

Coloured Raven's Progressive Matrices

	School	Condition	Standardized Mean Scores
	6	DG, CT	50.34
	2	DT, CT	50.21
Subset 1	5	CT, DG	50.06
	7	CG, DT	50.03
	4	DG, CG	50.02
	1	CT, DT	49.87
	8	DT, CG	49.78
	3	CG, DG	49.63

Note: The Scheffé test did not detect significant differences between any pairs of these schools on this variable.

Table 69

Picture Completion Scheffé Test Results

	School	Condition	Standardized Mean Scores
	2	DT, CT	50.26
Subset 1	7	CG, DT	50.24
	4	DG, CG	50.10
	3	CG, DG	50.03
	1	CT, DT	49.93
	5	CT, DG	49.98
	6	DG, CT	49.78
	8	DT, CG	49.70

Note: The Scheffé tests did not detect significant differences between any pairs of these schools on this variable

Table 70

Crichton Vocabulary Scale
Scheffé Test Results

	School	Condition	Standardized Mean Scores
Subset 1	2	DT, CT	50.59
Subset 2	1	CT, DT	50.04
	7	CG, DT	50.02
	6	DG, CT	49.95
	4	DG, CG	49.94
	3	CG, DG	49.86
	5	CT, DG	49.86
	8	DT, CG	49.66

CT = Convergent tests

CG = Convergent games

DT = Divergent tests

DG = Divergent games

School 2 (DT, CT) is consistently at or near the top of the list in level of performance, while School 8 (DT, CG) consistently has a low performance. This is in line with earlier reported results that test conditions enhance performance on convergent tasks. School 6 (DG, CT) is at the top of the list on two of the tasks confirming these results, although on Picture Completion this school did relatively poorly. A Scheffé test was done to determine whether subsets could be formed by administrative condition. Two such subsets were formed for the Coloured Matrices and the Crichton Vocabulary scale. Those schools which had test conditions had significantly higher scores than those schools which had play conditions. No such "condition" subsets were found on Similarities and Picture Completion.

Order, from a glance at the Scheffé test results, appears to exert an effect on level of performance in the case of Similarities. A Scheffé procedure for order produced two subsets with those who had the divergent tasks first producing the better convergent result. The exception is School 8, as seen from Table 67, where the lower performance may have been due to the game conditions. No subsets were found on the other convergent variables.

A Scheffé procedure was done to determine whether the conditions under which the divergent tests were taken influenced the convergent test results. No such subsets were found on any of the convergent tests, although the covariance analysis of the Crichton vocabulary test in Table 38 had found such an effect. In that latter analysis it was found that doing the divergent tasks under test conditions enhanced the performance on the vocabulary test. However, the more rigorous Scheffé test procedures did not confirm this result.

B. <u>The Divergent Tests</u>

Scheffé tests were also performed on each variable of the divergent tests. Log transformed scores were used. The level of significance chosen was .02 which meant that in order for significantly different subsets to be selected a range of 6.06 was required. On all 28 variables a significant F value was recorded and on 25 of the variables significantly different subsets were found. The three exceptions were for Fluency 1 and Fluency 2 on Uses of Objects and Originality 1 on Just Suppose. Tables 71 to 98 present the subsets on each divergent variable. Log transformed mean scores are given. The administrative procedure is abbreviated in the following way:

Table 71

Uses of Objects, Fluency 1
Scheffé Test Results

	Schools	Conditions	Log Transform Mean Scores
Subset 1	2	DT, CT	1.41
Subset 2	8	DT, CG	1.24
	6	DG, CT	1.18
	3	CG, CT	1.18
	7	CG, DT	1.16
	4	DG, CG	1.13
	1	CT, DT	1.11
Subset 3	5	CT, DG	1.02

Table 72

Uses of Objects, Fluency 2
Scheffé Test Results

	Schools	Conditions	Log Transform Mean Scores
Subset 1	2	DT, CT	1.42
Subset 2	8	DT, CG	1.25
	6	DG, CT	1.18
	3	CG, DG	1.18
	7	CG, DT	1.17
	4	DG, CG	1.14
	1	CT, DT	1.13
Subset 3	5	CT, DG	1.02

Table 73

Uses of Objects, Flexibility 1
Scheffé Test Results

	Schools	Conditions	Log Transform Mean Score
Subset 1	2	DT, CT	1.23
Subset 2	6	DG, CT	1.04
	8	DT, CG	1.02
	7	CG, DT	0.99
	3	CG, DG	0.98
	1	CT, DT	0.97
	4	DG, CG	0.97
Subset 3	5	CT, DG	0.79

Table 74

Uses of Objects, Flexibility 2
Scheffé Test Results

	Schools	Conditions	Log Transform Mean Score
Subset 1	2	DT, CT	1.32
Subset 2	8	DT, CG	1.14
	6	DG, CG	1.12
	3	CG, DG	1.09
	7	CG, DT	1.08
	1	CT, DT	1.06
	4	DG, CG	1.05
Subset 3	5	CT, DG	0.92

Table 75

Uses of Objects, Originality 1
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	0.97
Subset 2	6	DG, CT	0.62
	8	DT, CG	0.60
	7	CG, DT	0.57
	3	CG, DG	0.53
	4	DG, CG	0.51
	1	CT, DT	0.50
Subset 3	5	CT, DG	0.33

Table 76

Uses of Objects, Originality 2
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.27
Subset 2	6	DG, CT	0.99
	8	DT, CG	0.95
	7	CG, DT	0.92
	3	CG, DG	0.88
	4	DG, CG	0.87
	1	CT, DT	0.83
Subset 3	5	CT, DG	0.61

Table 77

Uses of Objects, Originality 3
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.44
Subset 2	8	DT, CG	1.12
	6	DG, CT	1.11
	7	CG, DT	1.07
	3	CG, DG	0.97
	4	DG, CG	0.97
	1	CT, DT	0.96
Subset 3	5	CT, DG	0.69

Table 78

Patterns, Fluency 1
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.31
Subset 2	6	DG, CT	1.20
	7	CG, DT	1.20
	8	DT, CG	1.19
	3	CG, DG	1.14
	1	CT, DT	1.12
	4	DG, CG	1.07
	5	CT, DG	1.06

Table 79

Patterns, Fluency 2
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.31
Subset 2	6	DG, CT	1.20
	7	CG, DT	1.20
	8	DT, CG	1.20
	3	CG, DG	1.15
	1	CT, DT	1.13
	4	DG, CG	1.07
	5	CT, DG	1.06

Table 80

Patterns, Flexibility 1
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.21
Subset 2	8	DT, CG	1.14
	6	DG, CT	1.13
	7	CG, DT	1.11
	3	CG, DG	1.05
	1	CT, DT	1.05
	4	DG, CG	1.01
	5	CT, DG	0.98

Table 81

Patterns, Flexibility 2
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.25
Subset 2	8	DT, CG	1.17
	6	DG, CT	1.16
	7	CG, DT	1.14
	3	CG, DG	1.09
	1	CT, DT	1.08
	4	DG, CG	1.03
	5	CT, DG	1.00

Table 82

Patterns, Originality 1
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.00
Subset 2	6	DG, CT	0.81
	8	DT, CG	0.80
	7	CG, DT	0.77
	4	DG, CG	0.66
	3	CG, DG	0.65
	1	CT, DT	0.58
Subset 3	5	CT, DG	0.50

Table 83

Patterns, Originality 2
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.35
Subset 2	6	DG, CT	1.17
	8	DT, CG	1.17
	7	CG, DT	1.12
	1	CT, DT	1.04
	4	DG, CG	1.03
	3	CG, DG	1.02
Subset 3	5	CT, DG	0.91

Table 84

Patterns, Originality 3
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.47
Subset 2	8	DT, CG	1.31
	6	DG, CT	1.31
	7	CG, DT	1.29
	3	CG, DG	1.19
	1	CT, DT	1.17
	4	DG, CG	1.16
	5	CT, DG	1.09

Table 85

Inkblots, Fluency 1
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.31
	6	DG, CT	1.28
	7	CG, DT	1.24
	8	DT, CG	1.22
	4	DG, CG	1.17
	1	CT, DT	1.17
	3	CG, DG	1.15
	5	CT, DT	1.14

Note: The Scheffé test did not detect significant differences between any pairs of these schools, on this variable.

Table 86

Inkblots, Fluency 2
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.31
	6	DG, CT	1.27
	7	CG, DT	1.24
	8	DT, CG	1.22
	4	DG, CG	1.17
	1	CT, DT	1.17
	3	CG, DG	1.15
	5	CT, DT	1.14

Note: The Scheffé test did not detect significant differences between any pairs of these schools, on this variable.

Table 87

Inkblots, Flexibility 1
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.21
Subset 2	6	DG, CT	1.18
	7	CG, DT	1.14
	8	DT, CG	1.11
	4	DG, CG	1.09
	1	CT, DT	1.06
	3	CG, DG	1.05
	5	CT, DG	1.04

Table 88

Inkblots, Flexibility 2
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.25
Subset 2	6	DG, CT	1.21
	7	CG, DT	1.18
	8	DT, CG	1.15
	4	DG, CG	1.11
	1	CT, DT	1.09
	5	CT, DG	1.08
	3	CG, DG	1.08

Table 89

Inkblots, Originality 1
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.09
Subset 2	6	DG, CT	1.09
	7	CG, DT	0.93
	8	DT, CG	0.93
	4	DG, CG	0.90
	3	CG, DG	0.82
	1	CT, DT	0.79
	5	CT, DG	0.78

Table 90

Inkblots, Originality 2
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.63
Subset 2	6	DG, CT	1.58
	8	DT, CG	1.52
	7	CG, DT	1.48
	4	DG, CG	1.44
	1	CT, DT	1.44
	5	CT, DG	1.40
	3	CG, DG	1.39

Table 91

Inkblots, Originality 3
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.63
Subset 2	6	DG, CT	1.58
	8	DT, CG	1.52
	7	CG, DT	1.48
	4	DG, CG	1.44
	1	CT, DT	1.44
	5	CT, DG	1.40
	3	CG, DG	1.39

Table 92

Just Suppose, Fluency 1
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.14
Subset 2	8	DT, CG	1.08
	7	CG, DT	1.04
	6	DG, CT	1.01
	1	CT, DT	0.99
	4	DG, CG	0.96
	5	CT, DG	0.93
	3	CG, DG	0.93

Table 93

Just Suppose, Fluency 2
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.14
Subset 2	8	DT, CG	1.09
	7	CG, DT	1.04
	6	DG, CT	1.01
	1	CT, DT	0.99
	4	DG, CG	0.97
	5	CT, DG	0.93
	3	CG, DG	0.93

Table 94

Just Suppose, Flexibility 1
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	0.99
Subset 2	8	DT, CG	0.92
	7	CG, DT	0.91
	6	DG, CT	0.90
	1	CT, DT	0.87
	4	DG, CG	0.86
	3	CG, DG	0.82
	5	CT, DG	0.81

Table 95

Just Suppose, Flexibility 2
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.07
Subset 2	8	DT, CG	1.02
	7	CG, DT	0.99
	6	DG, CT	0.96
	1	CT, DT	0.93
	4	DG, CG	0.91
	3	CG, DG	0.89
	5	CT, DG	0.88

Table 96

Just Suppose, Originality 1
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	0.53
	6	DG, CT	0.40
	8	DT, CG	0.38
	4	DG, CG	0.37
	1	CT, DT	0.34
	7	CG, DT	0.33
	5	CT, DG	0.28
	3	CG, DG	0.27

Note: The Scheffé test did not detect significant differences between any pairs of these schools, on this variable.

Table 97

Just Suppose, Originality 2
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.10
Subset 2	6	DG, CT	0.98
	8	DT, CG	0.98
	4	DG, CG	0.94
	1	CT, DT	0.94
	7	CG, DT	0.89
	3	CG, DG	0.81
	5	CT, DG	0.78

Table 98

Just Suppose, Originality 3
Scheffé Test Results

	School	Condition	Log Transform Mean Score
Subset 1	2	DT, CT	1.25
Subset 2	7	CG, DT	1.13
	6	DG, CT	1.11
	8	DT, CG	1.08
	4	DG, CG	1.03
	1	CT, DT	1.02
	3	CG, DG	0.93
Subset 3	5	CT, DG	0.86

CT = Convergent test

CG = Convergent game

DT = Divergent test

DG = Divergent game

The best performance on all the divergent variables comes consistently from School 2 (DT, CT), which in most cases forms a subset of its own, significantly different from the other seven schools. Most generally, the poorest performance comes from School 5 (CT, DG) which, in many cases, also forms a subset of its own. School 3 (CG, DG) has also a consistently low performance. These findings suggest that test conditions enhance performance on divergent tasks and that order exerts an effect. To do the convergent battery first seems to impair performance on the divergent tasks, whereas to do the divergent battery first seems to enhance performance. Taking all 28 variables into account, two groups of four can be distinguished.

(a) School 2 (DT, CT)
(b) School 6 (DG, CT)
(c) School 8 (DT, CG)
(d) School 7 (CG, DT)

(e) School 3 (CG, DG)
(f) School 4 (DG, CG)
(g) School 1 (CT, DT)
(h) School 5 (CT, DG)

A strong influence of order can be seen in the fact that three of the top four schools had the divergent battery first and three of the lower four schools had the convergent battery first. School 7 (CG, DT), which did well, had the divergent battery under test conditions. School 4, which is among the poorer performers had the divergent battery under game conditions. Indeed, three of the four best schools had the divergent battery under test conditions, and the other school of this group, School 6, had the divergent battery first, although under game conditions. Three of the four lowest schools had the divergent battery under game conditions and the other school of this group had the convergent battery under test conditions first. This suggests a strong condition as well as order effect.

Further Scheffé procedures were done in order to determine the influence of the three experimental conditions in terms of order, condition under which the divergent tasks were taken and the conditions under which the convergent "tests were taken.

When the effect of order was considered, subsets were formed on all 28 variables, with those schools having the divergent battery first having the higher performance level. Subsets were formed on 26 of the variables when condition became the dependent variable with those doing the divergent battery under test conditions having the higher score. The two exceptions, where no subsets were formed, were Inkblots, Originality 1 and Just Suppose, Originality 1, but even here the higher scores came in the schools which had the formal conditions.

No subsets were formed when the condition under which the convergent tasks were taken was considered, although on all variables those schools having test conditions had the higher scores.

Interaction Variables

Several interaction effects were found on both the convergent and divergent variables.

Table 31, which shows the analysis of variance results for the average convergent score shows that there is an interaction between the condition under which the divergent test was taken and the condition under which the convergent test was taken, a two-way interaction between the divergent task condition and order and a three-way interaction between these three variables. On examination of the breakdown table it is seen that the main effect of test conditions applies to only three of the four schools having this test regime. The school which diverges is School 6 (DG, CT). School 4 (DG, CG) which has play conditions for the convergent tasks has the higher score. Thus, it would seem that where divergent tasks taken under play conditions precede the convergent tasks, performance on the convergent tasks will be relatively high even if the latter is taken under game conditions. On the other hand, if the divergent tasks taken under game conditions, precede convergent tasks to be taken under test conditions overall performance will be lowered. Furthermore, divergent assessment under formal conditions preceding convergent tasks under game conditions interacts to inhibit performance on the latter, as seen in School 8 (DT, CG). This result means that Hypothesis 2, which states that being subjected to two types of conditions will not effect the level of performance on convergent tests, is rejected.

The average convergent scores in ranked order are presented in Table 99. Values are presented as corrected standardized means.

To re-emphasize, Schools 4 and 6 differ due to the fact that although both have divergent games preceding the convergent battery, School 4 has convergent games and School 6 has convergent tests. The order x convergent interaction is also seen in the very significant difference between Schools 2 and 8 where both have divergent <u>tests</u> preceding the convergent battery, but whereas School 2 has convergent tests, School 8 has convergent games. Thus, if the divergency type is formal and precedes the convergent battery, better performance is gained by having formal

Table 99

Average Convergent Scores of Each School in Ranked Order

	Mean
School 2 (DT, CT)	+.39
School 1 (CT, DT)	+.30
School 5 (CT, DG)	+.11
School 4 (DG, CG)	+.10
School 6 (DG, CT)	+.02
School 7 (CG, DT)	+.01
School 3 (CG, DG)	-.20
School 8 (DT, CG)	-.58

rather than informal convergent conditions.

These explanations do not apply when each convergent task is taken into consideration. In two of the results, School 6 (DG, CT) has the highest score (Tables 67 and 68), implying that the divergent game conditions preceding the convergent tasks enhances performance on specific convergent tests. The principle, that formal testing of divergent thinking preceding convergent performance under game conditions inhibits convergent thinking, still applies for School 8 (DT, CG) has a consistently low performance level.

A two-way interaction between the conditions under which the convergent test was taken and the effect of order is seen to affect the average divergent score in Table 32. An examination of the breakdown table provides the explanation. The main effects on the average divergent score were order and condition, with higher performance, found on those schools who had the divergent tasks first, and did the divergent tasks under test conditions. There is no interaction between these two variables on the average divergency score. The aberrant schools not fitting this scheme are School 7 (CG, DT) and School 4 (DG, CG), and to a lesser extent School 1 (CT, DT), and School 6 (DG, CT).

School 7 (CG, DT), which is among the top scorers had the divergent battery under test conditions as did School 1 (CT, DT). Both these schools had the divergent battery preceded by the convergent battery which, as we have seen, dampens divergent scores. However, this interaction effect suggests that if the convergent battery is taken under game conditions this has a less inhibitory effect on divergent thinking than if the convergent tasks are given under test conditions. The difference in performance between Schools 4 and 6, both of which had the divergent battery preceding the convergent battery, a feature which enhances divergent scores, but also took the divergent battery under game conditions, is due to the differences between the schools other than the experimental dependent variables. However, the results attained in these interactions suggest that Hypothesis 4, which states that being subjected to two types of conditions will not effect the level of performance on divergent "tests," is rejected.

The average divergent scores in ranked order are presented in Table 100. Values are presented as corrected standardized means.

The interaction between order of taking the "tests" and the conditions under which the convergent "tests" are taken is the most common interaction among the divergent variables seen in Tables 39 to 66. The interaction between

Table 100

Average Divergent Scores of Each School in Ranked Order

	Mean
School 2 (DT, CT)	+.81
School 6 (DG, CT)	+.11
School 7 (CG, DT)	+.10
School 8 (DT, CG)	+.10
School 1 (CT, DT)	-.13
School 4 (DG, CG)	-.21
School 3 (CG, DG)	-.27
School 5 (CT, DG)	-.57

order of taking the "tests" and type of condition for the divergent "test" is of the form that better performance occurs if the divergent task is taken under "test" conditions and precedes the convergent task. The interaction between type of condition for the convergent "test" and type of condition for the divergent "test" seen on Uses of Objects is explained by the fact that if convergent tasks precede the divergent tasks, the inhibitory effect of convergency is less if those tasks are given under game conditions, and divergent performance is under test conditions. Thus, School 7 (CG, DT) does much better than School 5 (CG, DG). This strengthens the rejection of Hypothesis 4.

Intercorrelation Results

By means of a regression analysis with the average convergent score as the constant and the average divergent score as the dependent variable, correlation coefficients between these two scores were computed to determine whether convergent and divergent thinking were significantly related, and whether this relationship changed with different experimental procedures. Therefore, correlations were computed for each school. These results are presented in Table 101.

An analysis of variance of these results <u>within</u> each school revealed that at the .02 level of significance only

Table 101

Correlation Coefficients of Average Convergent and Divergent Scores

School	Conditions	Correlation Coefficient (r)
1	CT, DT	.23
2	DT, CT	.01
3	CG, DG	.24
4	DG, CG	.29
5	CT, DG	.37
6	DG, CT	.34
7	CG, DT	.12
8	DT, CG	.35

those correlation coefficients or r values in School 5 (CT, DG), School 6 (DG, CT), and School 8 (DT, CG) were significant. Thus, only in these three schools were convergent and divergent thinking significantly related. In the remaining schools, the r values were not significant; i.e., convergent and divergent thinking were not significantly related.

It is noteworthy that these three schools, which showed a significant relationship between convergency and divergency had a "mixed" set of conditions; i.e., had their convergent tasks under one set of conditions and their divergent tasks under another set of conditions. There was one case of "mixed" conditions where this was not so, namely, School 7 which had the procedure CG, DT. In this school, as in the schools which had all test or all game conditions, convergent and divergent thinking are relatively independent.

Thus, for the greater part, Hypothesis 7 is rejected. Although there is some variation, according to condition, in the relationship between convergent and divergent "tests," the correlations were not high when both "tests" were taken under formal conditions: Rather they were low. Similarly r values were low under both game conditions. The predicted "lowest" r value between convergent and divergent thinking was for the procedures CT, DG or DG, CT.

Both produced significant r values and the conditions CG, DT produced a very low r value of .12.

The lowest r value of .01 was computed for School 2 (DT, CT). This compares with the r value of .23 for School 1 (CT, DT).

The difference in values between some schools having the same set of conditions seems substantial enough to claim that order of testing may affect the relationship between convergent and divergent thinking. A difference occurs between School 7 (CG, DT) and School 8 (DT, CG) where in the former the two variables are independent and in the latter they are related. The difference between School 1 and School 2 is also noteworthy. In contrast, the r values are very similar between School 3 (CG, DG) and School 4 (DG, CG). Similar r values also occur between School 5 (CT, DG) and School 6 (DG, CT). Thus, in four schools, order of testing does exist on effect. This means that Hypothesis 8, which states that order of giving "tests" will not affect the relationship between convergent and divergent thinking, is rejected.

Pearson Product correlation coefficients were computed between the four convergent variables, the Coloured Progressive Matrices, Picture Completion, Similarities and the Crichton vocabulary scale. In all cases the r values between each of the variables was significant, indicating that the construct validity of these four variables over all the different conditions is high. Table 102 presents these results for the eight schools combined.

The highest r value of .66 occurs between the two verbal tests of Similarities and the Crichton Vocabulary Scale. The lowest value of .36 occurs between the verbal vocabulary test and the non-verbal Coloured Progressive Matrices test.

Pearson Product correlation coefficients were also computed between each variable of each divergent test. All these were significant, values ranging from 1.00 between the two fluency scores of all tests to .51 between Fluency 1 and Originality 1 on Just Suppose.

These results are presented in Appendix F. It should be noted that all correlations are rounded off to two decimal points so that the 1.00 values actually represent values of .995 or more. Each table represents the eight schools combined.

These tables in Appendix F show that each variable within any one test is extremely high, particularly so when Fluency 1 is compared with Fluency 2; or Flexibility 1 with Flexibility 2; or a comparison is made of the three Originality scores.

When each variable from one divergent test is compared with each variable on another divergent test r values are

Table 102

Pearson Product Correlation
Coefficients of the Convergent
Tests for the Eight Schools Combined

	CPM	PC	SIM	VTS
CPM	1.00	0.40	0.41	0.36
PC	0.40	1.00	0.47	0.55
SIM	0.41	0.47	1.00	0.66
VTS	0.36	0.55	0.66	1.00

Key:

CPM = Coloured Raven's Progressive Matrices

PC = Picture Completion

SIM = Similarities

VTS = Crichton Vocabulary Scale

lower than the ones computed above, but they are all significant. Values range from .61 between the flexibility scores of Patterns and Inkblots, Uses of Objects and Just Suppose to .30 between the fluency scores of Inkblots and the Originality 1 score of Just Suppose. Generally, r values between fluency, flexibility, or originality between the different divergent tests lie above .40. Thus, the construct validity of the divergent tests is high across different conditions. These results are presented in Appendix G. Pearson Product moment correlation coefficients were computed. Results represent the eight schools combined.

Other Effects

Table 2 in Chapter IV shows the distribution of ages (in months) and Otis IQ scores. An Analysis of Variance was done to determine whether there was a significant difference between the eight schools on these variables of age and Otis IQ. On both variables there was a very significant difference, well above the required probability value of .02.

A. Age

Pearson Product moment correlations between age and each convergent variable revealed that there was no significant relationship between age and scores on the Similarities, Coloured Progressive Matrices, Picture Completion or the Crichton Vocabulary Scale. However, age did affect significantly the uncorrected average convergent scores. Table 103 presents the r values obtained during a regression analysis of age and convergent scores. Schools are ranked according to the magnitude of the r value.

There is a definite effect of order here as well as convergency type. Thus, Schools 1 and 5 are paired as are Schools 7 and 3. All had the convergency tests first but the first pair had formal conditions, the second pair had game conditions. Schools 2 and 6 are paired as are Schools 8 and 4. All had the divergent tasks first, but the first pair had test conditions, the second pair game conditions on the convergent tasks. The chances of getting this order between two variables over eight conditions is one in 40,000 so that this rationale formed the basis for correcting the convergent score for age. The convergent score for Schools 1 and 5 had an age correction of .13; Schools 3 and 7 had an age correction of .06; Schools 2 and 6 a correction of -.13; and Schools 8 and 4 a correction of -.27.

Pearson Product moment correlations between age and each divergent variable produced four significant r values on Fluency 2 of Uses of Objects, Originality 2 of Patterns, and Flexibility 2 and Originality 3 of Inkblots. This

Table 103

r Values Between Average Convergent Score and Age

School	Condition	r Values	
1	CT, DT	.14	.13
5	CT, DG	.12	
7	CG, DT	.09	.06
3	CG, DG	.02	
2	DT, CT	-.10	-.13
6	DG, CT	-.16	
8	DT, CG	-.23	-.27
4	DG, CG	-.31	

seemed to be a random distribution of significant correlation values for age and divergent score, particularly since there were 28 divergent variables. When the effect of age on the average divergent score by school was examined, no rationale for correction of scores by age could be determined so that the divergent score was not corrected for on this variable. Table 104 presents the r values obtained during a regression analysis of age and divergent score. Again, schools are ranked according to the magnitude of the r value.

B. Otis IQ Scores

Since there was a difference in IQ levels between the eight schools covariance analyses were included when each convergent or divergent variable was considered separately.

When the distribution of scores within any one school is examined it is seen that the IQ distribution is skewed in most schools with very few people in the low IQ range. This is particularly so with Schools 6 and 7. School 6, with a total number of 63, had a range of Otis scores between four to 64. Twenty-seven pupils had scores over 45 and there were 21 children with scores over 50. School 7 had a total number of pupils of 69 and a range of Otis scores between nine and 54. Twenty-six children of this sample had IQ scores of 45 or more and 10 children had scores of 50 or more. The distribution of Otis scores are

Table 104

r Values Between Average Divergent Score and Age

School	Condition	r Value
1	CT, DT	.11
8	DT, CG	.08
3	CG, DG	-.01
4	DG, CG	-.03
6	DG, CT	-.04
2	DT, CT	-.07
7	CG, DT	-.08
5	CT, DG	-.30

presented in the form of histograms for each school in Appendix H.

With respect to the relationship of the Otis IQ to the average convergent score, regression analysis revealed that the r value was different for each sex. For the boys the r value was .72; for girls the correlation between the average convergent score and Otis IQ was .69. The convergent scores were, therefore, corrected for each sex separately.

In the case of the rationale behind the relationship of Otis score and average divergency score, regression analysis revealed different r values according to the order in which such tests were taken. Where the order was convergent tasks preceding the divergent ones, the r value of divergency with the Otis was .19. If divergency came first then the r value was .36. These r values then became the correction scores applied to the average divergent score to correct for the Otis. Such r values also reveal the change in the relationship between divergent thinking and Otis IQ. Here order affects the independence of the two variables, divergent thinking being a more independent variable when convergent "testing" has preceded its expression. These results are summarized in Table 105.

Table 105

Correlation of Otis Score and Average Divergent Score

School	Condition	r Value
1	CT, DT	
3	CG, DG	.19
5	CT, DG	
7	CG, DT	
2	DT, CT	
4	DG, CG	.36
6	DG, CT	
8	CG, DT	

Other Correction Factors

No other correction factors needed to be made for the convergent score. However, with respect to the average divergent score each variable had to be log transformed due to the fact that raw scores were so extreme. For example, the scores on Fluency 2 of Uses of Objects ranged from 1 to 204! On this particular variable six pupils had a score lower than three and there were only three people who had a score over 85. The mean score value was 17.40. In such circumstances it was not surprising that the kurtosis value was 84.25. On the Originality 1 variable of Patterns, scores ranged from 0 to 60 with 105 children having a score under two and only two people having a score over 50. The mean raw score value for this variable was 7.10 and the kurtosis value was 10.79. This was the typical pattern for all 28 variables.

In addition to this there was a significant difference on each divergent subscore between boys and girls. Therefore, each divergent variable was corrected for in terms of sex gender in computing the average divergent total.

These corrected scores were used for the analyses of average convergent and divergent scores described in the early part of this chapter. Where each variable was considered separately for statistical analysis such a rigorous correction procedure did not take place. Despite this, the same major results emerge on both the average scores and the subscores providing strong support to the major conclusions to be discussed in the next section.

CHAPTER VI

DISCUSSION

Levels of Performance

The major findings of this research seem well-defined, the main results having been verified over several statistical analyses. The purpose of this discussion is to seek interpretations and implications of the main effects which were:

1. Children perform better on convergent tasks under test conditions;
2. Children perform better on divergent tasks under test conditions;
3. Order of tests exerts an effect on divergent scores. Children perform better on the divergent tasks when the divergent battery precedes the convergent battery.

For the average or total convergent score there are no order effects except with respect to interaction phenomena. Since adjusted scores were used in these analyses the explanations for the interactions can be accepted with confidence. The interaction effects, as applied to the convergent score imply that although children perform well on these tasks if formal administrative procedures are given, and that preceding such testing with a divergent task does not usually affect convergent performances, the type of condition applied to divergent testing does. Thus, if convergent tasks given in a formal atmosphere are preceded by divergent tasks, the best levels of performance occur when the preceding divergent test atmosphere is formal. Overall performance on convergent tasks is lowered if both types of conditions are applied. To precede formal convergent testing with divergent games lowers the overall convergent score. If divergent games are to precede convergent tasks it is better to give the latter under game conditions also.

The disadvantageous effect of preceding formal convergent tests with divergent games, however, does not always apply since results on certain separate convergent "tests" indicate that performance may be enhanced under such conditions. This is particularly so for Similarities where the effect of divergent games on convergent thinking was to enhance reasoning by analogy. In contrast this particular regime is extremely disadvantageous in a task such as Picture Completion. The worst conditions for convergent performance are those where a convergent battery to be taken under game conditions is preceded with formal divergent testing. School 8, which had this procedure, had a consistently low performance on the convergent tests.

For the average or total divergent score as well as most of the divergent variables there are the two main effects of condition and order. Surprisingly the interaction between order and condition was rare. The interaction effect on the average divergent score and on most of the divergent variables resulted from convergency type and order. Best levels of performance occur when divergent "tests" taken under formal conditions precede the convergent battery. Scores are generally lower if the convergent battery is given first. However, under circumstances where the convergent tasks do precede the divergent formal testing, better results are obtained if the convergent battery is given under game conditions. In other words, a preceding set of convergent tasks in a play atmosphere has a less inhibitory effect on divergent thinking, as long as the latter is administred in a formal atmosphere. The greatest inhibitory effect on divergent performance is to precede divergent tests with a convergent test battery given under formal conditions and then to give the divergent battery under game conditions. School 5 which had this regime had consistently low levels of performance.

Thus, to precede divergent tests with convergent tests, has an inhibitory action on divergent thought processes and this is exacerbated if the convergent tasks are formally administered together with the divergent tasks being taken in a play atmosphere.

It would appear to be less detrimental if the convergent tasks are given in a play-like atmosphere and the divergent tests are formally administered. However, it is better not to give the convergent battery until after the divergent tasks are given in a test atmosphere.

For both the convergent and divergent tasks, statistical significance of the rankings of level of performance was done with the Scheffé tests. School 2 had consistently high scores and had the highest overall convergent and divergent scores. Such a result implies that this regime is the most advantageous; i.e., DT, CT.

Other Scheffé test results produced less significant subsets than would have been expected from the obtained F values from the ANOVA which were highly significant. On the divergent tasks, School 5 (CT, DG) usually emerged as a separate subset, having low scores, but greater differentiation among the other schools was not apparent. The explanation for this may be due to the rigorousness of the Scheffé test itself. Since the Scheffé test is more rigorous than other procedures it leads to fewer significant results and Scheffé (1959) recommended that a lower level of significance be chosen when applying it. Had this been done in this research a greater distinction between the schools would have been observed.

Despite this, however, very definite results are obtained with respect to order and condition. The divergent thinking results are at variance with those investigations reporting better levels of performance on such "tests" under play conditions. Boersma and O'Bryan (1968), Vernon (1971), and Hargreaves (1974), for instance, found divergent thinking scores enhanced under a play regime. The research results on divergent thinking in the present study agree with those of Williams and Fleming (1969), Kogan and Morgan (1969), Leith (1972), and Channon (1974) which find divergent thinking scores elevated under test or evaluative conditions. Better performances under test conditions suggests that an evaluative context by imposing a mild stressful situation serves to induce the optimal arousal required to perform at one's best. It would appear that this same alertness is not induced under play conditions where the atmosphere is relaxed and the children have no pressure placed upon them to perform well.

However, this explanation alone is too simplified. The present research results would have merely added to the confused literature of contradictory evidence had it not been for the introduction of the variable of order which effected performance on the divergent tests but, generally, not the convergent tests. Indeed, it is possible that this strong effect of order may explain why previous research results have varied from one investigation to another. In this investigation having the divergent tasks first tended to produce better results. Thus, School 6, although having a game atmosphere, had the divergent battery first, and thus did better on divergent tests than School 1, which, although having formal conditions, had the convergent battery first.

The analysis of the findings, already described in this chapter, emphasize the inhibitory effect of convergent thinking preceding divergent thinking. The inhibition of divergent thought processes is strongest when the convergent assessment is under formal conditions. Most previous investigations into the effect of conditions on divergent thinking performance, have used evaluative procedures when giving their convergent battery and have not taken into consideration this feature of order and its presumed effect on set. In many such investigations the authors do not specify whether or not the convergent battery preceded the divergent battery. Boersma and O'Bryan (1968) are an example of where such a procedure is stated. They gave the convergent battery to both groups first. The group which had formal testing throughout, CT, DT had lower scores than those with the regime CT, DG; however, the present author claims that since this particular order depresses divergent behaviour, the signifiance of their results is limited. If the procedure had been switched to that of DT, CT for Group 1 and DG, CT for Group 2 their results may have been very different and more in agreement with those of the present investigation.

The reason that convergent thinking tends to depress performance on a subsequent divergent task may be due to the strbng set influence a convergent thinking context has. With respect to this point of set, Harris and Evans (1974) found that college students exposed to a convergent model showed fewer divergent responses than those exposed to a divergent model. Thus, in the present investigation having the convergent tasks first may have narrowed down the focus of these subjects, placing them in a certain cognitive set which they applied when the divergent tasks were presented. The strongest convergent model would have operated when convergent performance was under a formal context, although such a statement needs qualification. The influence of a formal convergent model is most detrimental if the context of the divergent tasks changes to a game condition. Under these conditions children seemed unable to switch successfully from one cognitive style to another. This set is not apparent when the convergent model is presented in a game-like context and the subsequent divergent sessions are formal.

The implication of this finding is that exposure to a convergent model should not precede divergent tests. At the present time elementary (and the secondary) school students not only spend more time on convergent-like activities, but divergent activities, such as writing an English essay, usually come after a long session of convergent-type tasks. The findings of this research suggest that rearrangement of the structure of the curriculum should occur so that the divergent activities come at the beginning rather than at the end of the school day.

In addition, assessments should place divergent-like problems before convergent problems, since the problem of order does not detrimentally affect convergent thinking if the convergent tasks are given in a formal context.

The findings that convergent performance levels are higher in a formal context has not been previously reported in the literature to such a major extent as in this investigation. Less extensive investigations, such as Williams and Fleming (1969) and Channon (1974) suggested that convergent scores would improve under evaluative conditions and this study provides definitive evidence to support that claim.

That a formal atmosphere is best for convergent and divergent thinking has some interesting implications for our educational system where many of the lessons are taken in an informal atmosphere. Many of the complex learning skills, such as mathematics are presented to the children initially in a more game-like context. Bruner, Jolly and Sylvia (1976) point out that the function of play is

> to reduce or neutralize the pressure of goal-directed action, the 'push' to successful completion of an act. There is a well-known rule in the psychology of learning, the Yerkes-Dodson law, that states that the more complex a skill to be learned, the lower the optimum motivational level required for fastest learning. Play, then, may provide the means for reducing excessive drive and frustration, (p. 15).

Yet the results in the present research would indicate that the conditions of play lower "the optimum motivational level" and the optimum arousal level so that while the anxiety level is low, performance is also at a low level. This may also be due to the fact that a game-like atmosphere is more ambiguous about what is required than the test atmosphere. The play-like context provides less obvious cues which the child can use to perform his tasks well. The evaluative or formal atmosphere, despite the fact that it is more stressful than play seems to provide the necessary unambiguous framework with sufficient reference points for the children to work at higher performance levels on convergent and divergent thinking problems.

The fact that, generally, tests of convergent thinking were not affected by order such as taking preceding divergent tests, whereas the performance on the divergent tests was so affected, is important. It suggests that divergent thinking is a more labile ability and more sensitive or more responsive to situational conditions. This implies that measuring divergent thinking or in attempting to improve divergent performance, attention should be paid to contextual effects. High level performance of divergent thinking has to be focussed as well as "expansive." Play conditions would appear to allow the individual to attend to too many environmental cues and preceding convergent tasks would appear to narrow the focus of thought too much so that very few possible responses can be considered.

During the literature review it was predicted that high formal, high convergent children would do better under test conditions, whereas those with a bias towards divergent thinking would do better under play conditions. Yet in this study children over a wide range of abilities did both the convergent and divergent tests better in a formal atmosphere indicating that Rohwer's (1971) conclusions appear to be wrong concerning better performance of imaginative conceptual tasks under informal conditions.

The evidence further suggests that since children, over a wide range of abilities, perform better under formal conditions on convergent and divergent tests, both abilities may be inherited. Since divergent thinking, unlike convergent thinking, was affected by order, which was not a predicted outcome, suggests that this type of thinking is more responsive to the internal state and condition of the organism in addition to the external conditions of evaluation. Hence, divergent, like convergent, ability may be inherited but whether such abilities are expressed or not depends on the environmental context.

An analogy from the plant kingdom seems appropriate. Some plants will open and close depending on certain environmental influences; however, their ability to open and close is inherited. Thus, although the ability is inherent in the organism, its expression is dependent on environmental factors. Similarly with children's abilities, where the expression of abilities depends on such factors as arousal, reading of environmental cues, mood and emotional state. In particular the results of this research suggest that divergent thinking has a more sensitive switching on and switching off mechanism where the off trigger is convergent thinking. Divergent thinking also appears to be more susceptible to conditions of arousal and the intentionality of the subject. While on the surface this might suggest that it is less an ability factor than a motivational one, this conclusion does not follow. Nor does it follow as explained above that because divergent tests are more influenced, apparently, by expectations, that divergent thinking is therefore less heritable in nature.

Scoring Methods of the Divergent Tests

A great deal of effort went into formulating a scoring scheme for the divergent tests which would be both reliable and meaningful. Many of the studies cited in this thesis use the scoring scheme of fluency and originality of content, which were the variables Fluency 2 and Originality 3 in this research. Ward (1969) used a different set of criteria for marking for fluency, particularly in his elimination of superordinates. This was Fluency 1 in this study and has been previously discussed in Chapter IV. Torrance, through his many studies, has devised a flexibility scheme (e.g., Torrance, 1962). Using flexibility as a score seems important for investigating divergent thinking since such a variable introduces the criteria of quality of the response rather than mere response quantity. It answers the question as to whether the subject remains within one set of ideas and preservates within that set, or is able to shift his set and move into another category. The problem is to determine how much shift is required to classify a person's flexibility. Thus, in this investigation two flexibility scores were created to record shifts between major categories and between more minor shifts, such as differentiating between a response which is characterized by movement and a response which is characterized by elaboration. Originality 1 and 2 were the scores recording the extent to which each subject used common or unusual categories. This was done on a frequency basis; i.e., in terms of how rarely each major or minor category was used. Such scores have not been used in other previous research cited in this thesis but would seem valuable in the sense that originality of idea category would seem to be essential in creative behaviour. Originality 3, on the other hand, recorded the unusualness of the response content.

Within any one divergent test the variables are all very highly related and in particular there is little difference between the two fluency scores, the two flexibility scores or the three originality scores. Thus, a future investigator may choose to use the Fluency 2 score for fluency since it is a simpler scoring method and avoids the issue as to whether superordinates are a form of repetition. Flexibility 2 has a slightly higher correlation coefficient with fluency on the four tests, which indicates that recording minor shifts in thought patterns is very closely related to fluency and cannot properly be distinguished from it. It would seem sufficient in future research to use only Flexibility 1, thus recording shifts in thoughts between major categories such as animal, animal detail, human and human detail. From the three originality scores, originality of content (Originality 3) is the most closely related to fluency. Together with originality of minor categories (Originality 2), these scores are very closely related to the flexibility categories also. The originality score which differs the most from both fluency and flexibility is originality of major category (Originality 1). This suggests that Originality 1 may be a more useful criterion in assessing divergent thinking performance. This new method is also a much more rapid scoring method than that of originality of content.

Construct Validity

The result section described the relatively high correlations which existed between each of the four divergent tests and the four convergent tests. As would be expected the r values between variables of the two non-verbal tests or between variables of the two verbal tests are higher than when a verbal and non-verbal test are compared. This indicates that as on verbal and non-verbal intelligence tests, some different cognitive processes are required. But at the same time the r values between the verbal and non-verbal tests are of sufficient magnitude to suggest that there are common cognitive abilities underlying both types of tests. The same reasoning is applicable to the convergent tests. The correlation results, therefore, suggest that both the convergent and divergent batteries have high construct validity.

The Relationship Between Convergent and Divergent Thinking

The generation of hypotheses is the essence of divergent thinking but also requires that the individual have a perspective on what he is generating hypotheses about. Thus, divergent thinking would appear to be somewhat dependent on some convergent processes and would help to explain the lability of its nature found in this research. If this is so then one would expect convergent and divergent thinking to be related as was predicted. Yet the results of this research concerning the relationship of convergent and divergent thinking would appear to contradict these assumptions.

Correlations computed between the average convergent and average divergent scores were generally low, although in three cases, School 5 (CT, DG), School 6 (DG, CT) and School 8 (DT, CG), the values were statistically significant. In the other five schools the r values were of sufficiently low magnitude to indicate that convergent and divergent thinking were relatively independent variables. Wallach and Kogan (1965) based their claim that convergent and divergent abilities were independent on the finding that none of their correlation coefficients exceeded .23. Among the non-significant correlations the highest score was .29, which was School 4 (DG, CG), School 3 (CG, DG) had a correlation coefficient of .24 and School 1 (CT, DT) had a r value of .23. School 7 (CG, DT), was lower at .13 and the lowest was School 2 (DT, CT) with a value of .01.

With respect to the first four schools, there is little difference between Schools 3 and 4, suggesting that order does not affect the relationship between convergent and divergent thinking when both cognitive tasks are given under play conditions. There is also little difference between Schools 1, 3 and 4, indicating that in terms of these three schools there is little difference in the relationship if both the convergent and divergent are given under the same conditions, either all test or all play. Because the r value is low, the predictions of Wallach and Kogan (1965), Lundsteen (1966), Boersma and O'Bryan (1968) and this investigator, that the correlation coefficients would increase between convergent and divergent thinking if only one condition was applied, were not fulfilled. However, the difference between School 1 (CT, DT) with an r value of .23, and School 2 (DT, CT) with an r value of .01, is of sufficient magnitude to suggest that the order of giving the tests may affect the relationship between the two cognitive modes under formal conditions. There is a more definitive separation of these two abilities when the divergent tests are given first.

This same tendency is seen when we compare School 7 (CG, DT) and School 8 (DT, CG). Here School 8 has an r value of .36 which is statistically significant, and School 7 has a non-significant r value of .13. In these schools the interaction of these particular conditions (game conditions of the convergent tasks; and formal conditions of the divergent tasks) with two different orders produces two quite different relationships. This effect of order of presentation of tests is not seen with Schools 5 and 6, both of which have formal conditions for the convergent tasks and game conditions for the divergent tasks. This is a similar experimental procedure used by Wallach and Kogan (1965) and others who have attempted to replicate their work, a review of which was given in Chapter III. In this investigation the correlations of Schools 5 and 6 of .37 and .34, respectively, are similar to those found by Hargreaves (1974) of .31 and higher than that found in the study by Vernon (1971) of .26, using a similar set of conditions. Under this particular regime CT, DG or DG, CT, the two modes of thought appear related in these studies as it is in the present investigation.

However, in not one of the schools is the correlation coefficient noteably high. Had the numbers in each experimental group been greater, a factor analysis would have been done with the prediction after Cropley (1966, 1968) and Cropley and Maslany (1969) that two oblique second order factors would have appeared, one of convergent abilities and one of divergent abilities. But it is predicted that future analyses would reveal one first order factor representing a large general factor (after Cropley & Maslany, etc.).

There may be other reasons for these relatively low r values between convergent and divergent thinking. The sample of 493 in the main study was carefully selected to be a representative sample of middle-class, white Australian children. They were, therefore, a fairly homogeneous group as in the Wallach and Kogan (1965) study. Cropley (1966) has pointed out that correlations between convergent and divergent thinking tend to be lower in homogeneous samples because of the restriction of range. In addition to this the average Otis IQ in each of the schools was fairly high ranging from 109 to 119, with a mean of 115.

It is possible, therefore, that the IQs in each school were of sufficient magnitude so that a minimum threshold value had been exceeded. The threshold theory suggests that substantial correlations between the two modes of thought exist only at the low intelligence levels. Beyond this the relationship decreases and after some critical value the two become independent. The IQ levels in each of the schools may have been high enough to lower the degree of relationship between convergent and divergent thinking.

CHAPTER VII

CONCLUSIONS

In terms of future research these findings have several implications. The relatively low correlations obtained between the convergent and divergent tests suggest that it is not possible to substitute one for the other in mental assessment. The present findings would also suggest that if one was studying the heritability of mental factors one would have to study the heritability of divergent thinking separately. This is because convergent and divergent thinking abilities appear to be relatively independent so that finding that convergent ability has a large heritable component tells us nothing with respect to divergent thinking.

Since the sample was large (493) and the experimental procedure so comprehensive, embracing the independent variables of both conditions and order, some rather definite conclusions can be stated. The investigation provides substantial evidence to suggest that convergent and divergent thinking assessment should take place in a formal context. Equally important is the finding that order of presentation of convergent and divergent stimuli will affect the performance level on divergent tests. Divergent tasks should precede convergent tasks. Convergent "tests" tend to inhibit divergent thought processes. This effect of set or order has not previously been reported in the convergent/divergent thinking literature. It needs further research, particularly to identify the mechanism involved in the inhibitory processes of convergent thinking on divergent tests. Then such questions arise as to the applicability of these findings to other age and socioeconomic groups, particularly that of order.

The research findings on the relationship of convergent and divergent thinking were less definitive, perhaps as was suggested because of the relatively homogeneous sample and the high average Otis IQ. However, the results did indicate that the higher correlations were scored in those schools where a combination of test and play conditions were applied. School 7 (CG, DT) was the exception. The lower correlations came from those schools where the same conditions applied to all tests, but it is possible that School 7 belonged to this group because convergent games may have had some resemblance to the divergent tasks which followed. The greatest separation of divergent and convergent thinking was found with School 2 (DT, CT). School 2 had also the highest level of performance on both the convergent and divergent tasks. Thus, convergent and divergent thinking under a DT, CT regime appears not only to be more productive but have a richer psychological meaning in terms of understanding these two cognitive processes. Although overall correlations between convergent and divergent thinking are sufficiently low to suggest that they may be separate clusters of abilities, it was felt that further statistical analysis might show that there is a more general ability underlying them both.

Convergent thinking is a more stable dependent variable, being influenced generally by situation and not order. Such thinking is perhaps the more crystallized and more organized capacity. Divergent thinking is a more labile cognitive process being influenced by both order and condition. It is situation responsive in the sense that the subject makes an appraisal of the situation and responds to the demand. It is also set responsive in that the preceding situation and type of task determines the nature of the response. Put another way, convergent thinking appears as a trait; i.e., a more stable and predictable characteristic of behaviour. Divergent thinking appears more of a mixture of ability (or traits) and a state, which is a more transient mode of cognitive adjustment made to certain problem-solving approaches. Being influenced by situation and set, divergent thinking can be "turned off" under certain circumstances even in the presence of appropriate stimuli. It can also be switched off after producing the relevant responses if the presenting stimuli calls for a convergent response. This is unlike convergent thinking which, once "turned on" tends to persist even in inappropriate situations. This could explain the success of School 2 (DT, CT) and the relatively poor performance levels of those schools where convergent thinking preceded the divergent thinking.

This research answered the question as to what set of administrative procedures is most effective in producing high performance levels in Grade 5 children. A mentioned earlier, questions remain as to their applicability to other age groups and to other populations. Such populations include those from different socioeconomic groups and those populations outside the Australian school system. Most importantly, the question left remaining involves the cognitive mechanisms or processes involved in enhancing or inhibiting convergent or divergent thinking. Why does DT, CT impair convergent performance, and CT, DG impair divergent performance? Some explanations for such findings have been given but others need to be explored. These include identifying personality and motivational influences, to determine the extent to which affective components influence the importance of test atmosphere and order or set on the expression of convergent and divergent thinking.

The present study, though providing clear findings, raises further questions as does most research. These questions seem important because of the intrinsic significance of higher order cognitive functions. It would seem that effectiveness of an individual depends upon the ability to shift from convergent to divergent modes of thinking. There are times to focus on a solution, and there are times to focus on generating alternatives. Given the fact that divergent thinking in this research was found to be more susceptible than convergent thinking to the effects of set, it would be valuable to find out whether children can be trained to shift easily without carryover effects from one mode of thinking to another. This assumes that the underlying abilities themselves remain unchanged and that set determined how the basic abilities will be used. Further research is also needed on the relationship between convergent and divergent abilities to determine how independent they are in fact.

The fact that divergent abilities were found to be more responsive to order effects suggests that divergent thinking has a more evanescent quality, as has been suspected by the use of the term creativity. The capacity for divergent thinking may be a latent matter for the individual if the proper state or mental set is not established to exercise the capacity. Broader issues also need research such as attempting to alter the relative levels of convergent and divergent thinking in arts and science students. The results of such training attempts should also have to be studied in relation to the amount of transfer to daily life performance as well as task performance in school-like situations or professional roles.

While there has been a long and productive history of research on intellectual abilities, much remains to be done if more children are to be brought to a fuller expression of their latent abilities. Also, the possible shifts in mode of thinking over the adult years need to be explored. The important topics of research for the adult years may not be so much "can we think" but "how we think," or less a matter of capacities than how these capacities are deployed. The present research barely scratched the surface of the nature of convergent and divergent thinking, but some of its implications can be seen. Much remains to be learned by extending studies like the present one to studying the entire life span.