Scientific ability, Spatial ability and Formal thinking in adolescence

by Christopher John Purcell B.Sc.
(B.P. Research Fellow)

Thesis submitted for Ph.D. to The University of Warwick.

The research was undertaken at the Department of Science Education of Warwick University as part of a research fellowship funded by B.P. Co. Ltd.

Submitted November 1984.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The author would like to thank British Petroleum plc for the research grant which made the work included in this thesis and other associated development work for pupils showing ability in science, possible.

Thanks also to Dr. Peter Screen of the Department of Science Education at Warwick University who supervised the work always being totally supportive of the project even in its early stages when the full pattern of the ensuing study had not totally developed.

In this respect thanks are also due to Mr. John Staples of the Department of Education who took a continued supportive interest in the work from its inception and to Mr. William Penn, Mr. Winston Collier (Rolls Royce) and Mr. Peter Coaker (Manager B.P. Educational Liaison Division) who were all members of the B.P. steering committee.

The project was made possible by the schools who co-operated in the work by allowing me entry to their premises often at no little inconvenience to staff and pupils. To list all the particular people involved would run the risk of inadvertant omission so I simply give thanks to all, both staff and pupils. I must however add special thanks to Mr. Bill Device of Atherstone school regarding his support and interest over the whole period of investigation.

Thanks to Mrs. Peggy North for the typing of the manuscript with great accuracy that eased my final writing up of the thesis.

C.J. PURCELL.

SUMMARY

The research undertaken in this thesis was based upon the desire to establish as early as possible in school life those pupils who have the potential to do well in science. Evidence is presented from the literature to suggest that spatial ability is an important indication regarding this and one that is not normally tested within schools. The fact that males show an average superior performance on such tests is discussed and the current available theories for this are explored. A possible view which integrates these presently dissociated theories is given.

As formal thinking and the appearance of the sex difference in performance on spatial tests are both associated with adolescence, the former is discussed and an argument is made to support the view that formal thinking possesses a spatial element.

Based on the evidence linking spatial and scientific abilities a test of the former was developed and after pretesting was used in schools to compare pupil performance in school with their attainment in the test. The results of these testings are given and their implications discussed both regarding indication of scientific ability and also development of spatial ability in adolescence. The latter raises the problem associated with this and similar studies where the length of study prevents a great deal of longitudinal investigation.

The thesis concludes that the developed test is of benefit to the classroom teacher and argues that the 'scientific nature' of such studies is not lost by the inclusion of subjective views.

INTRODUCTION

The content of this thesis was developed in association with the ESTEAM project (Enrichment Science and Technology for Extremely Able and Motivated pupils) which originated in 1968. The ethos and aims of the project have been discussed elsewhere by Screen (1981) and Perryman and Purcell (1983b). One aspect of the scheme was the running of Science Summer Schools in the department of Science Education at Warwick University. This required pupils in the first year of secondary school to be extracted from local schools and hence some form of testing was needed to help in the process of selection.

Many tests have been developed to identify particular abilities at an early age and so facilitate their subsequent nurture. There is evidence to suggest that spatial ability is indicative of ability in science and mathematics. The evidence for this is examined, along with the development of a test and its subsequent use.

Piburn (1980) suggests that there is a link between ability in science, spatial abilities and Piagetian formal thinking. The stages of Piagetian development and sex differences that appear in spatial ability at adolescence is suggestive of some biological influence acting alongside that of experience. Evidence for this is discussed with a view to integrating these ideas further.

The identification of Scientific Talents

The identification of gifted pupils has received much attention, albeit spasmodic and often motivated by national needs as in the U.S.A. (Ziv (1977)). However, relatively little has been done to narrow this identification to a specific talent, or talents, in particular those required in Science with the exception of the development of a test of scientific ability described by Stillman (1982).

Lesser, Davis and Nahemow (1962) gave a test for selection of scientific talent to a group of high ability pupils with a mean IQ of 150 at the age of 6 - 8 stating that it would seem important to identify talent at an early age. The Hunter test so developed contained 91 items and was based on previous work that had yielded encouraging results. The first test included a measure of knowledge of scientific principles and information together with space conceptualisation, vocabulary, number and reasoning ability. The test was correlated with achievement during the 3rd grade (age 9) course, and was found to be a good predictor, though the authors caution that the Hunter test may well only be measuring the same things as the achievement tests, which were themselves specially constructed. This immediately raises one of the problems with prediction on relatively short term studies such as this. How does one define the success which is being predicted? Comparison of results with tests of a similar nature seems self defeating. A further attempt at selection, centred on the secondary age range was based on the "Talent Search" (George 1979). This seeks to select pupils throughout the U.S.A. who will benefit from special provision and study.

Tests of this nature assume that one can measure scientific potential, as a specific ability, distinct from general intelligence. This view is opposed by Brandwein (1959) who suggests that scientific ability is just a development of intelligence in a particular direction governed by external influences. Cooley (1958) states that scientific talent is just a conglomerate of abilities of no particular form. In contrast to this Lesser et al (1962) cite among others Fehr (1953) and Guilford (1950) who subscribes to the view that there is an ability pattern common to scientists. Sharing this view are Super and Bachrach (1957) who list spatial visualisation, manipulative ability, ability to plan a study and to communicate. Although not specifically advocating spatial ability Denton and Postlethwaite (1982) are supportive of subject specific identification in that they found the number of pupils with all round high ability to be very small.

Longitudinal studies by Terman (1947) point to the fact that there are distinctive traits that characterise scientists from other gifted individuals, and although some of the selection techniques have received criticism (summarised by Ziv 1977) there nevertheless appears to be a collection of characteristics associated with scientific ability. Ann Roe (1952) in her study of 64 eminent scientists gave a distinct profile of the group Physical Scientists being characterised by both verbal and spatial abilities. Smith (1964) outlines what one would expect of a scientific type. It should, however, be emphasised that it is always the average type being discussed and not particular cases. It needs to be remembered that the facets of personality and intellect are varied and an individual will rarely conform to the average overall pattern. Human abilities cover a complete spectrum across which they are probably normally distributed, and provided this is not forgotten, analysis of the group can yield interesting trends that can be looked for in individuals.

An attempt to identify scientific ability in adults was made by Mandell and Adams (1948). A battery of 15 tests was given to help select Chemists, Physicists and Engineers. They correlated the results with salary levels and employer ratings of these scientists. Only one test which they called a 'hypothesis test' yielded promising results on all three groups. A surface development test was a moderate predictor of physics and chemistry but no other test spatial or otherwise yielded significant results. (No reference was made to the previous academic achievements of these scientists in the report). This reported use of a battery of tests further highlights the problem of how one estimates success in science and that, unless a long term study such as Terman's is carried through, one cannot be sure whether the possession of innate abilities resulted in the success or whether the course of study fostered and led to the development of these abilities.

From this work however there is a hint that spatial ability may have a part to play in scientific ability. There is also a considerable amount of evidence for this from studies other than those of selection already discussed. Smith (1964) indicates the need for further research in the area emphasising that the full significance of spatial ability may not be apparent until 'A' level and beyond. This could be taken to be supported by studies cited by Maccoby and Jacklin (1975), in which girls were shown to perform better on questions of a verbal nature, whereas boys dominate if the bias of the test is towards mathematical and spatial presentation. It would seem that as science advances, the knowledge of the subject matter is tested more and more by questions of the latter type.

Maccoby and Jacklin write:
'Little is known about the role of spatial imagery in mature complex thought but it now seems possible that the iconic representation does not give way to more advanced forms of thought but rather it is retained and utilised as part of the more advanced levels of information processing.'

Einstein was known to feel that visualisation was important in scientific thinking (French 1981). This view is supported by Arnheim (1970).

Yates & Pidgeon (1957) undertook research to establish a battery of tests to help in 11+ selection. What emerged was a series of tests each having a particular weighting to indicate its relative importance in the battery. The best prognosticator of success in secondary education was a battery containing a test of spatial ability, and they were led to remark that although it may seem surprising to find this in such a battery, the inclusion could be linked with future performance in maths and science. Greater research was needed to confirm or deny this.

Abranaval (1973) investigated the development of spatial ability in pre-secondary school children and found that spatial rotations and transformations were, for many, difficult mental tasks.

She says:
'Even more important may be the encouragement of spatial understanding based on rotations. The ability to find points of reference in science or in a figure, and to imagine the outcome of various rotations seems to us a high order ability that is necessary for much more advanced thinking in mathematics, science and related fields.'

Siemankowski and MacKnight (1971) demonstrated science success correlated with 3-D visualisation. This observation supported by Sheldon, Baker and Tolley (1972) and Baker and Tolley (1974) who found spatial visualisation correlated with achievement in chemistry. Further support is given by Vernon (1972), Peterson (1976), Baumann (1976) and Bishop (1978) who all evidence links between spatial ability and science.

A study by Lewis (1967) also adds to the argument that spatial abilities are important in science. It is of interest not only in that it connects spatial ability with science but also in that it relates this to abilities in GCE 'O' level science. He used items from The Dressel and Nelson file (1958). In some cases these were of modified form. This consists of science questions categorised by subject and the Bloom Taxonomy for use in American schools. Questions from this were taken in Physics, Chemistry and Biology, and classified by teachers of 'O' level science into:

1. Questions of Knowledge.
2. Questions of Comprehension and Application.
3. Questions of Evaluation.

These nine tests were then administered as part of a test battery.

Factorial analysis extracted a primary factor loaded on vocabulary, language usage and arithmetic along with high loadings on the science tests. The evaluation questions yielded lower loadings than the others. The second factor extracted had high loadings on the spatial tests and moderate ones on the evaluation tests with no significant loading on the other science tests. Lewis goes on to say that this factor may be regarded as one of scholastic aptitude and reasoning, the tests of evaluation being less governed by what has been taught on a science course. Thus, apart from supporting the influence of spatial abilities on scientific abilities there is also the inference that as evaluation is a higher level ability, ~~that~~ the full value of spatial abilities may not emerge until later in a science course when such abilities are called on to a greater extent.

But we have been presented the details of the support

It can be seen then that there is a substantial amount of support for the idea that spatial ability is an important factor in scientific ability.

Nor have we had much presented re What is spatial ability?

Definition of spatial ability?

McGee (1979 a) defines spatial ability as:

> 'The ability to manipulate, rotate, twist or invert pictorially presented visual stimuli.'

This is in contrast to the use of the term as applied to the ability to find one's way in the environment (Downs and Stea (1977)) although it is accepted that the two are probably related. Before elaborating on this definition it is relevant to trace the origins of intellectual abilities and in particular the spatial one.

Around the beginning of this century efforts were made to measure intelligence, but initially problems arose because there was no clearly formulated idea of what this quality might be. This situation held until Binet, who had been studying mental growth in school children, was requested by the Paris Authorities to produce a test which would identify pupils with special educational needs, i.e. those who were dull as opposed to those who were lazy. In response to this the Binet-Simon scale was produced and underwent revisions in 1908 and 1911. The ideas presented by this approach were widely adopted and Terman used them to produce the Stanford-Binet Scale published in 1916 and revised in 1937 and 1960. These tests were a mixture of 'aptitudes' graded so that a child could be individually tested and the level of performance relative to others of the same age measured. Each part of the test had levels of difficulty that an average child of a particular age would be expected to achieve. The highest age achieved was noted and related to chronological age to give a definition of IQ - $\frac{\text{mental age}}{\text{chronological age}} \times 100$. Thus a child of 10 with an IQ of 130 would have performed as an average 13 year old on the tests. This method of expressing mental ability is rather crude and has now been superseded by more refined methods. In general it may be said that IQ gives a pointer to a person's intellectual ability at a particular moment in time relative to other members of the population of the same age. Although it has its use as a predictor of academic achievement it says little of overall individual development and its use and limitations in the identification of gifted pupils has been highlighted many times, e.g. Ziv (1977) Dielisle and Renzulli (1982). Binet's approach was to regard intelligence as an overall ability rather than a sum of distinct abilities.

In contrast to this is the view that intelligence can be broken down into specialist abilities. The theoretical background for this arose from studies by Francis Galton in the latter part of the 19th century. He was interested in inherited characteristics and one study was of fathers' heights in relation to sons. When these two variables were plotted a scattergram of the points was obtained. By using fathers' specific heights and the corresponding average sons' heights he was able to plot a line showing the correspondence. The slope of this line was not at 45° to each axis as would be expected from an exact correlation. Instead the slope was less, indicating sons' heights regressed towards the mean. This now only has historical significance being an artefact in the mathematics (Mulaik 1972). The slope of regression line as it was called showed the amount of correlation between the variables and the slope of this line received mathematical expression from Pearson in the form of the product-moment correlation coefficient (r). This is defined in any introductory statistical text, e.g. Guilford and Fruchter (1956).

In 1904 Spearman used this to correlate the results of testing some school children and this gave rise to the study of factorial analysis. On analysis of the correlation matrix he could see that the tests could be arranged in hierarchical order dependant on the amount of correlation with other tests. He interpreted this to mean that the tests were of the form:

$$X = ag + bs.$$

where 'g' was a general factor and 's' an ability specific to the test. a and b were 'loadings' which indicated how much each factor contributed to the total test score X. The higher in the hierarchy a test the greater its correlation with other tests and hence the greater the 'g' factor's contribution. The inference from this is that if tests for pure 'g' and pure 's' existed, a person's score on each of these could be substituted in the equation to find the score expected on test X. The factors are considered additive.

The Spearman model assumes no correlation between specific abilities. Thurstone (1947) went further in that he took the view that 'g' did not exist, and stated that the mind could be considered as a series of specific factors. These factors, however, always tended to show a positive correlation, and so it appeared that Thurstone's analysis was only masking 'g'.

This is but a brief outline of the development of factorial analysis, and a full description of the methods used can be found in Mulaik (1972) and Thurstone (1947). The method of factor analysis has come in for criticism, e.g. Skemp (1979) who points out amongst other things that abilities are not simply additive. It should be added that in the illustrative example he does not use PRIMARY abilities.

Thomson (1939) puts forward a sampling theory which rather than considering the mind to consist of additive parts assumes it to be a mass of relatively undifferentiated parts. Performances on tests correlate because they have been sampling overlapping parts.

The model which has emerged though is one of separate factors which go together to make a whole. How these relate to each other depends on the specific model adopted. Vernon (1950) proposes a hierarchical model where 'g' sits at the top and other factors branch out below this - the greater the branching the less significant the factor. As an alternative the Guilford model (1967, 1969), is a three dimensional arrangement of 120 cells each of which is seen to represent a distinct intellectual ability. This model was used by Getzels and Jackson (1962) who took the convergent-divergent cells of the model and applied these to the identification of the creatively gifted. Such models should not be considered to have real existence, they merely provide a stimulus for further research.

A further model (Davis 1964) proposes 15 factors and is of interest here as it splits the intellect into three main factors, numerical, verbal and spatial.

	Perceptual Speed and Accuracy	Memory	Reasoning	Fluency	Visualisation
Numerical	PN	MN	RN	IN	VN
Verbal	PV	MV	RV	IV	VV
Spatial	PS	MS	RS	IS	VS

The origins of the spatial factor go back to the development of tests of mechanical aptitude, and their evolution is traced by McGee (1979 b).

Acceptance of such models enabled test batteries such as that of NFER (1981) to be developed as a means of predicting future promise rather than diagnosing present achievement. Similar batteries enable a profile of abilities to be drawn up and advice given on future employment by comparison with profiles known to correspond with success in the job. This acceptance also makes it possible to regard scientific ability as having a profile of certain abilities, one of which is spatial.

In his review McGee claims that two spatial factors have now been isolated, those of spatial orientation and spatial visualisation. (The latter being the more difficult). Even this division is disputed by Eliot (1980) who claims that the degree of refinement in tests is not sufficient to establish this. At present factorial studies provide different loadings on the same test depending on the method of administration and the population sampled. He does, however, go on to account for 12 different types of spatial test classified by inspection of 300 tests that will go to make up a comprehensive handbook of spatial tests (Eliot and Smith 1982).

The classification provides a more detailed definition of spatial tests as follows:

1. <u>Matching tasks</u>

    i) Maze copying

    ii) Embedded figures test

    iii) Figural drawing - drawing reproduced from memory

    iv) Figural combination - Tangrams

    v) Figural rotation.

2. <u>Manipulative tasks</u>

    vi) Blocks - estimate number from drawing

    vii) Object rotation

    viii) Paper folding

    ix) Surface development

    x) Perception tasks - seeing objects from different points of view

    xi) Combination tasks - two or more tasks in each item

    xii) Figural collage - mixed items to respond to.

Camp (1981) goes a step further with a definition similar to that of McGee but with the added rider that there should be a difference in performance between the sexes in favour of males. This point will be developed further as it is relevant to the use of these tests as indicators of ability in science.

Sex differences on spatial tests

It is an attractive explanation to use the difference in spatial ability as being a causal factor in the drift of girls away from science studies during adolescence. This view has been opposed on many fronts, e.g. (Hows 1980). One of the main arguments against, being that the noted difference is not large enough. Although some studies such as Abranaval (1973) claim to find a difference in performance at a young age, the general consensus in the literature is that the evidence points to a male superiority in such tests which does not appear until puberty. Maccoby and Jacklin (1975), McGee (1979a, 1979b) and Lloyd and Archer (1976) all discuss this in some detail. If the difference is manifest at puberty, it lends weight to there being some biological reason for this rather than one of an environmental nature. Differential experience of the environment is often used as an explanation of this sex difference.

If it were environmentally determined it seems unlikely that it would suddenly appear. Although inferring human behaviour from animal studies is a dangerous practice if taken too far, it is of interest to note that the laboratory rat shows a male superiority in spatial tasks (mazes), Buffray and Gray (1972).

The question arises that if the sex difference is environmental,can it be removed by training? The results seem to be inconclusive. Goldstein and Chance (1969) gave practice on tests of embedded figures and noted that in adults only females showed improvement. They explained this as being due to motivation and encouragement by examples. Boys already performing at a higher level would have less room for improvement. Contrary to this, Smith and Litman (1979) using Tangrams (two dimensional jigsaw on paper) found that with 11 - 13 year olds the girls showed superiority and found instruction only brought about improvement in the older boys. They were led to suggest that this could have a bearing on what age instruction is given. The fact that the girls were found to perform better is explained on the grounds that the Tangrams are probably testing a different spatial ability from many other tests. It could be suggested that such tests favour girls as they are more akin to dress patterns, but this would be expressing a cynical view of the environmental lobby. Why? In reviewing the same area Camp (1981) concludes 'Why does doll manipulation have more effect than trucks, and why does the ability difference only occur at 12?' This is a question that awaits a satisfactory answer.

Attempts have been made to explain this difference in performance in three ways:

1. It is genetic
2. It has its roots in hormonal differences
3. The difference is neurological, though both neurological and physiological differences will have a genetic component.

These will now be considered as they are relevant to the understanding and implication of the use of spatial tests especially with reference to 3.

1. A possible genetic explanation

Burt's fictitious data produced to support inheritance of IQ from identical twin studies was a set-back to genetic studies but there now appears to be considerable valid evidence to support Burt's conclusion that intelligence has an inheritable component. There are still arguments for and against, e.g. Eysenck and Kamin (1981) but studies such as those cited by Tyler (1963) have led to the view that inheritance is more important than environment. Gourlay (1979) concludes that 75% is inherited. The exact figure, even if one was possible, is not important to the present argument. If then, there is a certain component of inheritance in intelligence, as spatial ability is a factor of the intellect, it would seem reasonable to suppose some degree of inheritance in spatial abilities. However, there is the problem of sex difference appearing at adolescence. There are many characteristics which show sex differences in the population as a whole, e.g. colour blindness is more prevalent in males (Taylor 1962). This can be explained in terms of the linkage being on the X chromosome. In the body there are 23 paired chromosomes. Each carry alleles which govern certain characteristics. These are paired and may be dominant so that this trait will always be expressed or recessive, where expression only finds itself when it is not paired with a dominant allele (gene). In the female the sex chromosomes are similar and are known as X chromosomes. They differ only in the alleles they carry. In the male this is not the case - the sex chromosomes consist of an X and a Y only the former carrying alleles. When female eggs are produced each normally contains one of the X chromosomes and the male sperm may be X or Y. On fertilization several combinations are possible. So for example, if normal colour vision C is dominant and resides on the X chromosome possible combinations are:

X^cX^c	normal female
X^cY	normal male
X^bY	colour blind male
X^cX^b	female carrier normal vision
X^bX^b	colour blind female.

b = recessive gene for colour blindness.

Analysis of the frequency of these combinations in the total population gives a greater occurrence of the X^bY rather than X^bX^b.

McGee (1979b) reviews the evidence for an explanation of spatial ability sex differences in terms of an X linked recessive gene and says that the evidence based on the expected correlations in ability between generations is inconclusive. An immediate problem in such studies would seem to be due to the trait being measured. It is not as definite as black hair or blue eyes and it has to be corrected for performance relative to age. Studies by Defries et al (1976) and Fralley, Eliot and Drayton (1978) give little support to the thesis although Stafford (1961) Hartlage (1970) and Bock and Kolakowski (1973) do. It would appear that the explanation that was initially seen as simplistic is now realised to be more complex, if indeed any such explanation is possible.

Which one?

Traits such as colour blindness would be apparent from birth and the appearance of the differences at adolescence is therefore an extra problem. Goldberg and Meredith (1975) sought to perform a longitudinal study by using results from a younger age on Piaget tests claimed to be spatial and correlating these with spatial tasks. They were hoping to show spatial abilities to remain static in rank order in the population, i.e. sex difference at early age. This was not apparent. They did claim the ability appeared to remain steady but conclusions appear unfounded. There are, however, other ways in which sex linked inheritance is controlled.

The trait could be sex influenced, not being carried by the X chromosome, but instead only being dominant in a particular sex, e.g. baldness is dominant in males but recessive in females. Alternatively, it could be sex linked in that the trait is only expressed in the presence of certain hormones. Credence is added to this by the fact that women with Turner's syndrome having the genetic make-up XO are deficient in visio-spatial tasks (Money and Alexander 1966) which could indicate a necessity of male hormone for expression of the ability.

2. Hormonal influence

This leads into the second explanation of sex difference (i.e. hormonal) although it will already be apparent that the first, second and third enumerations may not be alternatives, but may require integration to produce a satisfactory explanation. The main problem with simple sex linkage seems to lie in a matter of degree. The ability to do spatial tests does not suddenly appear in one sex, learning? it is a differentiation in ability which appears, and so this suggests that it cannot simply be a genetic expression. There must be some mediating factor producing a spectrum of ability and this could be hormonal. This influence can be seen in secondary sexual characteristics, where men are genetically male because of XY make up, but some are more masculine than others due to the level of testosterone.

First it needs emphasising that there is little evidence for causal relationships between hormones and visiospatial ability but the changes in both of these around puberty is suggestive of a link. The original suggestion of hormonal control is usually attributed to Broverman et al (1968) who considered the effects of various chemicals on the sympathetic and parasympathetic nervous systems. Stimulation of the former is associated with activation and the latter with relaxation and suppression of arousal. The two systems then work together and it is their integrated effect that controls the state of the organism. which, state,

In noting that women are better on verbal tests and repetitively learned tasks whereas males have greater visio-spatial ability and greater control over closure on a task, they suggest that the female behaviour is related to activity, i.e. arousal of the sympathetic system or suppression of the parasympathetic system. Tasks of field independance introduced by Witkin (1962) and showing superior performance by males require the opposite type of reaction, i.e. one which allows for delayed closure on the task. Vernon (1972) reviews this field of study and while he does not say they are the same, concludes there is a strong correlation between spatial tasks and field independance. The use of the embedded figure task (EFT) for determining field independance suggest that the two may be synonymous. McGee (1979a) concurs with this view to the extent that sex differences noted in many tests can be traced to an underlying spatial component. This is true of field independance.

meaning?

not clear

The Broverman explanation related the sex hormones to this observed difference between the sexes. The argument was criticised by Singer and Montgomery (1969) on the grounds that the chemical data used was incorrect (replied to by Broverman et al (1969)) and Archer (1976) who claimed the sex differences [illegible]ted were only those that fitted the evidence. The theory does, however, provide a model on which to build. Stated briefly the conclusion from the available literature is this.

details

Norepinephrine production which amongst other chemicals in the body excites the sympathetic system (causing arousal) is inhibited, to a large extent, by the level of Monoamine Oxidase (MAO) which in turn is controlled by the sex hormones. They inhibit the release of MAO which facilitates Nonepinephrine production which in turn causes activation. The female hormone oestrogen is found to be a stronger inhibitor of MAO than androgen which would account for higher arousal in females. Along with this acetycholine production which is known to stimulate the parasympathetic system and so cause suppression is inhibited by oestrogen but not by androgen. The combination of these two effects accounts for greater arousal in the female.

Before proceeding it is noteworthy that Eysenck's view that personality must be taken into account when a test is given (Eysenck 1967) could be relevant to the present argument. His view is that there are two facets of the personality which have a bearing on performance, namely

1) Neuroticism - emotionality
2) Extraversion and introversion.

The polar extremes of neuroticism are lability (high arousal) and stability. The level of arousal is linked with drive according to Spence (1956) cited by Eysenck (1967). This in turn can be linked to the Yerkes-Dodson Law which relates performance to drive not in a linear fashion but as an inverted horse shoe. The point of inflection is the amount of drive needed for maximum performance on a task. Further, the degree of drive to give maximum performance is related to the nature of the task. Easy repetitive tasks have high drive for optimum performance whereas more difficult tasks require a lower drive. The point a person occupies on a curve for a particular task depends on their drive or arousal level. They may be at the point of inflection where greater arousal lowers performance or may require arousal to gain this state of maximum performance. Those high in activation then would be expected, on tasks requiring control of closure (especially under pressure) to perform less well than those of low activation. This corresponds to Broverman's females and males related to hormonal levels. This is further supported by the findings of Munkelt (1965) cited by Eysenck (1967) which showed significant differences between men and women, the former displaying stable characteristics and the latter labile. Women as a group tend therefore towards greater neuroticism than males and being more aroused would be expected to perform less well on tasks requiring suppression of closure especially if these tasks were performed under pressure.

Returning to Broverman's theory there are a number of studies which relate to this. Peterson (1976) studied a mixed group for the relation between hormonal levels and performance on visio-spatial tasks. She found that the worst performers on spatial tasks were those with high levels of oestrogen or androgen (as adjudged indirectly from bodily characteristics) i.e. those at the poles of the sexual spectrum. Those closer to spectrum centre were found to perform best on such tests. Now on Broverman's theory both polarities of the sexes would be expected to have high hormonal levels and high arousal, females on average higher than males. Although Peterson does not express this view her findings are consistant with the theory.

Show this - evidence

A further interesting study is that of Waber (1976, 1977, 1980) who observed that late developers (as measured by time of puberty) were better on visiospatial tasks than early developers. No difference was noted on verbal tasks. If one adds to this the expectation that early developers would have a high level of sex hormones (Tanner 1962) and so be at the extremes of the sexual spectrum then Peterson's studies could be fitted to these findings. Waber also found that later developers were more lateralised. This will be referred to later.

Why so? How do you know

meaning?

A problem with these and similar studies related to puberty is that hormone levels are indirectly inferred from physical characteristics, time of menarche or urine samples. Until such measures can be taken directly in a longitudinal study the results will always be open to dispute. Katchadourian (1977) says that such methods of measurement are now available. Archer (1976) gives a review of hormonal studies and concludes that they do not, at present, give an adequate explanation of sex differences in visiospatial ability. It is argued here that a pattern does emerge and that this is further supported by Hier and Crowley (1982).

Their study relates to a group of males who did not undergo a normal puberty due to a lack of male hormonal production at this time. The performance of this group with those of a control on tasks of visiospatial ability was found to be poor. The use of hormone replacement therapy did not appear to produce any improvement in this subsequent to puberty. These findings are suggestive of the influence of male hormones on the development of spatial ability prior to and at puberty. They suggest that a minimum level of these hormones is needed to enhance the development of visiospatial skills. These observations appear to fit Peterson's findings as the more masculine females would be expected to have sufficient levels of male hormone. Waber's findings indicate that it is not only this minimal level that is needed but also the longer associated growth period.

Could be other factors need to know much more about the individuals involved

3. Brain lateralisation and the sex difference

What?

This is an area which is becoming of increasing interest not just as an explanation of sex differences in spatial ability but also in other intellectual functions. The term lateralisation of the brain stems from the fact that the brain is divided into two, the left and right hemispheres. These halves are cross linked by a large bundle of nerve fibres known as the corpus callosum. It is known that the transfer of information between hemispheres is facilitated by this nerve bundle and that the function of each hemisphere seems to be essentially different.

The realisation of this followed operations, performed on patients with severe epileptic seizures, by Sperry, Gazzinaga and Levy in the mid-sixties (see Wittrock (1977) for a more detailed account). In order to localise the epilepsy the corpus callosum was cut. It was subsequently found that although these patients appeared normal that the now isolated hemispheres handled data differentially.

As the nervous system is cross linked the left hemisphere controlling the right hand side of the body and the right the left hand side, information placed in the left field of view would only be 'seen' by the right hemisphere. Making use of this it was found that only when objects were presented to the left hemisphere from the right hand field of view could they be named. In the right hemisphere they appeared to go unnoticed. If, however, an object was presented to the right hemisphere it could be selected from a group of objects presented to the left hand but could not be named.

It was concluded from this that the left hemisphere is the verbal centre, can handle difficult calculations of a serial nature and reasons logically using verbal ideas. In contrast the right hand side handles spatial tasks and musical appreciation. This is a brief summary of the overall findings, which are reviewed by Blakeslee (1980) and Springer and Deutsch (1981). The left hemisphere can be said to be adept in abstraction and serial data handling where the right is holistic in its view. It should be noted that this is the normal, about 1/3 of left handers (5% of the population) show reversed lateralisation. It has been suggested that the right recognises faces by overall view (Gestalt) where the left will do so by particular features (Downs and Stea 1977) although Bradshaw and Wallace (1971) claim some amount of serial processing in the right hemisphere based on face recognition. Cohen (1973) evidences some parallel processing of information in the right hemisphere based on reaction time studies. Clearly from this, the simplistic picture does not represent the true one. McGee (1979b) concludes that the evidence to support the general differences between hemispheres is overwhelming.

Having a left verbal hemisphere and a right spatial one provides an attractive vehicle to account for observed sex differences on tasks related to these abilities.

Generally lateralisation is not complete, some functions being shared by both sides. Females show less lateralisation than males or as in Waber's studies early maturers show less lateralisation than late. Blakeslee (1980) says that lateralisation has a marked effect on the person as greater laterality gives two specialists working together, whereas shared laterality produces two 'Jacks of All Trades'. No claim is made that one is better than the other, they are just different. Levy (1976) makes use of this argument to account for differential verbal and spatial abilities between the sexes. It is suggested that the hemispheres must cross talk down connecting nerves. If lateralisation of the verbal ability is weak more connections are needed to allow for cooperation. This impairs functioning on tasks that are visio-spatial but the shared ability aids verbal performance.

Corballis & Morgan (1978) seek to explain differences in lateralisation by a left right development gradient, the left (language) side leading the right. Hence with late maturers there is a greater degree of lateralisation due to the longer growing period. That this is not the detailed story is accepted and there is a suggestion of critical periods of maturation when various differential functions are manifest.

The explanation of sex difference in spatial ability by lateralisation is not yet apparent although sex differences in the two are. What is missing is the causal link if one exists.

Character traits

Observations related to character already mentioned in Roe's study of eminent scientists finds corroboration from other sources and these can be related to time of maturation. Smith (1964) devotes a considerable amount of discussion to the relationship between spatial ability and personal characteristics. What emerges is that scientists tend to be schizothymic (introverted with a tendency to tall thin stature) a trait that is also linked with high spatial ability. Eysenck also gives some evidence to suggest that those who show field independance tend to display intraversion. This is supported by Sheldon cited by Harrison et al (1977) who related body build to character. They also observe that age linked examinations favour early maturers as physical and mental maturity can be linked. They suggest also that early maturity has an effect on personal mental outlook. The early maturers tend to be more muscular than late, who tend to tall build and introversion. They say that in America early maturers tend to be more successful although what is meant by success is undefined. In conclusion it is added that the relation of temperament to body build may be fantasy, but it seems more likely to spring from the firm ground of human biology.

Kohen-Raz (1977) cites Broverman et al (1964) and quotes from this that 'Individuals with weak autonomizers (less masculine individuals with inhibited motor skills in whom stimulation of the autonomic or sympathetic nervous system is weak) mature late and are typically tall with schizothymic body build.' Further related to this Head (1980) using the Loevinger questionnaire related to ego development (for a summary see Hauser (1976)) as a measure of maturity, says that there is a clear sex difference at 12 to 14 years with girls being more mature but by 17 years this had disappeared. With boys at 14 years those expressing interest in the humanities are more mature than those opting for science ($p < 0.05$). Head (1982) affirms this view suggesting the different personalities and hence view of the world during adolescence will influence between a subject based on 'factual' evidence (i.e. Science) and one which is more open ended and questioning (at this age in school) i.e. The Humanities.

Also related to maturity are the cross cultural studies, the most frequently quoted, in relation to visio-spatial skills, being those of Berry (1976). With Eskimoes and Zambians no sex differences are noted and this is taken as supporting the environmentalists' case, in these cultures women getting similar experience to men. This can be approached from another angle consistent with the pattern already established. If one examines time of menarche (Harrison et al 1977) in relation to European countries where 12.8 - 13.2 years is the average, Nigerian and Eskimo girls show later maturation with ages of 14.3 and 14.4 years respectively. This is in keeping with late maturation and spatial ability. How these times relate to male puberty would be of interest if available. A further note of caution should be added regarding cross cultural studies. It should not be overlooked that these different cultures have different gene pools that have been established over many generations. Because of this to claim that the difference in performance of these cultures is due to environmental effects alone is not valid. Always predisposition must be considered and then the ensuing environmental effects.

Some further observations

Regarding the evidence for the role of sex hormones on visio-spatial ability it seems that little notice is taken of the cyclic nature of female hormone levels which fluctuate far more than the males. Messant (1976) also notes the effects of progesterone which is the other dominant hormone in the female cycle. This appears to oppose the effects of oestrogen when present in considerable concentration. A sudden drop in the level of this occurs premenstrually and one could surmise that with inhibition by progesterone removed the oestrogen would have greater effect on the sympathetic nervous system causing greater activation and poorer performance on activities requiring inhibition or those involving stress. There is evidence to support this.

Dalton (1960) found that classwork performance dropped markedly in the week preceding menstruation and Dalton (1968) on looking at girls aged 14 - 17 years showed a drop in pass rate of 5% of girls taking examinations during menstruation.

In 16 - 19 year olds the difference in pass rate between those taking the examination during menstruation and those in mid cycle was 13% in favour of the latter. She calls for the time of papers in multiple paper examinations to be split to take account of this. She says, 'The more this knowledge is appreciated the more sympathetically we can regard the vagaries of womanhood with the inequalities thrust upon her by the hormone variations in the menstrual cycle.'

Redgrove (1971) examined cyclic variations in performance and concluded that a difference was only noted when women were asked to work at peak capacity. She concludes 'that only jobs that present considerable tactile, visual, auditory or other similar demands do these differences become apparent.'

If, then, the cycle has an effect, as this limited evidence suggests, when a sample of post adolescent females is taken there will be (unless controlled) a complete range of position in the female cycle. This could account for some of the observed differences on spatial tests. The suggestion is that males and females could perform equally well on spatial tests if all were performing in an optimum manner. Because of the cycle all females will not be so performing and this lowers the mean score.

Can these theories be integrated?

Evidence has been presented which indicates that males as a group are superior to females on visio-spatial tasks. They are also more lateralised. The reasons for this are not resolved. There is also reason to associate these observations with ability in science.

It is noted here that many reviews of the literature give evidence for sex differences in spatial ability and then list the explanations for it as being inherited characteristics, hormones and laterality. There is, it would seem, an alternative view. If the spatial ability and laterality are considered as sex differences, one in performance and one in physiology then for each a genetic or hormonal explanation (or a combination of these two) can be sought, a causal link between laterality and spatial ability being largely speculative.

Support is given to this view in a recent paper by Newcombe and Bandura (1983) which confirmed that, as Waber had found, late maturers (in this case using a sample of girls ranging from 10 - 12 years in age) were superior in tasks of spatial ability. It also showed that early maturers were more feminine and participated less in activities that were considered to enhance spatial development. Amongst this group masculine traits were correlated with ability on spatial tests. These findings are consistent with the pattern established and it is concluded that the results can be explained in two ways owing to the fact that laterality was not found to correlate with maturation at this time. One explanation is that made above i.e. there is no causal connection between laterality and ability on spatial tests.

The second is due to the fact that no correlation was found between laterality and rate of development during this study but had been observed by Waber in an older group of 14 - 15 year olds. This led to the suggestion that there is an actual change in the patterns of lateralisation at puberty as proposed by Waber (1977).

One other point which has emerged from the foregoing discussion seems worthy of recognition at this point. It is conceded that the difference in ability between males and females is small and on the face of it not sufficient to account for the discrepancy of male-female study of science in school. However, the male-female designation can be viewed not as a sharp division but in terms of a complete spectrum of sex. Within this spectrum those high in spatial ability will be the late maturers who as Peterson found lie towards the centre of the sexual spectrum i.e. 'masculine-females' and 'feminine-males'. Further support is given to Peterson's findings by Berenbaum and Resnick (1982). If science is perceived as a masculine pursuit as discussed by Bamber et al (1983) then it can be argued that it will attract boys who see it as an enhancement of their masculinity and it will deter girls who wish to affirm their feminity. Bamber et al (1983) discuss the fact that girls studying science in the sixth form tend to hide their femininity. This explanation would seem to be acceptable to both the inherited characteristics and environmental views of the uptake of science by girls as it is suggestive of a predisposition which is probably fuelled by the present system which tends to present science as a masculine domain. It would also fit in with the observation that in single sex schools the uptake of physical science* is greater than in coeducational schools.

The implication is that in the former environment there is not the same pressure on masculine females to affirm their stereotype.

* It should be noted that when science is referred to here it is physical science, the uptake of biological science by girls in school being high.

Formal Thinking

The definition of formal thinking stems from Piaget's developmental model which will be considered in some detail in order to relate it to the previous argument. It should not be thought that this model is the only one and alternative views are briefly summarised. A fuller account of these may be found in Salkind (1981), Crain (1980) and Langer (1969) amongst others. As is often the case in such fields of study there are no absolutes and all possibilities of development are covered. There are three ways in which development might occur.

What?

1) The mind is blank at birth and all that is learnt is due to the environment. This takes no account of the individual's part in the matter, being a behaviourist view exemplified by Skinner. All action can be explained in terms of stimulus and ensuing conditioned response in the organism.
2) At the other end of the spectrum is the view held by Gessell that the mind develops independently of the environment following a genetically determined path.
3) The development of the mind is caused by an interplay between it and the environment, neither being dominant. This would represent a mid point view.

Not clear.

A more detailed account of the theories fitting into this spectrum is given by Bunge (1980).

Not appropriate for a PhD

Of these theories the most familiar and influential on educational thought is the developmental theory of Jean Piaget.

Where is Piaget's posn in 1–3 above

This has so much dominated interest in recent years that reviewers of the literature in Science Education (1979, 1980) have been moved to remark that the excessive amount of work in this area was unhealthy for general progress on the question of intellectual development. It is very much an example of a paradigm stifling alternative approaches to a problem. Alternatives to the Piaget model such as that of Ausubel (West and Fersham (1974) review this in relation to science) deserve more attention. The present account will be guilty of this bias but claims justification for this as the focus is on formal thinking implicit with Piaget's theories.

weak

Skemp (1979) makes the point that in a short space it is not possible to do any justice to the theory, and indeed this will not be attempted. The facets which seem relevant to the present argument will be extracted. Many books are available that will give a wider introductory view, e.g. Boyle (1969), Beard (1969) and Furth (1970). A comprehensive overview is provided by Flavell (1963).

Two further points relating to Piaget's work are that definitions of meaning are not easy, as he was not always consistent in word usage, together with the fact that it has been translated from the French suggests that some meaning is inevitably lost. The combination of the two makes interpretation difficult. Because the studies were over a long period there was an inevitable shift in thinking, so for example a greater recognition of genetic influences was made in his later work

weak

as yet untranslated (Vuyk (1981)).

In her books 'Piaget's Genetic Epistemology 1965 - 1980' Vuyk emphasises the necessity of having a knowledge of the background in which the ideas were developed. Such a knowledge can be gained from Rotman (1971) and Boden (1979). One of the chief influences in moulding his ideas was the philosophy of Bergson (Lindsay (1911)). Another, his interest in biology. The influence of the former can be seen in 'Insights and Illusions in Philosophy' (Piaget 1972). In this he argues that there needs to be a merger between the scientific method and philosophy. Only by use of the former can thinking be placed on firm ground and true knowledge realised otherwise it remains metaphysical.

Piaget's work has received criticism in that it is unscientific, the theories not being presented in a way that could lead to falsification. This on first sight seems ironic in view of what has just been said regarding philosophical method and remains so if one adopts Popper's theory of scientific advancement. This theory, however, has one major drawback and that is that it is unable to take account of time, because implicit in the ability to falsify is the reproducibility of conditions, so on this basis theories such as that of Darwin do not qualify as scientific because they occurred through time and cannot be reproduced in order to falsify; they are metaphysical.

It was just this problem that Bergson sought to solve by the realisation that biological systems could not be studied in the same way as inanimate ones because they were always changing in time. Kantian philosophy claimed that the only real knowledge was that supplied by the mathematical sciences. Bergson opposed this view on the grounds that in order for it to function, mathematical reasoning must discount individual characteristics and it does this by abstraction. He saw this as removing the biological sciences from the field of science. If scientific method is confined to those characteristics that repeat themselves, then individuality is ignored and results will only apply to the mass. This manifests itself in the use of statistical methods in psychology.

Bergson or Kant?

To overcome this, he suggested that a second form of knowledge was possible, and this was intuitive knowledge, a knowledge that moved in time,did not depend on symbolism to represent it and is necessary in positive science. He saw the theory of life and the theory of knowledge as being inseparable, one pushing the other on indefinitely by a circular process.

Seen in this light, the establishment of Piaget's theory by use of the clinical technique which involves interviewing an individual takes on new meaning and also becomes scientific. Piaget classed himself as a Genetic Epistemologist and was seeking to study the development of knowledge from man's origins to the present day. In studying the development of thought in the individual he was seeking to find this mirrored in the development of man's thinking through history. This may be compared with the idea that as a human embryo develops in the uterus it mirrors the evolutionary steps that man has gone through. This was the background that shaped this theory.

Stated briefly, Piaget sees man the organism having to develop in an existing world which he has to cognise and adapt to. The better this is achieved the better the process of life is being coped with. Initially there are a few genetic instincts which set this reaction off. It departs from the view that the behaviourists take of reactions being purely reflexive by ascribing an innate drive and an ability to restructure experience, to the organism. As new situations arrive ways are found of coming to terms with them. Any new experience is assimilated and schemas (internalisations of this adaption) are constructed. New equilibria are attained as the organism keeps overcoming new problems in the environment by the gradual change of schemas and their incorporation in the developing hierarchy. This process of adaptation is called accommodation. Assimilation and Accommodation are intrinsically tied to create this equilibrium with the environment. As schemas cannot undergo dramatic changes development is slow. It is because of this rather than biological restraints due to maturation that an accelerated development is impossible. Thus the organism's development is due to a dynamic inter-action with the environment causing a gradual evolution. This is not Darwinian in nature but distinctly Lamarkian where characteristics gained in life are passed on to the offspring, i.e. schemas are constantly being modified by experience and so the 'inherent characteristics' of a schema are passed on from the more advanced schema by interaction with the environment. Thus the organism constructs a model of 'reality' within itself.

This slow development of the organism was seen to occur in distinct stages (Piaget 1950) and may be summarised as follows:

1) Sensori - motor intelligence	Birth - 1½ to 2 years	
2) Pre-conceptual thought	2 - 4 years)	) preoperational
3) Intuitive thought	4 - 7 years)	thought
4) Concrete operations	7 - 11 years - adolescence	
5) Formal operations	11 - 12 years	

The last three stages may be broadly seen as a movement from egocentric behaviour when the child cannot see things from another's point of view to one of detachment where all points of view can be appreciated. To this is added the ability to handle data. At first only one thing at a time can be considered so that in the intuitive stage conservation is impossible. If water is poured from a cylindrical beaker into a shorter wider one, the child will take the view that the volume has changed. This is explained by Lunzer (1968) as being due to the fact that the simultaneous comprehension of these two variables, height and width is not possible. One is focussed on to the exclusion of the other.

In the concrete stage this is not so, conservation becomes possible and egocentric talk disappears, although Vygotsky (1962) adopts a different view of this aspect. He sees this conversation with the self as being internalised, possibly due to social pressure, to become thought. Reversibility of action is now possible but only with concrete reality. The ability to move thinking back and forth in time is not yet apparent.

In the formal stage this attachment to the concrete present is no longer apparent. Hypotheses can be made on what might or might not happen. As Peel (1967) says:-

> ' Inhelder and Piaget insist that we have in this tendency of the adolescent to formulate and test hypotheses the real explanation of much of his idealism, theorising and criticism of parent and adult values. He is formulating hypotheses about religion, society and fundamental values and in putting these to the test he finds much that to him is deficient in parents, teachers and other adults. So he sees himself in the Messianic role and sometimes becomes the dreamy difficult person so often encountered in school and at home.'

In problems involving many variables, isolation of particular ones is made possible by holding the others constant. A feeling is developed for proportionality and hence proportional reasoning which is seen as a most important ability, particularly in science (Karplus et al 1977) becomes possible, meaning that combinations and permutations can be handled.

Piaget is a constructionist and the stage of formal thought is reached when a sufficiently complex model of the world is internalised and can be acted upon internally so that transformations can be made in the absence of concrete reality. It is not until the internalisation is complete that a person can explain and talk about transformations being made. In the concrete stage transformations are possible but only with external objects the representation of which has not yet become fully internalised.

As discussed by Rotman, Flavell and Brown and Deforges (1979) Piaget's theory has a number of problems associated with it due to semantic difficulties. Donaldson (1978) criticises Piaget for paying lip service to the importance of this in the clinical method and ignoring it in practice. There is also the lack of ability to falsify which has been discussed and the formulation and actual existence of the 16 binary functions linked with INRC. (Details of the latter are discussed by Flavell (1963)). In spite of this what has emerged is a work of monumental proportions that has served as impetus for a great deal of research. Lastly there is the question of proposed development stages which has been subjected to criticism, Brainerd (1978) who opposes stages and the ensuing critique both for and against discusses this in more detail.

Concept of stages needs that examined for many [illegible]

Not part of Piaget's theory anyway

in what?

The stages - do they actually exist?

The question of stages is one that needs closer consideration related to spatial abilities because sex difference in these and the appearance of the stage of formal operations both coincide with adolescence. The existence of stages with well defined boundaries is not essential to the present thesis, simply a transition to formal thinking, the acquisition of which appears to be linked in some way with puberty. Piaget's view of stages changed as his ideas developed. As has been mentioned he did not originally favour biological controls but attributed slow development to the slow change of schemas. Further formal thinking was not necessarily related to adolescence, rather to the attainment of adulthood in the post school world of work, i.e. formal thinking and puberty were not linked (p. 337 Growth of Logical Thinking, Inhelder and Piaget (1958)). In the same volume, (p. 243) however, mention is made of maturation of the nervous system being one of the components that contributes to structure formation along with experience and social influences. According to Vuyk his view of definite stages and lack of Physiology in their development is changed to a less pedantic view of stages and an admission that biological maturation plays a part. Related to the spectrum of development theories Piaget has moved along the continuum from 1 to 2.

A modified view of stages is taken by the Neo-Piagetian theory of Pascal Leone (1979), the theory of constructive operators (TCO). Before an output can be made these schemas have to be acted upon by various 'silent operators' which add weight to particular schemas making their use more probable until eventually the most suitable emerges for use. To account for novel performance there is a mechanism which can over-rule a selected scheme if it will not produce the desired effect (presumably it could then undergo modification?). As already mentioned the embedded figures test requires suppression of immediate reaction. In the paper cited the theory is related to EFT (Embedded Figures Test in which a simple figure has to be picked from a more complex background).

To account for learning and development an M factor is proposed. This increases in seven steps known as M levels and is related to Piaget's stages. The total M energy available is seen as being used to restructure schemas and generally reacting to novel situations. This idea relates to information theory proposed by Miller (1956) where intellect is treated as a black box. The amount of information put into the box at a given time can be increased until greater input causes no further change in output. The box is then seen to be handling information at maximum capacity. Thus the M levels are in keeping with stage development seen as the increasing ability of being able to handle variables. It has been suggested that tests of formal thinking are distinguishable in degree of difficulty by the number of variables requiring simultaneous consideration (Levine and Linn (1977)).

Shayer (1970, 1972, 1976) has used the existence of stages as an argument for restricting the age at which pupils are introduced to certain concepts in science. The CSMS (Concepts in Secondary Mathematics and Science) Science Reasoning Tasks Shayer et al (1981) based on Inhelder tasks (1958) have been developed as a means of assessing the stage a pupil has reached. A similar view of stages related to what is taught is taken by Renner and Lawson (1975) and Renner (1976). Brown and Deforges (1977) however, opposed such a definite approach to stage attainment.

If these stages do exist and in spite of opposition there is considerable evidence to suggest that they do, are they actually imposed by some underlying biology? Harrison et al states:

> ' The stages of mental functioning described by Piaget and others have many of the characteristics of developing brain or body structures and the emergence of one stage after another is very likely dependant on i.e. limited by progressive maturation and organisation of the cortex.'

Young (1978) concurs with this.
A similar view is taken by Van der Vlugt (1979) who describes a three level model of the brain functions based on evolutionary development. The first is the prereptilian which handles unconditioned reflexes, the second paleomammalian which controls conditioned reflexes and the third neomammalian which controls the higher levels of functioning. This is used to explain Piagetian development in that at each successive stage the higher level incorporates and controls the lower one.

Which view is correct, do the stages have some existence due to biological maturation or not? If there were no underlying controls one would expect acceleration through the stages to be possible, especially in the case of more able pupils. Is there any evidence for this?

Lovell and Shields (1967) and Webb (1974) report no precocious advancement towards formal thinking with high ability groups, although they outperformed fellow pupils in the concrete stages. The former make the point that with mathematics a group of 140+ IQ performed well but lacked the ability to deal with proportions.

Another study that is supportive of underlying biology is that of Epstein (1979) who relates stages to the rate of brain growth. Graphs of rate of brain growth against age show peaks and troughs, the largest being at adolescence and the girls leading the boys. These he relates to Piagetian stages and also postulates a sixth stage claimed to have been detected by Arlin (1975). Although the number of brain cells are not thought to increase after birth, the myelination (insulation of nerves which results in increased signal transmission) and dendritic growth to form new cell connections is known to continue well into adolescence (Rose 1973). That changes in brain function are occurring is further evidenced by S. Taylor-Parker (1978). She relates EEG (Electoencephalograph) plots against age, which show distinct peaks and troughs, corresponding to the seven M levels of Pascal Leone. A further study of EEG traces was made by Kraft (1977).

The activity of each hemisphere was monitored while children were asked to carry out Piagetian tasks. Following this she made the suggestion that the stages are a measure of the amount of integration between the hemispheres. Taking this a little further, with the view that Piaget's theory describes a gradual shift from egocentricity to detachment and a movement from information being handled in a serial manner to one where both parallel and serial processing are possible, with ability to handle more variables, it is not too great a step to relate this to hemispheric function. The serial thinking due to the dominant left hemisphere being gradually linked to the holistic ability of the right hemisphere as the nerve tracts of the corpus callosum become more myelinated.

Although not conclusive, there is a considerable amount of evidence to support the view that formal thought is a definite stage associated with adolescence. Further, its attainment would appear to have some underlying biological controls linked to puberty.

The nature of formal thinking: does it possess a spatial element?

There is connection between scientific thinking and formal operations, as far as Piaget is concerned they are implicit with each other. A case for linking the former to spatial ability has also been stated. Piburn's study (1980) found a correlation between spatial ability and formal thinking, in particular proportional reasoning.

This association finds support from Piaget himself. In the Child's Conception of Space (Piaget and Inhelder (1956)) the development of spatial awareness is investigated and the conclusion is reached that this development is not innate but requires a gradual transition through Topological understanding, where concepts of closure and proximity are developed, and projective understanding to finally emerge as an internalised operational conception of Euclidian space. The child moves from a perceptual view of space where things are seen to a state of formal operations where spatial transformations can be made in the absence of the objects. The spatial concepts become internalised actions and not merely a series of recalled images. This transition depends on development reaching a stage where differing view points can be co-ordinated.

That proportional reasoning depends on spatial concepts is suggested. Two kinds of reasoning are postulated: firstly a sub-logical mode which is able to construct a holistic field and is spatio-temporal and secondly one of logical reasoning where abstraction is necessary and similarities are sought.

The sublogical mode of thinking is eventually subsumed into the logical mode. This could be linked with hemispheric functions being integrated. Newcombe and Bandura (1966) suggest the possibility of hemispheric reorganisation in adolescence as does Waber (1977). What its function in thought then becomes is not clear and it reminds one of the earlier quote from Macoby and Jacklin Page 5 of this thesis.

Two papers by Wood (1976, 1977) also suggest a connection between visio spatial ability and proportional reasoning. In an analysis of boy girl performance in mathematics he found that most of this difference arose from the ability to deal with questions of proportional reasoning, in other facets of the 'O' level course boys and girls seemed to perform equally. Further support is given to a male advantage on Piagetian tests which involve proportional reasoning in a recent longitudinal study by Flexer and Roberge (1983). Is it possible that factorial analysis could yield a common factor between the two or at least give a greater insight into formal thinking? Prior to a consideration of such studies, the problem, that of marking formal reasoning tasks for use in factorial analysis is not straight forward, should be mentioned. Normally only levels of attainment are given, not marks out of 100.

Bart (1971) tested 90 pupils, three groups of 30 and of average ability aged 13, 16 and 19 years respectively. 32 were male. He claimed no need to control for sex as no sex differences had been noted in tests of formal operations. This is a point of interest that will be returned to, for the present it is noted that this is a variable which could have been eliminated. The tests given were clinical Inhelder tasks and written tests of formal thinking. The first factor common to all tests he attributed to 'g'. The second factor split according to the mode of presentation whether clinical or verbal. Bart, Baxter and Frey (1980) concluded that there is no connection between formal reasoning and spatial ability while cautioning on conclusions drawn from their particular sample which consisted of students studying gymnastics and aged 18 - 25 years.

Lovell and Butterworth (1966) investigating proportionality in formal operations found a general factor. This is confirmed by studies of Lawson and Norland (1976), Lawson, Karplus and Adi (1978) and Shayer (1979). Driver (1980) suggested that perhaps 'g' is being rediscovered. Until larger ranging test batteries are used little can be concluded about tests of formal thinking other than they appear to share a common factor.

Bart claimed there was no need to control for sex as no sex difference was known. Neimark (1975) reviewed the literature and found only one study showing sex bias. Blakeslee (1980) cites some examples of male dominance. Graybill (1975) found sex differences favouring boys, but these results were criticised by Ciesla (1976) on the grounds of poor statistics.

Linn and Thier (1975) found girls superior to boys but put this down to better verbal skills. Maccoby and Jacklin (1975) found little evidence of sex difference.

why?

The overall picture appears to support Bart. Is this not surprising? Isn't it to be expected that more studies would reflect the results of Linn and Thier? The reasons are twofold:

1. Much criticism of Piaget's work is based on semantics e.g. Ausubel, Novak and Harrison (1968) Hughes (1979) Donaldson (1978). He himself states that verbal competency is important in tests of formal reasoning. Piaget (1950).
2. Girls lead in development. Epstein's curves and early puberty demonstrates this.

Why then is no sex difference seen if puberty is a controlling factor? Girls reach this earlier and also have greater verbal ability which should aid the verbal nature of the clinical approach.

A study that could possibly have a relationship to this is that of Fairweather and Hutt (1972) although in caution it should be noted that Fairweather (1976) had completely changed his view from that expressed in the paper with Hutt. From a study investigating and trying to account for sex differences he argues, in the 1976 paper that there is little foundation for such sex differences having reality.

The 1972 paper recounted the testing of subjects for straight forward reaction time and results indicated a clear lead by boys after the age of eight. If, however, a choice had to be made, related to channel capacity, before reacting to the stimulus girls maintained the advantage shown before the age of eight. Further channel capacity showed a step-like increase with age. Girls showed an increase around 8 years and boys 11 years. Using the Broverman et al study, it is suggested that these increases relate to oestrogen and 17 ketosteriod concentrations which occur in the respective sexes at those ages.

The channel capacity of the sexes is tending to converge in adolescence and so any sex difference would seem to be disappearing. If channel capacity is, as it seems, relevant to formal thinking then no sex difference would be expected. The extra verbal ability displayed by girls would, however, still lead one to expect to detect a sex difference. This may not present a problem because although of importance verbal ability alone will not produce success on Piaget's tasks. As already stated it is essential that the process has been internalised before it can be verbalised. Thus if internalisation of relevant schemas has not taken place no amount of verbal ability will enable a solution to the problem. The full relationship between formal thinking and verbal ability is not known, and although its importance has been stated it is not apparent what the expected effects should be.

In the solution of spatial tasks, channel capacity holds less importance than reaction time. This statement does, however, require qualification and this will be considered in the following consideration of cognitive style. Support for the importance of reaction time is given by Rovet and Netley (1982) who found that a group of girls suffering from Turner's Syndrome performed poorly on spatial tests as compared with a control group. The strategies for solution appeared to be the same as the normals, the poor performance being due to slow reaction time in processing information. Berg, Hertzog and Hunt (1982) also found that speed of processing was the governing factor in problems of mental rotation.

Related to this it is of interest to note that a recent article in New Scientist (Cohen and Shelley (1982)) considers the connection between processing speeds and IQ. This takes no count of accuracy and ability to hold an image in spatial tests.

Eysenck (1967) evidences that there is a connection. Extroverts are more likely to make errors in a task than introverts, hold images in memory less well and are field dependant.

The question of cognitive style in terms of field independancy (FID) will now be considered in greater detail.

Cognitive Style and Information Processing

A concern with Piagetian theory which has not yet been mentioned is that not all of the adult population graduate to formal thought in that they fail on Inhelder type tasks. This observation is not restricted to those who are low performers in other aspects of adult life. The problem is discussed by Neimark (1979) and she argues that it may be the underlying difference in cognitive styles which is responsible. The reason she suggests is that because of the generally open ended nature of formal operational tests the field independant person is more likely to be successful, not because of any difference in overall ability but because being field independant will perform at an optimal level in such conditions.

For this to be so a correlation needs to be demonstrated between performance on formal operational tests and Neimark concludes that at present there is only limited supportive evidence. More recently Flexer and Roberge (1983) reported a longitudinal study of the development of formal operations along with that of Cognitive Style. They concluded that although the performance on the Embedded Figures Test (EFT) for Field Independance - dependance (FID) correlated well from year to year the same was not true of that on formal operation tests. This they say is to be expected because of the individual development of pupils within the age range 11 - 13 years. When the effects of IQ were removed from the analysis no advantage was demonstrated by field independants. The point is made though, that the formal operational tests were of a paper and pencil type and this brings the results into line with Neimark's view that field independants only exhibited superiority when tasks are unstructured ones of the clinical type used by Piaget. It is also noted in passing that boys were found to operate better than girls on tasks of proportional reasoning which is in keeping with similar findings already cited. No sex difference was noted on the EFT.

Lawson (1980) correlated field independance with formal operational thought in a group of biology students average age 22.6 years although adding that this cognitive style is not necessary for formal thinking. Rather, in line with Neimark's view, the field independants merely perform better in this particular test environment.

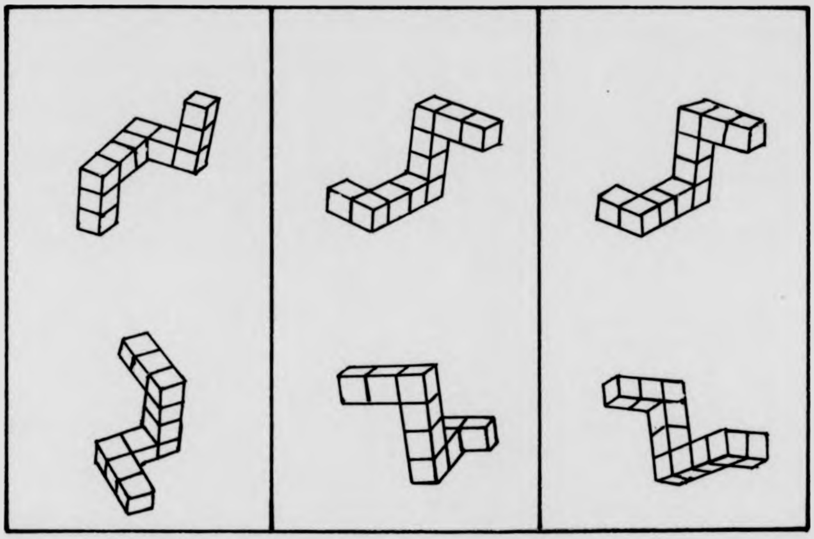
Linn (1980) addresses the same problem and concludes that the context within which a problem is presented will determine how readily a person will perform at optimum level i.e. A person who can reason proportionately in one setting will not necessarily do it in another. This lack of carry over is often observed in school between subjects such as mathematics and science. The overall view then does little to deny Neimark in that cognitive style does perhaps produce a hidden bias which favours field independants on the open ended type Inhelder tasks. The argument does however become confused in a matter of definition i.e. are formal operations characteristic of the thinking of all adults or are they defined in an operational manner so that formal operational thinkers are those who are successful on the open ended unstructured Inhelder tasks presented in a clinical manner. Acceptance of this operational definition leads to the view that field independants are formal thinkers and as Neimark says 'known field independants have in their number mathematicians, scientists and academically gifted adolescents.' Thus if field dependance hampers formal thought on tasks of the Inhelder type which are science based and if as has been argued earlier Spatial Ability and Field Independance are synonymous or at least heavily loaded with a common factor then it can be concluded that spatial ability is indicative of ability in science.

big jump dependent on if

The problem then arises as to the actual nature of spatial abilities i.e. What exactly do these tests measure? The question of processing speed has already been mentioned and this warrants closer examination as it could call into question the theory of visualisation being associated with such tests. Perhaps in the same way as Neimark suggests that field independance - dependance provides an artifact in the observed differences in performance in formal operations so too does processing speed in the performance on tasks of spatial ability.

In order to clarify this problem it is necessary to examine the types of processing that are available for solving spatial problems. This has been done at some length by Pascual-Leone and Goodman (1979) who use the embedded field task in order to illustrate the theory of constructive operators (TCO) mentioned earlier. It is suggested that the solution of embedded figures tests can be arrived at in two ways. The first strategy adopted by field independants is to look at the presented figure which has to be extracted from a complex background. Then taking this whole representation it is searched for in the complex background. Alternatively a less parsimonious approach can be adopted which involves the subject extracting features, such as lengths of line and angles between lines, from the given figure and looking for these one by one within the complex background until the whole figure is extracted. This style is characteristic of field dependants and is processing of the verbal type requiring abstraction rather than an holistic overview. It becomes less effective as the problems become more complex and greater mental effort is needed to control all the relevant schemas. It should be noted that Field independance-dependance is not seen as two distinct areas but as the two poles of a continuum.

What further evidence is there to suggest that those who favour holistic processing which characterises field independance are better at tests of spatial ability? Shepard and Metzler (1971) suggest two ways of solving spatial rotational problems as illustrated over . The two figures were presented on a video display and the testee had to decide whether the figures were identical or mirror images. The time to react was recorded when an answer button was pressed.



3 examples of stimuli used by Shepard

One way is to build up a verbal description of the first drawing and to establish by producing a similar description of the second whether or not the two are identical. The other is to establish a neural representation, which as pointed out by Arib (1972) should not be considered 'as crude as actually setting up little blocks and turning them round in the head because we are talking of a functional relationship rather than strict geometric relationship', and to rotate this to match the current input from the other figure. Introspection by those tested suggested that the latter strategy was used and this was supported empirically when it was found that there was a linear relationship between the angle through which the figure had to be rotated and the time to reach a solution. Similar results were also obtained in a paper folding task (Shepard and Feng (1972)) where it was demonstrated that a linear relationship existed between the number of folds required and time of solution.

This does not seem to exhaust the possible strategies the solution style suggested for field dependants on the embedded figures test being another possible method. Also consider the 3-D views exercise in the appendix - which is akin to Piaget's mountain task where the testee has to decide what view would be obtained from various directions. To do this rather than rotating the diagram the person can actually imagine themselves in the field as if they were actually looking from that viewpoint. Mohammed goes to the mountain rather than the reverse. A similar strategy could be applied to the figures of Shepard and Metzler so that the view of each figure from a selected direction has to be matched.

The point in question though is not the differing ways a solution is arrived at but which one gives the best i.e. most rapid results and this seems to be the holistic type processing of the field independant. Thus spatial tests against the clock will yield higher scores for field independants although most people will eventually arrive at the correct solution. Thus a view is reached which marries together reaction time and the ability to hold an image as being decisive factors in the solution of spatial problems.

Summary

From the evidence presented there are some overall patterns. Scientific ability does seem to be associated with spatial ability which in turn is linked with personality characteristics. Changes in formal thinking and spatial ability both occur at puberty and although the nature of the former is not well defined there may or may not be a causal link. Both seem linked with proportional reasoning ability.

The assembled picture is complex and Bergson's view that science and mathematics fragment whereas a person is holistic and changes with time is brought to mind. Clearly this is an important point and will be referred to again later in this thesis. Individuality should not be overlooked and it should be remembered that statistics only indicate the trend within a group or suggest the probability that an individual will conform to this.

The development and use of a spatial test

The development of this particular test originated from initial observations that spatial ability could be used as an indicator of science ability. The review of the literature in the field of spatial ability has reaffirmed this view. The age group the test would be used with was centred around the change from middle to secondary school. There appeared to be no readily available tests that called for the ability to manipulated objects presented pictorially in two and three dimensions, thus a test believed to examine these abilities was constructed. If the test proved of use in this intended capacity there would be a demand to include more bias towards this type of questioning within a more conventional test of science abilities. These often are largely tests of memory rather than application of knowledge. A spatial test should be free from this, although it can always be argued that there will be an unfair advantage to some on the grounds of environmental influences. It would also be the type of test that many pupils would enjoy doing.

This view of the spatial test as an indicator of ability in science received later encouragement from Camp (1981) who used a test of spatial ability to select groups of males and females of high and low spatial ability. They were then clinically tested with physics problems thought to require spatial thinking to solve them. The problems posed had to be visualised in an objective manner. She concluded that those of high spatial ability were most able to achieve this.

As there is less objective testing within the system, due to cost and removal of the 11+ examination prior to secondary school entry it is possible pupils who favour spatial thinking, because the ability may never be tested, will tend to be overlooked. It has been stated that in our schooling we emphasise the left hemisphere (Wittrock et al 1977). Rubenzer (1979) reviews the importance of the right hemisphere relating to creativity and problem solving. Beckman (1977) and Olson (1977) both note the possible use of spatial testing as a means of identifying able pupils. It has been said in a DES report of 1977 that it is the more able pupils who are being neglected. A method of objective testing such as that of the spatial test could well serve to indicate some pupils who are not achieving their potential.

Reasons for development of the test

The stimulation for the development of the test came from selection procedures used by schools sending pupils to the Science Summer Schools held annually at Warwick University. A report of the original summer school held in June 1979 was published in a monograph by the Institute of Education at the University of Warwick (ESTEAM (1980)).

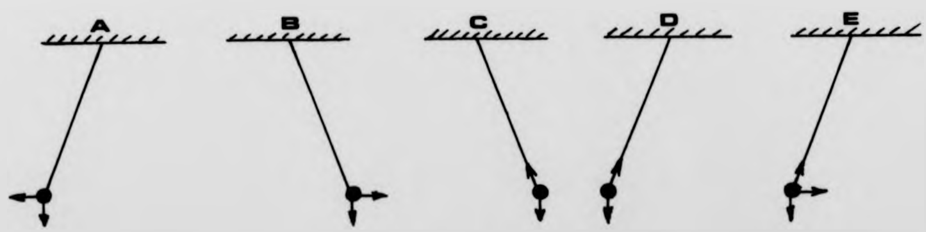
Only three pupils were selected from each school and it is doubtful if any sort of case could be presented that would justify such selection being totally satisfactory. No claim will be made here that a single test could be used for such a fine judgement, rather it could be used to help with the initial screening of the population. It should also be added that as well as being of benefit to the pupils selected the summer schools were originally intended as inservice training for teachers with the intention of producing teaching materials suitable as enrichment for able pupils in science. A number of these have now been published by the ESTEAM project (1983).

With the onset of comprehensive education the number of pupils in a single school has increased. In many cases this has resulted from the combination of schools on separate sites. Initial screening to obtain a group from which three pupils were eventually selected would usually be subjective and defy any interform comparison. A common easily administered test would be of help in this situation.

Selection was for attendance at a Science summer school and so if such a test could be indicative of potential in science rather than general ability then it met the aim of the project even better.

Introspection led to the feeling that visualisation and spatial ability were called on in science especially at more advanced level. The following example will illustrate the point.

A pendulum hangs from the roof of a railway carriage of a train that is travelling at a constant speed around a horizontal curve bending to the right. Which of the following diagrams correctly represents both the directions of the forces acting on the pendulum and the steady position of the pendulum as seen looking forward from the rear of the train?



It is subscribed here that the solution of this 'A' level physics problem would need a spatial representation to be constructed from the verbal and diagramatic description. This view has been concurred with by others presented with the problem. From this background the test was developed.

Would this not be more appropriately placed in the description of the pendulum investigation?

Contents of the test

In the original planning it was also intended to include some verbally presented tests which called for subsequent visualisation. As an example of this Hughes and Hughes (1937) give the following problem which they claim requires imagery:

'A squirrel is clinging to one side of a tree, and a man is standing opposite on the other side of the tree. The man walks round the tree, but the squirrel also moves round the tree so as to keep just out of the man's sight. They continue this movement until each has gone entirely round the tree. Has the man gone round the squirrel

a) in the sense of having been in front, behind and on both sides of him?

b) in the sense of being north, south, east and west of him?'

Examples of this and those requiring visual to verbal interpretation can be seen in appendix B.

After trials of questions of this nature, it was apparent that marking was difficult. The use and interpretation of language presented the main obstacle to this, problems immed- [illegible] iately coming to light from the example above. Camp (1981) found this a major problem in the clinical use of her science problems. The development of such questions as a written test was abandoned at that stage in favour of the more object- Which? ive form.

This does not mean that such questions do not test skills necessary in problem solving. It is in fact probable that they represent the type of approach necessary in much advanced problem solving. The problem often has to be visualised before the logical methods of mathematics can be used to make a spatial transformation and the final interpretation of results may again call on this skill. Such requirements become greater in advanced and degree work, and so Smith (1964) suggests it is not until 'A' level and beyond that the full value of spatial skills becomes apparent. It could be said that by abandoning the afore-mentioned approach the issue had been side stepped rather than met head on. In retrospect there could be some truth in this although rather than abandoned it could be considered shelved, and perhaps be returned to as a basis for producing questions of an holistic nature, taking into account both the left and right hemispheric abilities. The breaking down and reassembly of parts is certainly an important feature of problem solving (George 1980).

The spatial test adopted was objective and initially divided into six parts. The items were all intended to require visual manipulation of the figures presented. Retrospectively it is seen that they fit in with Eliot's (1980) paper already discussed.

They are based on the more advanced exercises associated with formal thinking in 'The Child's Conception of Space' (Piaget and Inhelder (1956)). It should be re-emphasised that Piaget would not have advocated the testing of these abilities by objective methods. Rather questioning would have to be clinical so that it could be assessed whether or not the child could verballise the visual transformation he or she was being asked to make. Only when the verbalisation was successful could it be taken that the transformation was truly internalised.

In the book they say for example that 'The child who is familiar with folding and unfolding in work at school is two to three years in advance of those who lack this.'

Also in regard to surface development (see cut-out) it is argued that this rests on the coordination of view points. Herein lies the difference between perception and transformation. Although the child may have all the required view points to hand it is not until they are coordinated into a single fluid schema that a solution to the problem is made. The transformations have not been fully internalised and it is this internalisation which is the crux of Piagetian ideas. Internalisation only comes about by the action of the subject on the surroundings, it is no good he or she being a passive observer. The internalised world is constructed by a person's interaction with the environment (Piaget 1971).

The full list of items is as follows. See Appendix A for the actual items.

1) <u>Turnabout (figural rotation)</u> This consists of five figures, one being a mirror image of the others. Mental rotation is required to bring each figure into the same orientation, to match them and identify the odd one which is a mirror image of the other four.

2) <u>Cut-out</u> - standard surface development exercise already referred to.

3) <u>Jig-cubes</u> - This is a three dimensional drawing of a cube which could be cut into pieces along dotted lines shown on its surface. There are two of these cubes and randomly orientated pieces that would result from the cubes being so sliced. The pieces have to be allotted a cube. This would probably be classed by Eliot as a combinatory task requiring rotation and combination.

4) <u>Folds</u> - standard type mental paper folding.

5) <u>3-D views</u> - These are three dimensional drawings from which the testee has to pick an appropriate view from different indicated directions. This would be classed as a perception task.

6) <u>Cross sections</u> - Three dimensional drawings, the cross sections of which had to be visualised if a transverse cut were across the plane indicated. Following initial trials of the test with a group of girls aged 12 to 13

years it became apparent that cross sections were unsatisfactory because of the problem of giving instructions that could be easily understood prior to the test.

The Administration of the Test

Each section of the test was preceded by instructions and a trial. If the instructions were not understood they were reviewed until all candidates felt they knew exactly what was being asked of them. The section items were then administered in a fixed time. As a result of preliminary trials cross sections was replaced by another section as follows.

7) Pencil and belt - This like the cross sections problem was not a conventionally accepted test item. The testee was asked to visualise the movement of a pencil and belt relative to each other to produce a drawing.

This section was eventually somewhat reluctantly discarded because of similar difficulties encountered with cross section in giving clearly understandable instructions.

The timings for each section were initially obtained from a small group of 14+ year old girls in very preliminary trials. Following more trials these were later altered to produce a distribution that avoided bunching at the top end, indicative of too many completing the test.

Discussion of test reliability and validation

Before the analysis of results obtained from using the test a discussion of what the test was intending to achieve and how far this has been met will be made. This should not be interpreted as a drawing of conclusions before obtaining results, merely a discussion of methodology used.

The aims of the study were to produce a test of spatial ability and to use this as a tool possibly predictive of success in science. This, as has been made clear by the review of the literature does not assume that spatial ability is the only ability needed for success in science which would clearly be a riduculous claim, rather that spatial ability is not generally tested in the school population, whereas other abilities are more obviously tested by conventional examinations. This then implies that the test should be a true measure of the ability it claims to measure and further it should correlate with success in science.

The question of its predictive value of success in science is not possible with a short term study that can only test and observe sections of the population at one point in time.

If it was shown that people who were successful in 'O' level science courses showed greater spatial ability than those on arts courses this would tell us little. If they were tested during the examination they could have built up a background of experience reflecting the choice of course, and so unless one assumes spatial ability is a trait little altered by experience, in that those who are good spatially have always been so (i.e. the rank order in spatial ability is static within the population irrespective of the age it is measured at and independant of experience) these results cannot be then extrapolated to the age of 12 years. In addition changes in adolescence make such an extrapolation unreliable. The requirement is a longitudinal study and one that could be exceptionally well planned to cover all aspects of development.

Given the limitations of transverse study, what measurements are possible? The first is to correlate results of school examinations. This raises two problems:-
Firstly the examinations themselves are not reliable and generally exactly what skills are being measured is not clear. Secondly the construction of another independant test to measure scientific ability, adopted by Camp (1981) is unreliable.

It must always be remembered that it is possible that the significance of spatial ability may not emerge until higher levels of education are reached.

Thus in a short period of investigation one can examine school performance. It must however be concluded that given the limitation of time and resources a predictive test of scientific ability cannot be validated. The test is required to be a reliable and valid measure of spatial ability. These terms need further clarification in the context of the present test. In simple terms the question is does the test measure spatial ability and if so, how accurately does it achieve this?

Accuracy of measurements (reliability) is frequently of concern in the construction of a test. There will always be a random error in methods of measurement and normally this can be eliminated by taking a number of measurements and assuming that in a mean value the random error will disappear. If the measure were the length of a plank of wood then all the variance would lie in the error. With a complex system such as man, there will be variance in the true measure depending on chance experience and mood. Similarly if the plank's length depended on humidity a variance would be introduced uncorrelated with random error. Thus reliability is defined to be the proportion of the variance of a set of measurements that is the true variance. (Guilford and Fruchter(1966)). It is also important to realise that reliability when quoted it only valid under the conditions it is measured. The plank's length would show greater true variance in England than in the Sahara desert where humidity is normally very low and steady. If conditions of measurement were static the true variance would be expected. With random errors eliminated, the same value would be obtained on each measurement and reliability would be one. In practice this is never achieved. There are three recognised ways of measuring this reliability

(i) Test - retest - this involves the correlation of two sittings at the same test. It has the drawback of carry over of experience altering performance and possible memory of particular item answers. This can be reduced by the increased time interval between testing, but unfortunately introduces the unknown factor of maturation. Even so, carry-over is always unknown. This point is illustrated by attempts to measure growth spurts in mental capacity at adolescence by Kohen Raz (1977), Lyng Bergt-Olov (1965) and Berglund (1975). The assumption in these had to be that variance was due to maturation in test retest, but as they point out, carry-over of experience is always an unknown even at yearly intervals.

(ii) Alternative forms - two forms of the test are constructed and administered at the same sitting. The problems raised here are in ensuring the two forms are parallel and in maintaining motivation.

(iii) Split half technique - the test is divided into two halves using for example odd and even numbers. Problems arise in deciding how to make the split, although there are ways around this (Richardson and Kuder (1939)). These become greater as the test becomes more heterogeneous (it measures mixed traits).

Guilford and Fruchter (1966) differentiate between power tests which allow time to complete questions and speed tests where all items are not completed. They say that split half methods, and indeed any internal consistancy method of estimating reliability is inappropriate so that alternative forms becomes the best method of estimating reliability. With the present test an alternative form was not constructed at the outset and although this approach is advocated by Guilford and Fruchter it is brought into question by Evans and Waites (1981).

Notwithstanding that statement, Heims, Watts and Simmonds (1972) use a split half method to estimate reliability of their shapes analysis test. The same is true of Lowes (1961) in development of NFER spatial test 3 (closed test for examining age group in question) although it is not clear whether or not he adopted equivalent half approach to this, i.e. split is made before rather than after testing so the halves are separately timed. Consideration should also be given to item analysis, but again the methods adopted examine the numbers passing or failing an item and in a speed test when many items are not attempted this is not appropriate. This will be discussed further after validity has been examined.

Validity is a question of determining whether or not the test is one of spatial ability. Even though it appears to be from its construction, it is possible that it can be solved by verbal type reasoning rather than the spatial one assumed.

There are two recognised ways of testing validity. The first is to take another test of spatial ability and correlate the new test with the accepted measure. Tests however will never be pure measures but will be made up of a number of factors, and so although there may be a degree of correlation it cannot be concluded that the correlation is of the particular trait the new tests purports to measure.

The way to overcome this objection is to include the new test in a battery which is known to contain a number of factors including the one of interest. A factorial analysis will then yield loadings of the new test in each of the separate factors it contains and hence the extent to which it measures the factor in question can be seen from the loading. Possible objections to factorial methods have been raised earlier, but aside from these is the problem of obtaining a large and suitable test battery. Smith (1972) uses this treatment of the shapes analysis test. With nine tests in the battery he was only able to extract a g/k factor as opposed to separating them. However in most instances the factor analysis approach is the method adopted to ascertain validity.

With the spatial test there is, to some extent, a way of showing validity without these methods. This is embodied in the definition of spatial ability adopted by Camp (1981) which stated that such tests demanded the manipulation of pictorially represented data and must show a sex difference in favour of males. Lowes (1961) made use of this approach although he added that for full validity factorial methods should also be applied.

The development of the present test will now be considered in more detail.

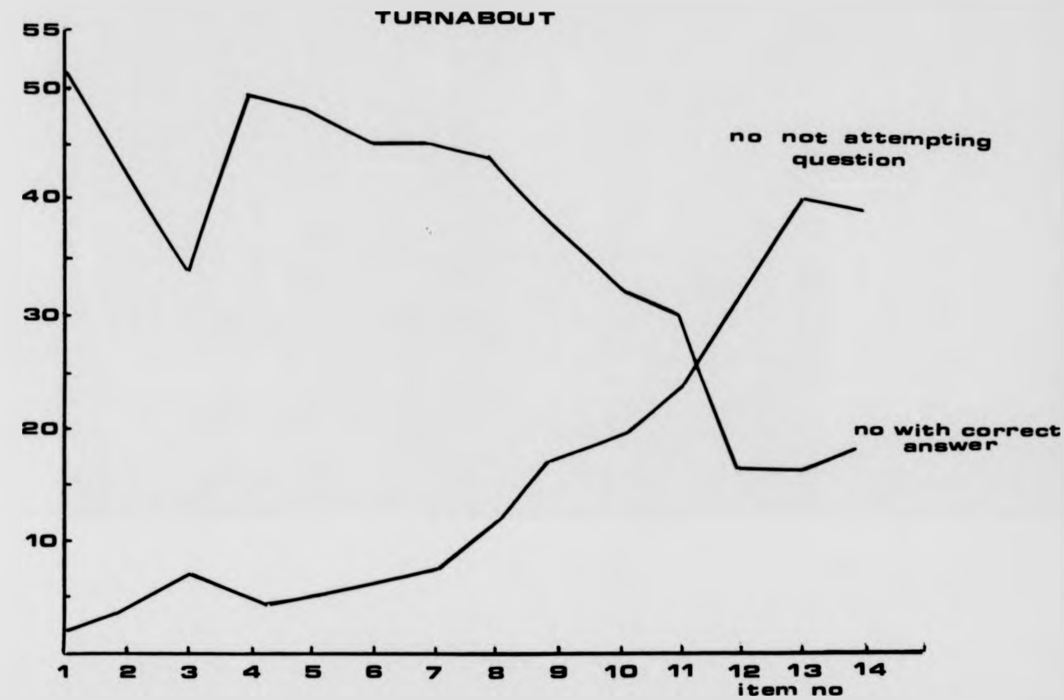
Results of initial trial items attempted

These results were obtained by testing a single group of Girls aged 12 - 13 years (N = 57). The group consisted of two forms that had been selected by 12+ examination for grammar school admission. The test was administered in the school hall at a single sitting. The girls were given as much time as they required to read the instructions to a section. If any sections were not understood after the trial item answers had been given further explanation was made. The item was then run with the times as indicated.

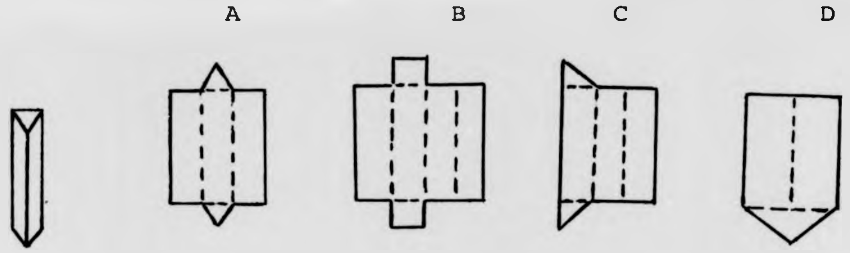
1) TURNABOUT - 4 MINS. Mean - 8.9 Standard Deviation - 4.1

A B C D E

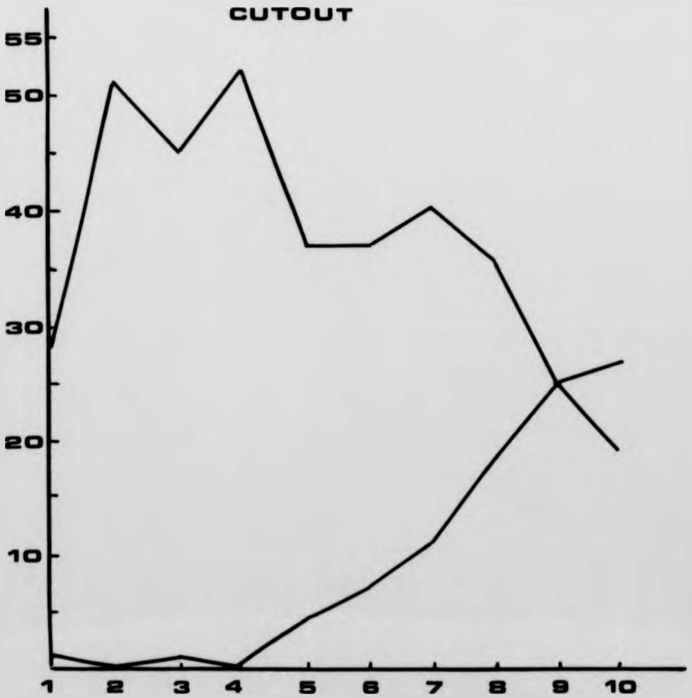
Correct Answers		E	D	A	D	A	B	D	B	D	A	D	E	A	C
Items Answered	A	1	4	34	2	48	2	0	0	0	33	0	0	16	0
	B	0	5	7	1	2	45	0	44	0	1	0	1	0	0
	C	1	1	4	1	1	4	1	0	1	2	4	4	0	18
	D	1	43	3	49	1	0	45	1	38	2	30	3	1	0
	E	52	0	2	0	0	0	3	1	0	0	0	16	0	0
Not attempted		2	4	7	5	5	6	7	11	18	19	23	32	40	39


TURNABOUT
55
50
40
30
20
10
no not attempting question
no with correct answer
1
2
3
4
5
6
7
8
9
10
11
12
13
14
item no

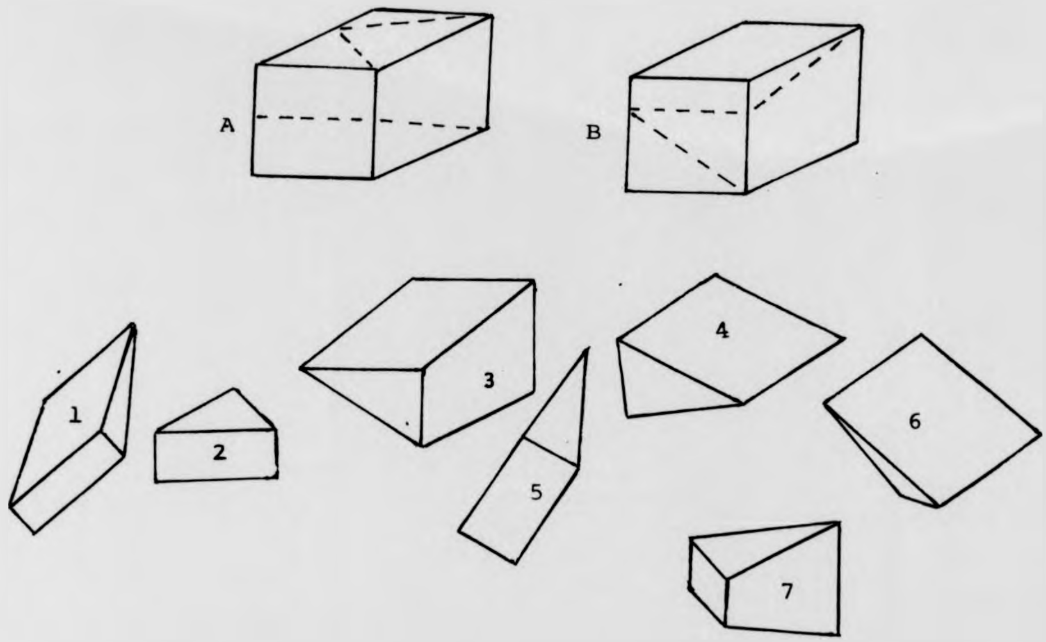
2) CUTOUT - 3 MIN. Mean = 6.2 Standard deviation 2.6



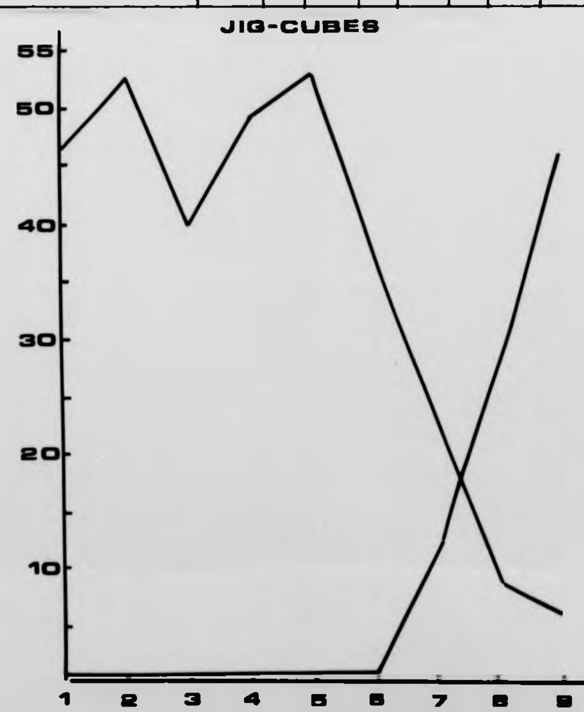
Correct Answers	B	A	C	A	D	C	B	A	C	D
Items Answered A	23	51	7	52	2	5	2	36	31	19
B	28	1	3	0	4	3	40	1	4	1
C	2	3	45	2	10	37	2	2	25	1
D	3	2	1	3	37	5	2	0	2	9
Not Attempted	1	0	1	0	4	7	11	18	25	27



3) JIG-CUBES - 5½ MIN   Mean = 5.2   Standard deviation = 1.6

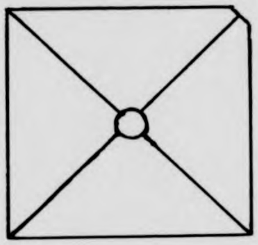


Correct	47	53	40	49	53	37	23	9	6
Wrong	9	3	16	7	3	26	23	20	5
Not attempted	1	1	1	1	1	1	11	28	46

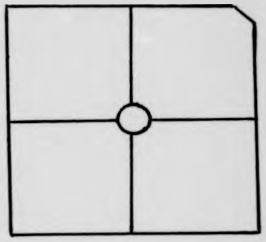


4) FOLDS - 3 MIN   Mean = 4.4   Standard deviation = 1.7

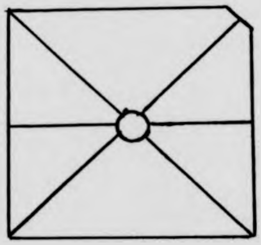
A



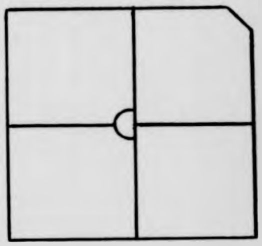
B



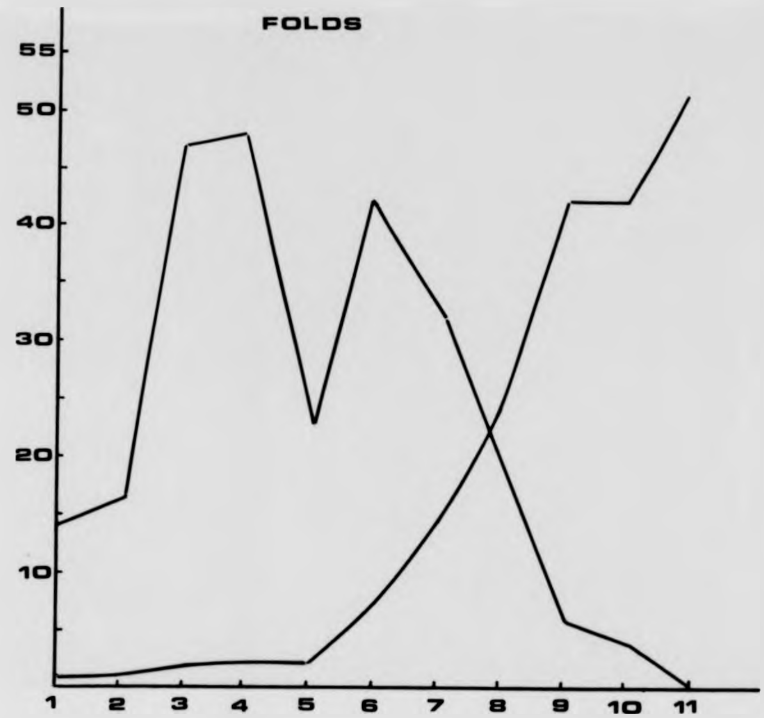
C



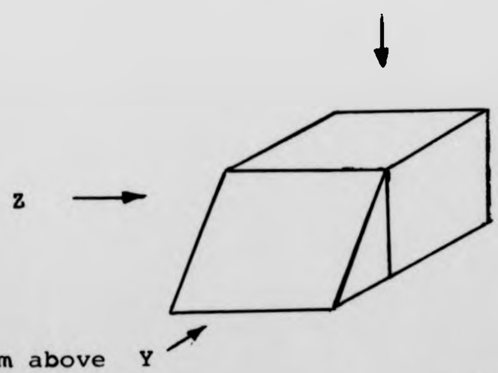
D



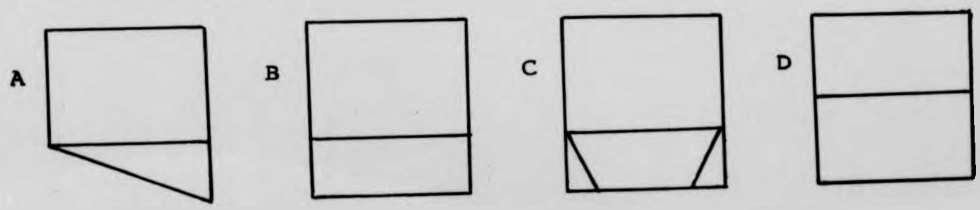
Correct Answers	A	C	B	D	A	B	C	A	C	B	B
Items Answered A	14	0	2	1	22	2	8	20	1	1	2
B	42	40	47	1	19	42	2	4	0	4	0
C	0	16	4	5	12	5	33	3	6	9	2
D	0	0	2	48	2	1	0	6	8	0	1
Not attempted	1	1	2	2	2	7	14	24	42	42	52



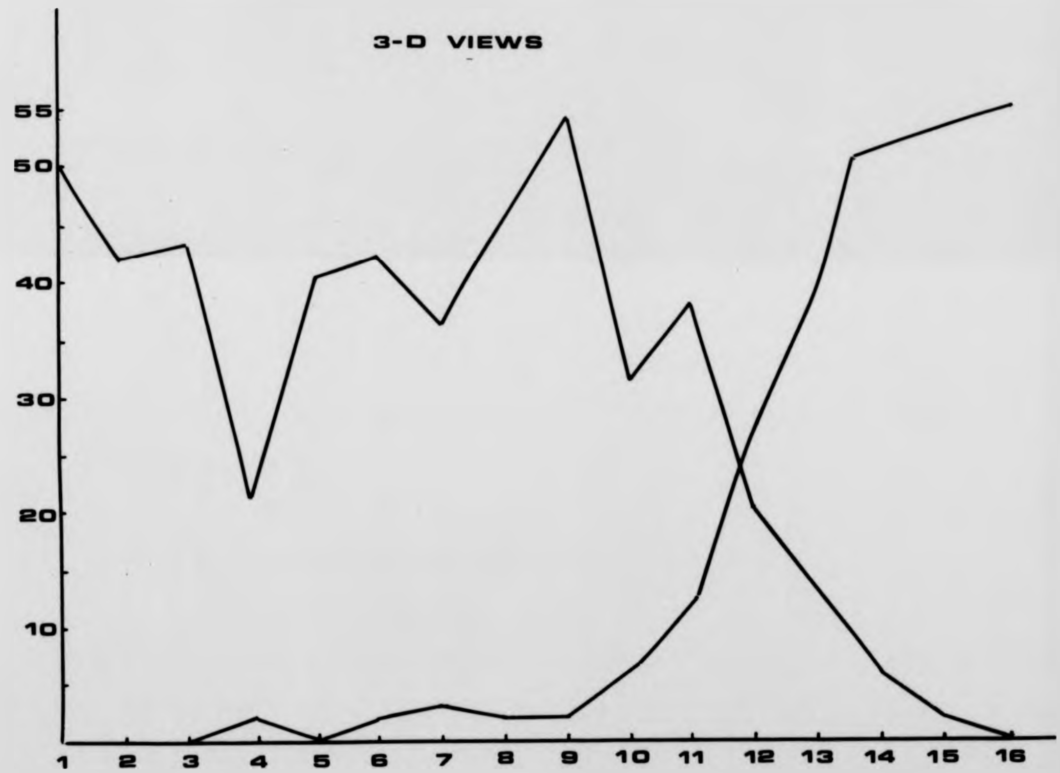
5) 3-D VIEWS - 4 MIN Mean 8.4 Standard deviation = 2.9



View from above



Correct Answers		A	D	A	B	A	B	A	B	B	C	B	D	A	B	C	A
Items Attempted	A	50	5	43	11	40	4	36	2	3	2	1	2	13	0	0	0
	B	6	8	9	21	0	42	6	45	54	10	38	3	2	6	2	1
	C	1	2	2	6	4	9	1	3	0	31	6	5	1	0	2	0
	D	0	42	3	17	13	0	11	5	0	8	0	20	2	1	0	1
Not attempted		0	0	0	2	0	2	3	2	2	6	12	27	39	50	53	55



6) X-SECTION - see appendix, abandoned because of reasons explained.

The Correlation Matrix for the test sections and final total was as follows:

	Turn	Cut	Jig	Fold	3-D	Cross Section	Total
Turn	-	32	33	21	22	15	73
Cut	32	-	18	21	33	34	63
Jig	33	18	-	23	32	16	55
Fold	21	21	23	-	52	-10	55
3-D	22	33	32	52	-	16	69
Cross section	15	34	16	-10	16	-	34
Total	73	63	55	16	69	34	-

The total score was correlated with a verbal reasoning score from middle school.

N = 53 Verbal Reasoning Mean = 119.6

Standard Deviation = 5.46

Product - moment correlation = 0.40

Discussion

How do you know it's [illegible] [illegible] ability

This breakdown of how the questions were answered shows that with some exceptions all items are easily answered and the variance is mainly caused by the time allowance for the section. Such a statement becomes less certain with items towards the end of each section because of the reduced numbers attempting an answer. This will always be a problem in item analysis of speed tests.

From this initial trial it was apparent that

1) TURNABOUT - all items appeared to be possible as the correct answer dominated. Eighteen candidates completed this section with seven returning all correct answers. This indicated that the timing required adjustment. After further trials three items were added to this section and the time retained.

2) CUTOUT - all items except the first caused few problems. The two items A and B of the first question were equally popular as the correct answer. This confusion of responses has not been as noticeable on ensuing trials. Similarly bunching at the high end of the mark distribution again indicated that a reduction in time was needed.

3) JIGCUBES - as would be expected the early items, in which less information had to be processed, caused few problems. Few finished this section so the time was retained. Following further trials with wider bands of ability the time allowance was extended by a further minute.

4) FOLDS - the first two questions proved to be the only ones on the entire test where more wrong than right answers were given. This was initially thought to be due to the fact that

two answers were possible as the word <u>inside</u> was not then included in the instructions. The subsequent rewording did not produce a change in this tendency. The two items were retained.

5) <u>3D VIEWS</u> - all items except 4 appeared to be correctly answered by the majority. The timing was thought appropriate but further trials caused its reduction to 3½ minutes.

The results indicated the problems with item analysis in a speed test and also the difficulty of judging reliability by an internal examination as opposed to the method of test-retest. The items that are thought to be the most difficult to answer are at the end of each section and because of the timing restrictions, are attempted by fewer candidates. Subsequent trials do show all items to be satisfactory. This statement includes the wording which introduces each section with the exception of cross-section and pencil and belt which were rejected owing to dissatisfaction with this. It is possible that if the test were to be used only with high ability groups that pencil and belt could be used although rejected for broader ability.

One encouraging thing that has become apparent in the administration of the test is that guessing seems to be minimal as indicated by the distribution of answers so it is not felt that this needs consideration in item analysis.

The subsequent trials with mixed groups resulted in the time allowances given in the appendix. These now give an approximation to normal distribution on the separate sections and on the total score when used with mixed ability groups in the age range 11 - 13 years.

An examination of the correlation matrix also gave reason to leave out the cross section item. The overall pattern indicated that each section of the test appeared to be making a contribution to the final total while testing differing aspects of spatial ability.

The correlation of results with middle school verbal reasoning score was made but owing to a small spread of ability (standard deviation = 5.46) and lack of full information as to its origin is considered to be of little significance.

Trials other than the one quoted will not be reported in detail as they merely involved practice in administration and led to adjustment of timings. One exception to this was the testing of a group at Banbury School studied by King (1982). The details of this are included as they do involve some longitudinal study which raises some questions. In addition there is available a large amount of well documented information on these children.

<u>Results from the Banbury Group</u>

This group consisted of forty pupils, twenty boys and twenty girls. They were divided into ten boys and ten girls of high ability and a lower ability control group by initial screening. Their progress during the first year of secondary school (11+ intake) was closely followed. During the year the top group was shown to contain fifteen pupils with IQ's of 129+ measured by individual testing using WISC short form intelligence test (1976). In addition to many other tests the spatial test was administered to the group. Although correlations do not reach any level of significance the ones relating to the spatial test are included in the present thesis as they are the only comparisons that exist between standardised test scores and this test.

Intergroup and Intersex comparison of spatial test scores

Numbers of candidates under consideration will be given in the normally accepted form N = throughout the rest of the results in this thesis.

	Mean	Standard deviation
All	45.3	9.7
Boys (N = 20)	47.8	7.9
Girls (N = 20)	42.6	10.8
Control (N = 20)	38.4	7.4
Able (N = 20)	51.0	6.5

Individual IQ and Spatial Test (N = 21)

The group contains twenty-one as one of the control pupils scored high enough to be considered able.

Test	Mean	Standard Deviation
Spatial	51.3	7.1
IQ	132	9.5

Product moment correlation factor (r) = 0

Correlation of Spatial test with various tests

Three figures will be given for these so camparison can be made.

1) Correlation for all N = 39

2) Correlation for able N = 19

3) Correlation for control N = 20

N.B. One pupil is missing from the able group as full data was not available due to absence.

	VRQ	Arithmetic	AH3	Lang. Apt.
Spatial 1)	0.58	0.41	0.57	0.56
2)	0.00	0.28	0.35	0.50
3)	0.44	-0.11	0.06	0.00

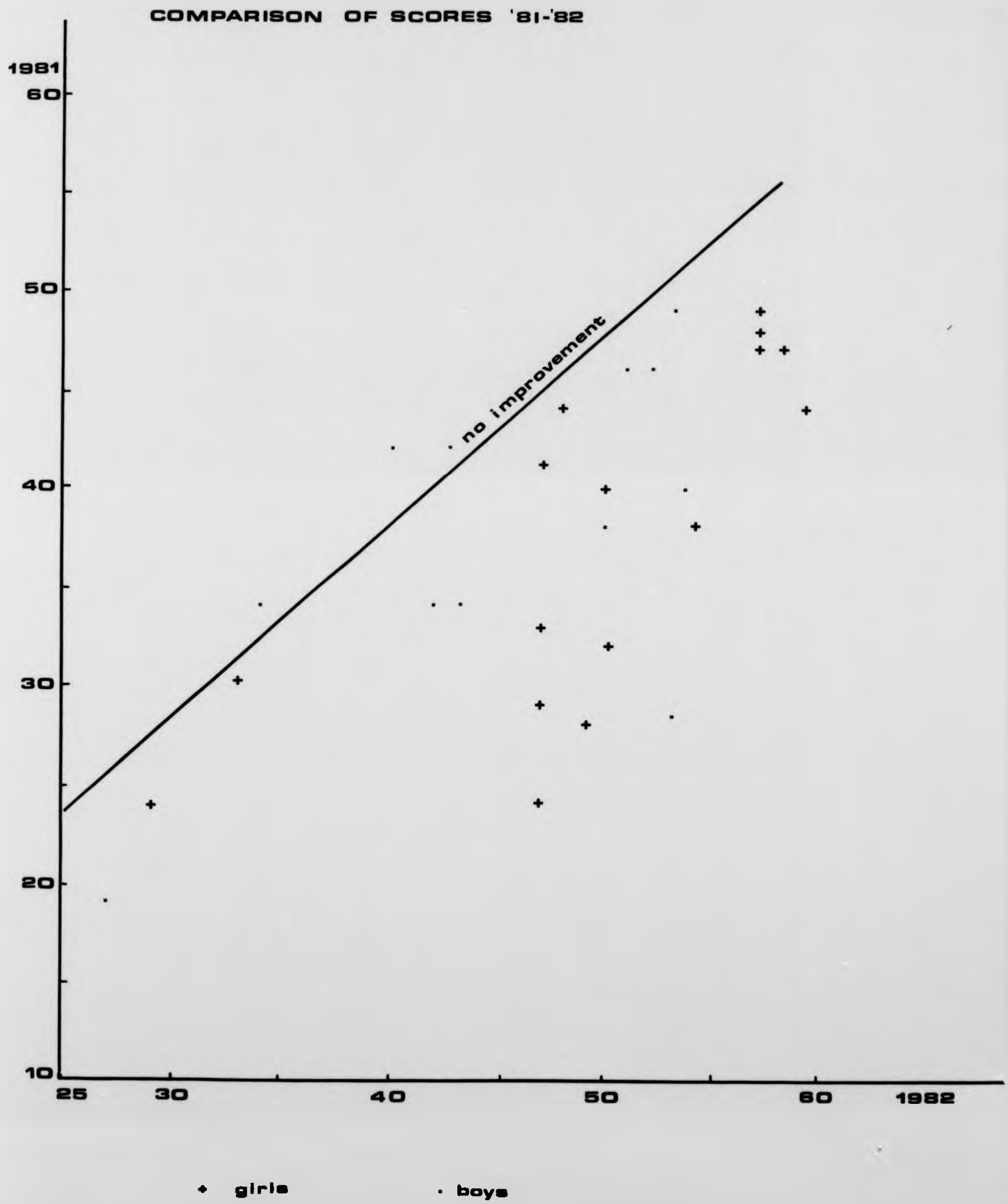
Some of the group were retested with the spatial test twelve months after the first sitting. All pupils were not available at time of testing The results are not directly comparable as the first administration was in trial form and the second in appendix version. Pencil and belt was not therefore common so the score on this was subtracted from total score in 1981. The 1981 and 1982 scores are shown on a scattergram to aid comparison. It should also be remembered that different timings were used. The full changes were:

1) Turnabout - time same 3 items added.
2) Cutout - time reduced from 3 min. to 2¾ min.
3) Jigcube - time increased from 5½ min. to 6½ min.
4) Folds - time same
5) 3D views - time reduced from 4 mins. to 3½ min.

<u>Comparison between sexes 1981 and 1982</u>

	Boys (N=14)		Girls (N=16)	
	1981	1982	1981	1982
Mean	45.7	46.9	42.6	49.3
Standard Deviation	10.6	9.0	9.3	8.5

No differences were significant $p < 0.05$ using t-test.


COMPARISON OF SCORES '81-'82
1981
60
50
40
30
20
10
25
30
40
50
60
1982
no improvement
+ girls
· boys

Results of Further Trials in the 12+ Age group

The time of testing in May means that most pupils would be 13+. The testing was carried out on the top bands of two comprehensive schools. Again interschool marks are not comparable owing to time alterations. The pencil and belt was included in both making the scores from 75 as opposed to 63 in the final version. The scores are included as the sex difference found between the scores of all groups will be referred back to in the final conclusions.

Results from Aylesford School

Aylesford is a comprehensive school on the outskirts of Warwick.

	ALL (N=51)	BOYS (N=21)	GIRLS (N=30)
Mean	48.7	52.7	45.3
Standard Deviation	10.7	8.1	11.4

Differences in means between boys and girls $p < 0.01$

Results from Heart of England School

A comprehensive school situated in Solihull.

The top band forms were tested on separate occasions.

	ALL (N=30)	BOYS (N=15)	GIRLS (N=15)
Mean	45.6	49.7	42.7
Standard deviation	9.2	7.2	9.2

	ALL (N=28)	BOYS (N=17)	GIRLS (N=11)
Mean	43.4	47.7	37.9
Standard Deviation	10.0	9.8	7.2

Differences in means between boys and girls $p < 0.01$

Results of final appendix version of the test on large groups of mixed ability

Coundon Court School

This is a ten form entry comprehensive school situated on the outskirts of Coventry. The intake is from a mixture of rural and urban background. Coventry has a large independant sector of schooling. The intake is at 11+ but time of testing would again mean that the majority of pupils would be 12+.

The test was administered to the ten teaching groups at a time when they would normally have science. The ten separate testings were carried out by the writer during one school week. At the start of each session pupils were told that the testing was part of a research programme which was investigating the way in which people visualise things. They were allowed to ask any questions they wished relating to this and the test was given in the manner already described. The mast majority of pupils appeared well motivated and the overall impression was that they quite enjoyed doing the test.

<u>Results of testing</u> (scores on individual sections are included)

A = ALL (N= 263)   B = Boys (N = 153)   G = Girls (N = 110)

	Turn			Cut			Jig		
	A	B	G	A	B	G	A	B	G
Mean	7.6	6.8	7.5	5.9	6.0	5.6	4.7	4.7	4.9
Standard deviation	4.4	4.7	3.9	2.1	2.3	1.9	2.0	2.1	2.9

	Fold			3D			Total		
Mean	4.7	4.6	4.8	6.2	6.5	5.8	29.1	29.6	28.6
Standard deviation	1.9	2.1	1.9	3.2	3.4	2.9	10.0	10.6	9.1

A = ALL (N = 263)   B = BOYS (N = 153)   G = GIRLS (N = 110)

The difference between the means of the boys and girls is not significant.

A significant sex difference had been noted on previous trials and as these were with selected groups the Coundon scores were re-examined by removing all scores below 20. This could be considered as removing the band of lowest ability in spatial tests.

Results of Coundon Court with scores below 20 excluded

	Turn		Cut		Jig		Fold		3D		Total	
	B	G	B	G	B	G	B	G	B	G	B	G
Mean	9.5	8.5	6.9	6.0	5.3	5.2	5.1	5.2	7.6	6.3	34.2	31.3
Standard Deviation	4.1	3.4	1.8	1.7	1.8	1.5	1.9	1.8	3.1	2.7	7.9	7.1

B (N = 121) G (N = 90)

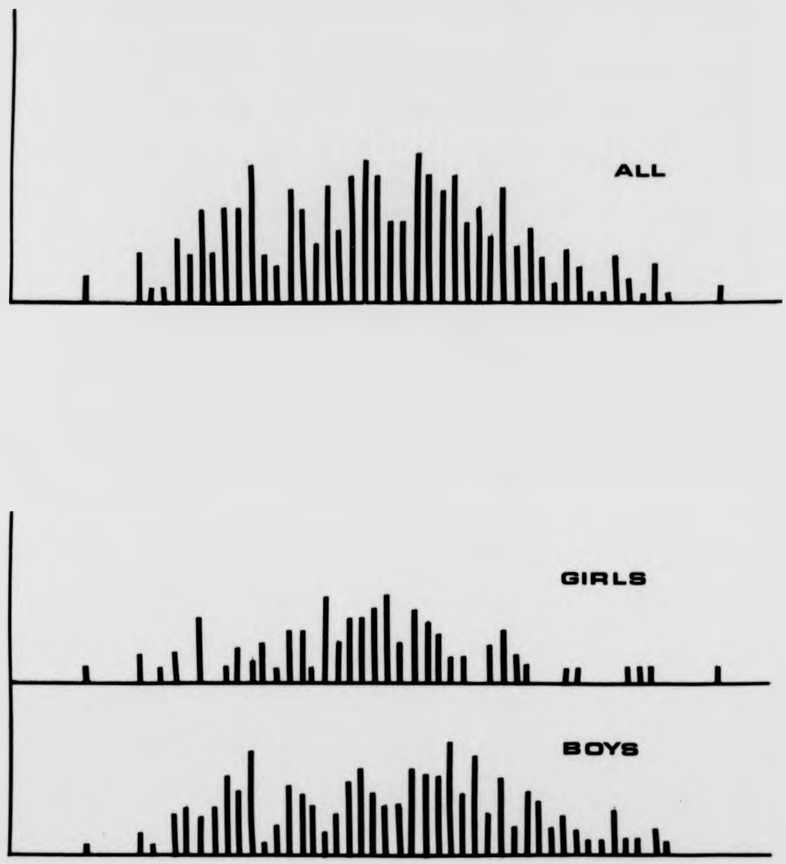
The difference in means is significant $p < 0.01$

Product - Moment Correlation values

The individual tests were correlated between themselves and the final total. These are given below, all values reaching significance at the $p < 0.01$ level.

	Turn	Cut	Jig	Fold	3D	Total
Turn	-	34	42	35	42	81
Cut	34	-	37	35	39	63
Jig	42	37	-	36	39	66
Fold	35	35	36	-	37	62
3D	42	39	39	37	-	74
Total	81	63	66	62	74	--

SCORE DISTRIBUTION COUNDON



Atherstone School

Atherstone school is a comprehensive school of eight form entry. It is situated in the Warwickshire town of Atherstone and was formed by the joining of the town's grammar and secondary modern schools which are situated close to one another. The intake of the school is very mixed in social background and is probably less depleted of pupils at the upper end of the ability range than is Coundon Court. This is due to two factors; the grammar school origins and the relative inaccessibility of an alternative independant sector. The intake is at 12+ but again time of testing meant pupils would mainly be 13+.

The test was administered at a single session, the two lowest ability groups of the intake being omitted. The reason for this was that the forms, unlike Coundon Court, were already banded in ability and as the test was part of a programme to isolate an able group of pupils within the year, it was not seen necessary to include these two groups. Although all pupils sat the test at the same time, the groups had to be split due to the fact that they could not all be accommodated in the school hall. The writer administered the test to a group of the two top ability band forms and two from the middle band. The remaining two groups from this middle band were given the test by two members of staff. The procedure already outlined was used with the two groups.

Results of All groups combined at Atherstone school

	Turn			Cut			Jig		
	A	B	G	A	B	G	A	B	G
Mean	8.8	9.9	7.5	6.5	6.5	6.5	5.5	5.7	5.3
Standard Deviation	5.2	5.2	4.8	2.5	2.6	2.3	1.9	2.2	1.5

	Fold			3D			Total		
Mean	5.1	5.3	4.8	8.1	9.0	7.0	33.9	36.5	31.0
Standard Deviation	2.1	2.3	1.9	3.2	3.3	2.9	10.8	11.8	8.8

A (N = 155)   B (N = 82)   G (N = 73)

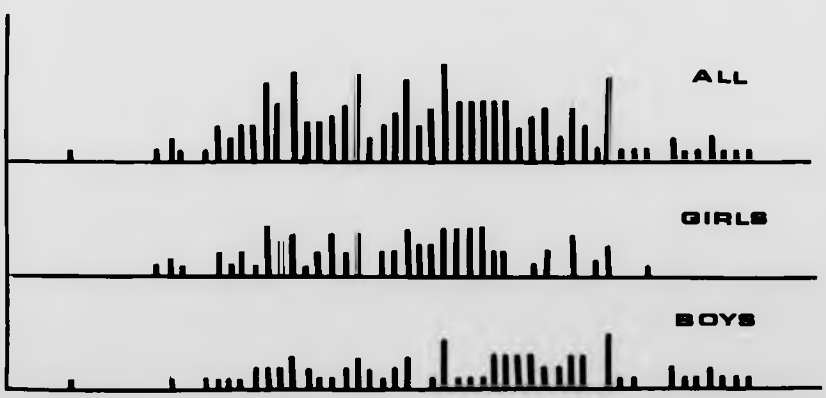
The difference between the means of boys and girls in the final total is significant at the $p < 0.01$ level.

Product - moment correlation values

	Turn	Cut	Jig	Fold	3D	Total
Turn	-	32	42	28	44	81
Cut	32	-	38	35	47	64
Jig	42	38	-	21	50	66
Fold	28	35	21	-	36	55
3D	44	47	50	36	-	77
Total	81	64	66	55	77	-

All values are significant at $p < 0.01$ level.

Score Distribution from Atherstone School



Atherstone top forms 1982

R1	Mean	Standard deviation	Number
Boys	41.0	8.4	N = 11
Girls	36.3	5.9	N = 13
G1			
Boys	39.3	8.7	N = 14
Girls	36.9	10.4	N = 14

In the following year (1983) this whole year group was tested again using the spatial test with the same times as the previous year. The results are as follows. Note the number of pupils is different from the first sitting as only pupils who took both tests can be included.

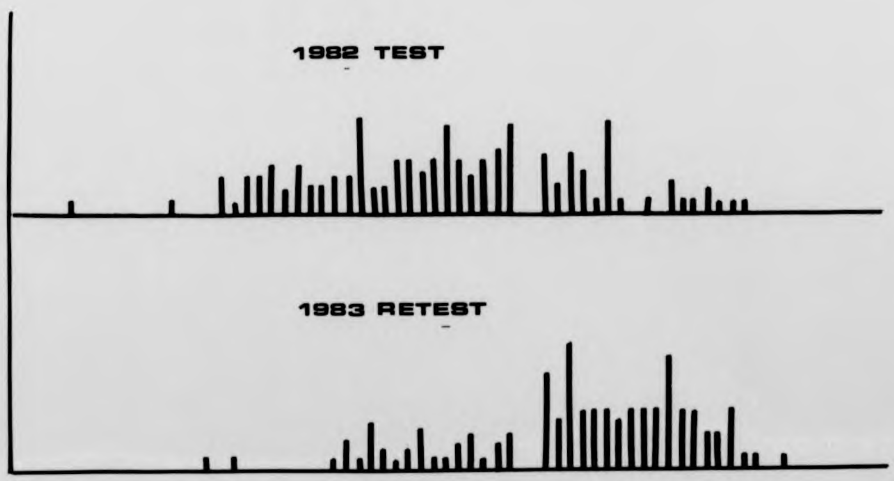
N = 135.

		Mean	Standard deviation
1983	Turn	12.5	4.4
	Cut	8.4	1.8
	Jig	6.9	1.8
	Fold	6.2	1.9
	3D	11.4	3.1
	Total	44.9	9.4
1982	Turn	9.11	5.0
	Cut	6.8	2.4
	Jig	5.9	2.1
	Fold	5.3	2.3
	3D	8.3	3.4
	Total	35.0	10.6

Product moment correlation for 1982/1983 test - retest

		1983						1982					
		Turn	Cut	Jig	Fold	3D	Total	Turn	Cut	Jig	Fold	3D	Total
1983	Turn	-	41	24	18	43	76	63	16	28	22	46	58
	Cut	41	-	30	30	36	51	40	46	30	26	38	43
	Jig	24	30	-	35	43	43	23	30	31	27	43	30
	Fold	18	30	35	-	35	39	42	43	48	45	47	48
	3D	43	36	43	35	-	73	39	37	45	27	69	60
	Total	76	51	43	39	73	-	60	33	36	25	64	79
1982	Turn	63	40	23	42	39	60	-	28	41	28	47	77
	Cut	16	46	30	43	37	33	28	-	36	39	47	54
	Jig	28	30	31	48	45	36	41	36	-	27	47	53
	Fold	22	26	27	45	27	25	28	39	27	-	34	47
	3D	46	38	43	47	69	64	47	47	47	34	-	74
	Total	58	43	30	48	60	79	77	54	53	47	74	-

Mark Distribution 1982 and 1983



Comparison of Boys and Girls on test retest

Girls N = 67

Year	Test	Mean	Standard deviation
1983	Turn	11.7	4.6
	Cut	8.5	1.3
	Fold	6.5	1.6
	Jig	5.6	1.8
	3D	10.3	3.1
	Total	42.6	8.3
1982	Turn	7.9	4.8
	Cut	6.8	2.3
	Fold	5.4	1.5
	Jig	4.9	2.0
	3D	7.2	3.0
	Total	32.0	8.6

Correlation matrix

		1983						1982					
		Turn	Cut	Jig	Fold	3D	Total	Turn	Cut	Jig	Fold	3D	Total
1983	Turn	-	23	10	21	31	76	60	1	11	26	40	58
	Cut	23	-	16	38	30	51	23	31	5	9	20	31
	Jig	10	16		31	32	46	8	22	9	12	25	23
	Fold	21	38	31	-	32	57	37	38	34	31	41	59
	3D	31	30	32	32	-	72	20	37	26	8	65	48
	Total	76	51	46	57	72	-	54	32	26	27	63	72
1982	Turn	60	23	8	37	20	54	-	13	21	20	28	78
	Cut	1	31	22	38	37	32	13	-	21	15	35	49
	Jig	11	5	9	34	26	26	21	21	-	1	30	46
	Fold	26	9	12	31	8	27	20	15	1	-	9	41
	3D	40	20	25	41	65	63	28	35	30	9	-	67
	Total	58	31	23	59	48	72	78	49	46	41	67	-

BOYS N = 68

Year	Test	Mean	Standard deviation
1983	Turn	13.1	4.1
	Cut	8.3	2.1
	Jig	7.4	1.8
	Fold	6.8	1.9
	3D	12.5	2.9
	Total	47.2	9.9
1982	Turn	10.3	4.9
	Cut	6.9	2.4
	Jig	6.4	2.5
	Fold	5.6	2.5
	3D	9.4	3.3
	Total	37.8	11.7

Correlation Matrix

		1983						1982					
		Turn	Cut	Jig	Fold	3D	Total	Turn	Cut	Jig	Fold	3D	Total
1983	Turn	-	58	31	7	50	75	62	31	36	16	47	57
	Cut	58	-	44	32	51	57	56	56	44	37	57	53
	Jig	31	44	-	29	43	33	26	38	37	33	49	27
	Fold	7	32	29	-	23	15	37	50	51	52	41	33
	3D	50	51	43	23	-	71	48	40	51	37	65	63
	Total	75	57	33	15	71	-	59	34	35	18	59	81
1982	Turn	62	56	26	37	48	59	-	41	48	29	55	75
	Cut	31	56	38	50	40	34	41	-	45	57	59	60
	Jig	36	44	37	51	51	35	48	45	-	35	50	50
	3D	16	37	33	52	37	18	29	57	35	-	45	46
	Fold	47	57	49	41	65	59	55	59	50	45	-	75
	Total	57	53	27	33	63	81	75	60	50	46	75	-

Results of 1983 Intake at Atherstone

		Mean	Standard Deviation
Turn	Boys	10.0	4.7
	Girls	8.2	4.5
	All	9.1	4.7
Cut	Boys	6.1	4.7
	Girls	5.6	2.4
	All	5.9	2.5
Jig	Boys	5.8	1.9
	Girls	5.2	2.0
	All	5.5	2.0
Fold	Boys	4.8	1.9
	Girls	4.8	1.6
	All	4.8	1.8
3D	Boys	8.4	3.9
	Girls	6.6	3.6
	All	7.5	3.8
Total	Boys	35.2	10.4
	Girls	30.49	9.9
	All	32.8	10.4

Difference total means of boys and girls $p < 0.01$

Correlation Matrix

	Turn	Cut	Jig	Fold	3D	Total
Turn	-	19	36	20	45	76
Cut	19	-	33	19	33	54
Jig	36	33	-	38	48	67
Fold	20	19	38	-	43	54
3D	45	33	48	43	-	81
Total	76	54	67	54	81	-

4th Year Atherstone 1983.   Top Chemistry Group

	Turn			Cut			Jig		
	A	B	G	A	B	G	A	B	G
Mean	13.3	13.8	12.7	8.3	8.3	8.2	6.8	6.9	6.7
Standard Deviation	3.4	2.9	4.1	1.6	1.6	1.7	1.4	1.2	1.7
	Fold			3-D			Total		
Mean	6.4	6.2	6.8	12.19	12.4	11.8	47.3	47.8	46.5
Standard Deviation	2.1	2.1	2.1	2.1	2.2	1.9	7.2	5.8	9.2

ALL (N = 49)   BOYS (N = 30)   GIRLS (N = 19)

Analysis in individual groups

	Group	Mean	Standard Deviation	Number
A	All	49.8	6.5	N = 24
	Boys	51.0	6.7	N = 12
	Girls	48.7	6.4	N = 12
B	All	45.2	7.5	N = 25
	Boys	38.7	7.6	N = 7
	Girls	47.8	5.9	N = 18

Correlation Matrix all

	Turn	Cut	Jig	Fold	3D	Total
Turn	-	21	51	30	44	81
Cut	21	-	07	07	25	43
Jig	51	07	-	44	28	64
Fold	30	07	44	-	23	56
3D	44	25	28	23	-	69
Total	81	43	64	56	69	-

Caludon Castle

This is an all boys school with a good academic standard in the City of Coventry. The top science groups were given the spatial test to try it out on older pupils. The results were compared with school performance in science (and mathematics for Year 3).

5th Year Chemistry at Caludon Castle

N = 25

	Mean	Standard Deviation
GCE Estimate*	21.6	6.9
Mock	63.2	13.1
Space	44.9	6.8

*A = 30
B = 20
C = 10

Correlation Matrix

	GCE	Mock	Space
GCE	-	86	-05
Mock	86	-	-14
Space	-05	-14	-

Caludon Castle 3rd Years

N = 30

	Mean	Standard Deviation
Space	39.6	9.3
Maths Test (1)	54.7	13.8
Maths Test (2)	60.8	14.6
Science Tests	65.9	14.1

Correlation Matrix

	Space	Maths (1)	Maths (2)	Science
Space	-	-32	-25	-30
Maths (1)	-32	-	87	69
Maths (2)	-25	87	-	59
Science	-30	69	59	-

Times for each section were as follows

Turn	3½ mins
Cut	2½ mins
Jig	5 mins
Fold	3½ mins
3D	3 mins.

Caludon Castle 1st Year (11+)

N = 55

	Science Test*	Attainment+	Space
Mean	64.7	3.3	38.5
Standard Deviation	11.6	1.3	8.9

* Science test - Total of 6 months progress tests on Scottish integrated science course.

\+ Attainment - Teacher assessment rated 1 to 5.

Correlation Matrix

	Science Test	Attainment	Space
Science Test	-	78	16
Attainment	78	-	24
Space	16	24	-

Campion School

This is a secondary school in the Warwick-Leamington area which has intake at 12+. The area is non selective but pupils do have a free choice of the school they attend. This may have an adverse effect on the level of intake ability. The spatial test was given along with a number of other tests in order to investigate this.

Campion School First Year January 1983

N = 165

	AH3	School Exam	Science Test	Data Analysis	Space	Word Assoc'n	Question-naire
Mean	99.2	51.2	21.8	20.2	27.0	40.3	26.2
Standard Dev.	10.9	21.8	8.0	6.3	9.6	16.6	9.2

Correlation Matrix

	AH3	School Exam	Science Test	Data Analysis	Space	Word Assoc'n	Question-naire
AH3	-	70	65	67	46	57	03
School Exam	70	-	68	73	48	50	-03
Science Test	65	68	-	69	47	60	06
Data Analysis	67	73	69	-	50	64	02
Space	46	48	47	50	-	41	05
Word Association	57	50	60	64	41	-	14
Questionnaire	03	-03	06	02	05	14	-

Comparison of Boys and Girls on Spatial test

	Boys	Girls
Mean	29.5	24.8
Standard Deviation	10.1	8.6
	N = 79	N = 86

Difference between means significant $p < 0.05$

Marks from Warwickshire Middle Schools 1982 June (Pupils 11+)

TELFORD	ALL	34.5	11.1	N = 84
	GIRLS	33.4	9.7	N = 40
	BOYS	35.4	12.3	N = 44
WOODLOES	ALL	28.3	9.9	N = 45
	GIRLS	27.8	8.4	N = 23
	BOYS	28.8	11.6	N = 22
BROOKHURST	ALL	33.8	13.7	N = 33
	GIRLS	28.3	12.7	N = 11
	BOYS	36.5	13.5	N = 22
BUDBROOK	ALL	33.5	11.2	N = 38
	GIRLS	31.5	10.5	N = 16
	BOYS	35.0	11.7	N = 22
TOTAL	ALL	32.8	11.3	N = 200
	GIRLS	31.0	10.3	N = 90
	BOYS	34.2	12.4	N = 110

Test retests from Junior Schools 1982 - Secondary 1983

	Mean	S. Deviation
1982	30.19	10.62
1983	35.22	10.59

Product moment correlation between scores (r) = 0.83

DISCUSSION OF RESULTS

Validity and Sex difference

Although the test was designed as one of spatial awareness and resembles other valid tests of this ability it was not correlated with any other spatial test. Due to this its present validity of being spatial is based on the definition of Camp (1981) which defines such tests as being those which require mental manipulation of presented diagrams and also show a sex difference producing a higher mean for boys than for girls.

All the results obtained do not show a clear cut advantage for boys in the younger age bracket. Those obtained from the Coundon Court group show no significant difference in the mean scores between the sexes. All groups who were tested in the original trials were in the first year of secondary education in Warwickshire schools and were a year older than the Coundon Court group the former having 12+ entry to secondary school and latter 11+ being in the Coventry Education area. Testing was carried out towards the end of the school year so that pupils would nearly all be 13+ in Warwickshire and 12+ in Coventry. All these older groups showed a sex difference in favour of boys ($p < 0.01$). This at first led one to suppose that the appearance of the sex difference in the older but not the younger group was due to their being further into adolescence. This pattern of males obtaining significant higher mean scores has been noted in all subsequent testing and along with the test design supports the existence of a spatial element in this test. A closer examination of the means for individual sections on the test shows that the sex difference is a feature of each and this contributes to the final total difference.

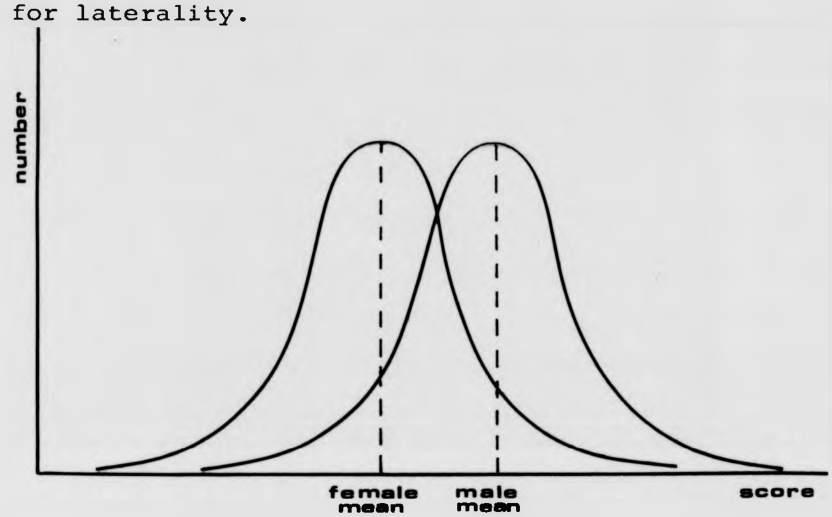
The original analysis of the Coundon Court and Atherstone (1982) scores appeared to be in agreement with the view from the literature that the sex difference becomes apparent during the adolescent period. This is in agreement with the younger group from Coundon Court displaying no significant sex difference and the Atherstone (1982) group showing one of 5.5 ($p < 0.01$). This contrasted with Lowes (1961), who in the development of a spatial test, found a sex difference in a group 10.8 to 11.2 years. These results were for an 'entire' city's year group and numbered 1045 pupils. Subsequently test results from some Warwickshire Middle schools also showed a sex difference and were not supportive of a developmental explanation as these pupils were in the same year as the Coundon Group. Although individual school differences were erratic the summation of these results showed a 3 point difference with the boys scoring the higher. Given these results the Coundon scores were examined more closely. At Atherstone the bottom two forms were not tested. The reason for this was that the testing was part of an identification process, which will be discussed in greater detail later and consequently did not take count of the lower forms.

For this reason the lower scores of the Coundon group were examined taking a cut off point of 20. There were fifty-two pupils in this group consisting of 20 boys and 32 girls. The scores of the candidates however showed that within this range the boys fell into the lower region of the distribution. This was shown to be the case when excluding the fifty-two produced a 2.9 sex difference in means. Similarly an examination of the scores above 40 showed twenty boys and fourteen girls. With the

Atherstone pupils twenty-one boys and five girls scored above 45.

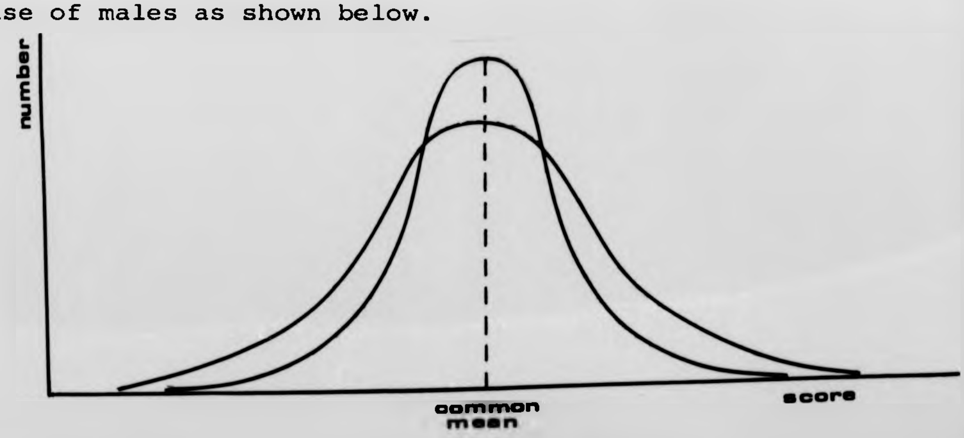
It appears then that these results do not support a spread of results with the means for the sexes offset as Blakeslee (1981) claims for laterality.

i.e.



These curves were estimated from growth curves and not from actual measurements. It is not meant to imply that laterality and spatial ability are the same, a point discussed earlier, rather the curves are used to illustrate a sex difference.

The results obtained are more indicative of the same average mean for males and females but with a greater standard deviation in the case of males as shown below.



This type of distribution is not unique to spatial ability. Taylor (1962) while not supporting the statement empirically says that men display a finer sense of perception experiencing both ends of the ability spectrum. Hendrickson and Hendrickson (1982) are also supportive in a BBC Today interview, which followed a paper in nature (Blinkhorn and Hendrickson (1982)) relating to evoked brain potentials and IQ measures (Ravens Progressive Matrices) stated that

"More males are mentally defective than women, at IQ 130+ there are twice as many males and at 175 the ratio is 1 to 55." (in favour of males) Eysenck and Kamin (1981) also discuss the question of a greater male variance in ability while debating the nature-nurture question.

It appears that there is no sex difference when considering the whole population but in higher ability groups, which is the main area of concern in this study, sex differences will be seen.

Atherstone school was an established grammar school which became comprehensive and this in an area where alternative schooling is not readily available. Coundon however is an urban school in an area where a substantial independant sector exists and results in some selection. Thus the higher ability pupils are likely to be missing from the population in the latter school resulting in no sex difference. In general one would expect sex differences to show because the lower end of the population will be removed from the general school population.

For the test to show further validity a factorial analysis would be required, using this test in a suitably chosen battery as has already been discussed.

Reliability - test-retest and Maturation

As the tests at Coundon Court and Atherstone were given the same timings it may have been possible to infer some maturational effects from differences in means. However in the light of the previous discussion this would be somewhat tenuous. Further testing at Atherstone and elsewhere does however provide some longitudinal results which are of interest regarding this. An early test-retest with a one year interval was carried out on the Banbury group studied by King (1982) but this was hampered by the availability of pupils and the fact that the form of the test was different so although it demonstrated a marked increase in performance by most pupils and in particular the girls, the number was too small to be of any significance. Observed improvement could be due to actual increase in processing ability caused by both experiential and Physiological maturation or it could be due to a carry over of experience. Inevitably using test-retest to demonstrate improvement is fraught with problems as it is difficult to differentiate these contributive factors.

In the case of this test as opposed to one with a large element of knowledge time would appear to exclude actual memory of answers, the familiarity of having taken the test being the carry over effect. Referring back to the evidenced activation levels between males and females this could provide an explanation for improvement in the girls as with familiarity an element of stress encountered at first sitting may well be reduced. The adverse effects of stress on women were noted by Wilson et al (1975).

That recall of items was improbable at a one year interval was concurred with from introspection by some of the pupils from the Atherstone 1982 group who took the test again in 1983. In this case timings, format and place of sitting of the test were identical. Inevitably some pupils had moved or were absent and consequently all did not take both sittings. However the number who did were sufficient to make the results significant. Reference to these results (Page104) shows that the mean of the group had increased by 10 points from 35.0 to 44.9
The distribution of the marks over the two years is shown for comparison. A further analysis of performance of boys and girls interestingly shows that the mean score for both has increased by the same amount so that the observed sex difference at the first sitting is again apparent which is not what one would have inferred from the Banbury results.

That the increase in scores is maturational is borne out by the setting of the test to two fourth year groups at Atherstone. The times of sitting were reduced to values given in the appendix but even so pupils scored highly at this first sitting. Although these groups were not representative of the school population, being top science groups, the number of high scores even at these faster times were well in excess of those shown by the whole population of second or third years. This increased speed cannot be put down to experience of the test.

The correlation between the test and retest scores was found to be 0.79 which indicates that although most pupils appeared to increase their scores it was those who came out best first time maintained this advantage. This is in spite of the fact that no change in test time was made leaving less room for those who scored well to improve. One might expect that the resultant closer packing of scores to the top end would produce a greater intermixing of rank order at second sitting than has in fact appeared. Certainly a number of high scoring pupils completed sections at first sitting and had no room to improve.

Further support for this result was obtained by a rather round-about route in that the spatial test was used in a Warwickshire secondary school to aid selection for the Science Summer School in 1983. A number of these pupils were in the middle school sample already referred to and had therefore been tested a year earlier. An increase of 5 points was seen on the mean score of the group but the number was too small to warrant a comparison of scores between the sexes. The correlation between the two scores was 0.83 and is in close agreement with that obtained from Atherstone. Comparison of results also shows similar mean scores and standard deviations for both groups in the first year of secondary school. This would be expected as the catchment areas of both schools are similar.

<u>Is the test indicative of scientific ability</u>?

What then of predictive validity? Is the test of use in the selection of able science students? As already discussed the immediate answer to this is no, as to give an affirmative reply one would require correlation of test performance in a longitudinal study at least to 'A' level and possibly beyond. There are however some positive answers that can be gained by examination of the results.

King (1982) found the marks on the spatial test were significantly better for her group, selected as able, than for her control group. There was however no significant correlation between this test and the WISC short form IQ test administered individually to the able group. This indicates the test is measuring something other than IQ. This however is not specific to science.

That the test is indicative of ability in science can be inferred from the results obtained from testing two Atherstone 4th Year top science groups. Here unfortunately the groups are relatively small and contain less girls than boys. The difference in mean scores, which are themselves high, between the sexes is small and is reduced to zero if the score of the lowest girl is ignored. One would not have inferred such a result by extrapolating the second and third year results. Inevitably one is left with some unknowns such as would the rest of the year group show no sex difference. This seems improbable from their presence in earlier years and the observations from the literature that male superiority would be expected to increase as puberty is reached. Although inferred the conclusion is that all who do science score well on the spatial test and that girls following science exhibit high spatial ability. There is always the lurking problem of nature - nurture in taking a group at a particular point in time as one does not know if the spatial ability has developed from the study of science or whether it preceded it. These science groups were undoubtedly of high ability so the possibility that no sex difference would be found within the top academic classes in previous years was investigated and this yielded a sex difference in favour of boys.

There is some further evidence for science being linked with spatial ability and this is given by intercorrelations in performance from two groups of pupils tested during their first year in Warwickshire secondary education i.e. 12+. The first group from Aylesford school numbered 53 (23 boys and 30 girls). They were given the spatial test which included pencil and belt at this time. The performance on this was compared with performance on NFER AH3 test and the percentage mark obtained in science over the school year. The correlation between AH3 and science marks were lower than between spatial and science and the AH3 and science also showed some correlation, though this was small enough to indicate the two tests were measuring a different pool of abilities.

A complete first year group from Campion School was given a number of tests by the science department at the school. There was some concern that the school was losing its higher ability pupils to other schools in the area, not because of private sector education but because of the availability of a choice as to which school pupils went. The results in this respect are not relevant to the present thesis but in passing it is noted that marks on the spatial test are lower than means established at other comparable schools. The tests consisted of AH3, school science examination results, two science tests designed for the purpose by the school science department (one testing recall and the other inference and reasoning), the spatial test, a word association test designed by Perryman (see Perryman and Purcell (1983a)) to look for ability in science and a science interest test based on that described by Ormerod and Wood (1983). The scoring on the latter was such that a positive response judged to be indicative in science was scored one and a negative one zero there was no degree of preference for pupils to select in response to the questions.

Reference to the correlation matrix shows the tests with the exception of the questionnaire correlated quite highly. The highest correlation was obtained with the years performance and AH3 result. All the science results correlated around 0.5 with the spatial test. The word association test and the spatial test have an intercorrelation of 0.41 and these results are in line with the use of the spatial test and word association test being jointly used as indicators of ability in science in that they both show correlation with that ability as measured by the tests set. It has to be accepted that the school marks may not be pointers to long term success in science but within the limits of short term study gives limited support for the use of spatial ability as an aid to measuring science ability.

Further testing to correlate school science performance with the spatial test was carried out with the first year (11+) intake at Caludon Castle School which is for boys only. The intercorrelations between the test and performance in school science was small.

When the third year and fifth year group were tested this pattern was repeated. Although too small in number to be of significance the chemistry 'A' level set of eight showed an exact correlation between estimated 'A' level grades and scores on the spatial test. These results giving low correlations are in agreement with Vernon (1972) who says that observed correlations are lower in selected groups. In this case all results, except that at Campion (where some correlation was found) are highly selected representing the top ability bands of secondary education. Thus the lack of correlation with such groups does not necessarily negate the view that the spatial test could be used as an indicater of scientific potential. It should be noted that all science groups at Caludon Castle scored highly, a mean of 44.9 and a standard deviation of 6.8 for the fifth year illustrates this and indicates that the time allowed for the test could be shortened further. What Vernon's observation means is that no test will be such a finely tuned instrument that it will discriminate between high ability pupils and at the same time provide an adequate test for the whole population. The latter was the original purpose of the spatial test.

So far the discussion has centred around group trends and statistical relationships. Within the limits of transverse study these have given support for the use of the spatial test as an aid in screening a group for ability in science.

Although not specifically seeking out scientific ability the test was used for screening purposes at Atherstone School to help select a group of more able second year pupils (12+). When selected certain provisions would be made for the group and it was hoped that its members would provide a nucleus from which school activities could grow.

The school had marks from the Middle schools which although they were all NFER Standardised tests the tests used were not common to all schools. In one case the scores were referred to as IQ scores which clearly they were not. These schools also provided a rank order based on school performance. Also available were the end of year performances in English and Maths although again the examinations taken were not common to all groups. The spatial test provided another indication of ability and was common to all groups.

Using these marks a short list was drawn up requiring outstanding performance on at least one measure, but preferably more, to qualify. This was then used by the school as a basis for further staff consultation across subjects and a final thirteen were picked. This omitted two pupils identified as gifted in middle school. Within the group of thirteen only two did not perform well on the spatial test. One of these was selected for high ability in English. The other while showing a degree of general ability has subsequently proved to be arts oriented. It should be said that even after the procedure selection was by no means confirmed in the ensuing year which affirms that while selection may be something achieved at a particular moment in time it is, in reality an ongoing process. There were of course pupils who scored highly on the spatial test and only longitudinal study will show whether or not they demonstrate ability in science.

One feature of this selection procedure and testing groups in other schools is that pupils have been brought to the teachers attention by the use of the spatial test who may otherwise have gone unnoticed. This aspect is something which cannot be conveyed by statistical analysis but nevertheless is seen as a further piece of evidence supporting the use of the test. This however is a subjective judgement and it could be argued that as such it should be omitted from a report of this nature. This leads on to the question of whether or not subjective judgement has a place in a scientific study and this question will be considered before drawing a final conclusion.

The question of subjectivity

The problem raised regarding subjective judgement is one that requires discussion as it is relevant to the use of tests and their general interpretation. While it is possible to produce a psychometric measure and standardise it there are still problems that surround the interpretation of marks.

Neimark (1979) discusses this stating that competence is not always related to performance as there are many other variables underlying test performance probably the most important being motivation. This has been brought out many times in the present study. The question most often asked by candidates is "Why are we being asked to do this test?" Although an adequate answer to this is given before starting it is not always possible to explain all reasons. This highlights the importance of motivation created by the existence of a needful goal which is the central feature of Skemp's (1979) theory.

Evidence indicates that pupils do try in this spatial test. This is probably helped by the fact that at this age a motivation to do well in school tasks exists across the population. Increased age may make lack of motivation more significant. One means of avoiding this is to select by voluntary means (e.g. Wilson et al (1975)) but the very words introduce an immediate bias in the sample. Success in a test then is indicative of motivation and ability and this must always be considered although the former is very difficult to assess.

The environment for the test is also an important factor. In administering the test in various surroundings it is immediately apparent that when the group is large in number pupils are more reticent to clarify any problems that might exist than in a smaller group when the atmosphere is less impersonal.

Other problems of measurement and variables have emerged in the discussion of validity and reliability of a test. These lie chiefly in the fact that generally the analysis depends on statistics and these can never give the full picture surrounding a course of testing such as has been undertaken in this study. On numerous occasions an analysis of results has shown pupils who are not performing well in school science but have done well on the spatial test. When examined in more detail there are often extraneous circumstances that could account for lack of correlation between performance on the spatial test and in science. For example the pupil has missed three months schooling or home background is not conducive to development at school.

For these reasons it is clearly not possible, and it would be unreasonable to expect, that a spatial test on its own will differentiate between the performance of high ability pupils but it is a valuable screening test for the whole population. When all the marks are eventually gathered together it is inevitable that subjective judgements are called for in any programme of identification,if they are not then a great deal is missing from the analysis.

Macquarie (1973) says

'It is the tendency of modern scientific humanism to regard the being of man as no different from the being of things and therefore amenable to study by scientific method. Of course that is not to deny the value of the social sciences (psychology and sociology) but it is to indicate their severe limitations.'

This then leads back to the problem regarding the nature of science already touched on when discussing Bergson's ideas and the claim that Piaget's clinical method cannot be regarded as scientific. Can scientific method be defined so that the position is clarified? To seek an answer to this it is necessary to consider the philosophical view-points.

Popper's view (Popper (1959)) is that for a theory to be scientific it must be possible to state it in such a form that it can be tested and refuted. It is not by confirmation of ideas that advancement is made but by failure which causes the theory to be modified. This view can be compared with Piaget's developmental theory where schemas are constantly being reorganised in the event of failure to cope with the environment and seems to have roots with the statement 'that people learn by their mistakes.'

There are problems associated with this view of science. Instances are seen when apparent refutation has not meant a theory being abandoned in the long term. Take for example the problems raised by wave-partical duality of light where theories that appear to be mutually exclusive do in fact flourish together.

The question of time dependant theories has already been mentioned. For example Darwin's Theory of Evolution cannot be tested so that it can be refuted. Does this mean that it is not a scientific theory? There does then seem to be a problem when applying Popper's view of science to biological science. This may eventually be resolved if the biological science can be reduced to physical in that all underlying mechanisms governing animal behaviour are understood by a process of reductionism. Until this is achieved (if this is in fact possible) then Popper's view is not sufficient to encompass all studies regarded as scientific. Conversely Popper's philosophy could be rigidly adhered to so that it excludes much of what is presently regarded as science.

An alternative view is held by Kuhn (1970) who tries to account for the actually observed advance of science through history where Popper's theory does not provide an adequate explanation. Science is seen as a problem solving exercise within the accepted structure and rules of the academic community. He accepts that while studies are in an immature phase they do arise from real world problems but as they develop the scientific community create their own working structures and rules known as Paradigms. When these are established progress of study can be made in isolation provided by the paradigm. Major advances in scientific thinking are seen to occur in revolutionary phases when paradigm shifts are forced upon the community because during this period competing paradigms develop, one eventually overthrowing the others because it is better able to account for observations than was the old paradigm.

Lakatos (1978) tries to combine Kuhn's historical development with Popper's more logical approach. He views science as progressing in a main paradigm but constantly associated with this are competing views which are continually being refuted and reappraised. These theories become self selecting in that they can be degenerative so that they are constantly having to be modified to account for new observation or progressive where they are self generating, spontaneously producing new theories.

There is then no real concensus of what is meant by scientific method. On these grounds then it is not essential for identification procedures to rely on tests alone and exclude subjective judgement. In fact the claims of psychometrics to be a science is argued against at some length by Evans and Waites (1981).

Thus it is argued that there is a place for some degree of subjective judgement at the end of a study of this type because this provides access to the individual which tests,no matter how well they are statistically trialled,can provide. To read into this study that spatial ability is ever going to provide a total means of identification would be wrong although it is evidenced that as a screening device it is useful to the practicing classroom teacher.

The individual selection at Atherstone has already been discussed and it was noted that this has not since proved totally satisfactory. The procedure adopted there was fairly exhaustive and was adopted as school policy and it is doubtful if this sort of programme could be contemplated by an individual teacher.

In an ideal world Martinson (1975) suggests an initial blanket screening followed by individual study as a means of identification. In detail this consists of:

1) Group IQ tests
2) Creativity tests
3) Teachers nominations
4) Data of school performance
5) Teachers notes on the traits of pupils
6) Parents notes on the traits of pupils.

When complete the pupils picked out would be given individual IQ tests and a case study would be produced on each pupil. Such a process needs to have an in-built on-going factor to account for differential maturation. Clearly this sort of programme holds little prospect of being implemented in a school because of the time it would take.

It is noted that this is looking for general ability. In looking for more subject specific ability, which they provided evidence to suggest did exist, Denton and Postlethwaite (1982) investigated the use of checklists which would be incorporated under 5) in the selection procedure just discussed. Although of use the method only worked satisfactorily when there was a second teacher in the classroom as a passive observer to complete the lists. In the real classroom this will not normally be possible because of staffing problems. If then any identification is attempted it is always going to fall short of ideal and for the individual teacher must mean a large degree of subjective judgement being used.

This is encouraged by Denton and Postlethwaite's observation that teacher judgement within subjects was better than they would have been led to believe from other such studies regarding the identification of general ability. They go on to say that where teachers fall down in identification in subjects such as physics is that they do not look for subject specific skills when seeking to identify pupils. Thus any help in this respect will be useful to the classroom teacher.

Conclusions and implications

It is re-emphasised at the start of this conclusion that the study was centred round the practical proposition of providing the practicing teacher with a means of screening a group of pupils to identify those who have potential to do well in science.

It is concluded from a survey of the literature that there is evidence to suggest subject specific abilities and that this supports the view that spatial ability provides one important indication of scientific potential. This has been supported empirically by the results which show little evidence for sex difference in older able science groups while a male advantage in such tests is demonstrated in all other facets of the testing.

Further measures of validity as an aid to identifying scientific ability are inaccessible to short term study and only individual longitudinal study could possibly answer this question. Even then there are problems regarding definition i.e. what is scientific ability? Denton and Postlethwaite (1982) used 'O' level results as a means of assessing ability. This however seems an inadequate measure especially when, as Smith (1964) suggests, it may not be until 'A' level studies that the true value of spatial ability begins to show. It would therefore seem that there is some justification for taking evidence from the literature that people who show long term success in the field of science display a certain profile as being indicative of the importance of spatial ability.

The problems of definition have been noted in the case of scientific ability and have to be considered again on the question of the test's validity to have a spatial factor. By adopting Camp's (1981) definition the test has been shown to be one of spatial ability on the grounds that it was designed as such and shows a significant sex difference between males and females. Further validity could be obtained by an exhaustive factorial study although the use of this method has been criticised by Evans and Waites (1981) and Skemp (1979).

That the test is reliable has been shown by a test retest carried out at a year's interval although it could be argued that such a test-retest would have greater value if carried out over a shorter time interval.

There is evidence from testing different age groups and from this test retest that there are considerable improvements made by pupils on spatial tests during adolescence supporting the findings of Kail, Pellegrino and Carter (1980). They observed that development of mental rotations continues into adolescence. It is noted that this is true for all the subtests of the spatial test.

The results do not support the findings that females improve on retest to the extent that sex differences disappear. On the contrary it was observed that sex difference remained after one year both sexes showing the same improvement. This leads to the conclusion that differences are caused by maturation and are not supportive of differential experience producing the sex difference.

This is an obvious area where longitudinal study should again provide some more information as to the effects of maturation at this time and could be evaluated in the light of the evidence that late maturers are better at spatial tests.

The spatial test provides a rapid means for screening a group. It is relatively free from language and knowledge bound content. It provides an objective measure of an ability which is not usually noted in general school performance. Because it can be attempted by a wide ability range it provides a useful addition to the knowledge that the classroom teacher may have of a group.

Caution must be taken in the use of such results, as indicated they may or may not have greater implications associated with them than might first be expected. The truth of this will only be established by a large number of wide ranging studies. No norms have been established for the test so that teachers will not be inclined to interpret cut off points as being exact taking no count of the error of measurement that is necessarily associated with any test mark. The intention is that the test will provide stimulus to the practicing teacher to investigate the groups of pupils they teach rather than providing them with a measure which they accept as being 'accurate'. They may then look for other methods of identification in the form of checklists or of using a more comprehensive test battery. It is only by the ongoing interest of teachers that a successful programme of identification can be implemented.

To this end the test around which this study has centred was published by the ESTEAM project (Perryman and Purcell (1982a)) and also includes the word association test referred to in the discussion of the Campion results which adds a further method of identifying pupil ability in science and is amenable to repeated use.

REFERENCES

ABRANAVAL, E. (1973) - Development of spatial awareness and representation in children: final report (EDO 80573) Washington: George Washington University Dept. of Psychology.

ARIB, M.A. (1972) - The Metaphorical Brain New York, London - Wiley Interscience.

ARLIN, P.K. (1975) - 'Cognitive development in adulthood: a fifth stage' Developmental Psychology Vol. 11. pp 602-606

ARCHER, J. (1976) - 'Biological explanation of Psychological Sex Differences' in Lloyd, B. & Archer, J. (eds.) (1976) Exploring Sex Differences London, New York, San Francisco: Academic Press

ARNHEIM, R. (1970) - Visual Thinking London: Faber & Faber Ltd.

AUSUBEL D.P., NOVAK J.D. and HARRISON H. (1968) - Educational Psychology - a cognitive view N.Y. Holt Reinhart and Winston.

BAKER, S.R. & TOLLEY, L.H. (1974). - 'Visualisation skills as a component aptitude for chemistry. A construct validation study' Journal of Research in Science Teaching, Vol. 11. No. 2 pp.95-97

BAMBER, J.H., BILL,J.M., BOYD,F.E. and CORBETT, W.D. (1983) - In two minds. Arts and science differences at sixth form level. British Journal of Educational Psychology Vol. 53 pp 222-233

BART, W.M. (1971) - 'The factor structure of formal operations' British Journal of Educational Psychology Vol. 41 pp 70-79

BART, W.M., BAXTER, J. & FREY, S. (1980) - 'The relationship of spatial ability and sex to formal reasoning capabilities' The Journal of Psychology Vol. 104 pp 191-198

BAUMANN, R. (1976) - 'Applicability of Piagetian theory to college teaching' Journal of College Science Teaching, Vol. 5 pp 94-96

BEARD, R.M. (1969) - An Outline of Piaget's Developmental Psychology London & Henley: Routledge & Kegan Paul

BECKMAN, L. (1977) - The use of block design Subtest as an identifying instrument for spatial children. Gifted Child Quarterly Vol 21 No. 1 pp 113-116

BERENBAUM, S.A. and RESNICK, S. (1982) - Somatic Androgeny and cognitive abilities Development Psychology Vol 18 No. 3 pp 418-423

BERG, C., HERTZOG, C., HUNT E. (1982) - Age difference in speed of mental rotation. Speed of processing governs rate on tests. Developmental Psychology Vol. 18 Part 1 pp 95-107

BERGLUND, G.W. (1965) - Mental growth, a study of test abilit between the ages of 9 and 16 years Scandanavian University books

BERRY, J.W. (1976) - Ecology and Cognitive Style Halstead Press John Wiley and Sons

BISHOP, J.E. (1978) - Developing Students Spatial Ability Science Teaching Vol 45 part 8 pp 20-23

BLAKESLEE T.R. (1980) - The Right Brain London, Basingstoke: MacMillan Press

BLINKHORN, S.F. and HENDRICKSON, D.E. (1982) - Average evoked response and psychometric intelligence. Nature Vol 295 pp 596-597

BOCK, R.D. & KOLAKOWSKI, D.D. (1973) - 'Further evidence of sex linked major gene influence on spatial visualisation ability' American Journal of Human Genetics Vol 25 pp 1-14

BODEN, M.A. (1979) - 'PIAGET' Brighton Sussex - Harvester Press

BOYLE, D.G. (1969) - A Student's Guide to Piaget London: Pergammon Press

BRADSHAW, J.L. & WALLACE, G. (1971) - 'Models for processing and identification of faces' Perception and Psychophysics Vol 9 pp 443-447

BRAINERD, C.J. (1978) - 'The stage question in cognitive development theory' The Behavioural and Brain Sciences Vol 2 pp 173-182

BRANDWEIN P.F. (1959) - The Gifted Student as a Future Scientist
New York: Harcourt Brace & Co.

BROVERMAN D., BROVERMAN K., VOGEL W. & PALMER R.D. (1964) - 'The automization cognitive style and physical development.
Child Development Vol. 35 pp. 1343-1359

BROVERMAN D., KLAUBER E.L., KOBAYASHI F., & VOGEL W. (1968) - Roles of activation and inhibition in sex differences in cognitive abilities' Psychological Review Vol. 75 pp. 23-50

BROVERMAN D., KLAUBER E.L. YOUTAKA K., & VOGEL W. (1969) - Reply to "comment" by Singer & Montgomery on 'Roles of activation and inhibition in sex differences in cognitive abilities' Psychological Review Vol. 76 No. 3 pp 328-331

BROWN G. & DEFORGES C. (1977) - 'Piagetian psychology and education: a time for revision' British Journal of Educational Psychology Vol. 47 part 1. pp.7-17

BROWN G. & DEFORGES C. (1979) - Piaget's theory ; a psychological critique
London & Henley: Routledge & Kegan Paul.

BUFFRAY A.W.H. & GRAY J.A. (1972) - 'Sex differences in spatial and linguistic abilities' in Ounstead, C. & Taylor C.D. (eds.) Gender Differences: Their Ontogeny and and Significance
Edinburgh & London: Churchill Livingstone

BUNGE M. (1980) - The Mind Body Problem: A psycho-biological approach
London: Pergammon Press

CAMP C.A. (1981) - Problem Solving Patterns in Science: Gender and Spatial ability during adolescence E.D.D. University, Massachusetts (Dissertation abstracts international KSN81-10311)

CIESLA J.L. (1976) - 'Comment on sex differences in problem solving ability' Journal of Research in Science Teaching Vol. 13 No. 6 pp. 573-74

COHEN C. (1973) - 'Hemispheric differences in serial versus parallel processing. Journal of Experimental Psychology Vol. 97, part 3 pp. 349-56

COHEN D. & SHELLEY D. (1982) - 'High I.Q. as high speed thinking' New Scientist 16th Sept. pp. 773-775

COOLEY W.W. (1958) - 'Attributes of Potential Scientists' Harvard Educational Review Vol. 28 pp. 1 - 18

CORBALLIS M.C. & MORGAN M.J. (1978) - 'On the biological basis of human laterality' The Behavioural and Brain Sciences Vol. 2. pp. 261-336

CRAIN W.C. (1980) - Theories of Development Concepts and Applications Englewood Cliffs, New Jersey: Prentice Hall Inc.

DALTON K. (1960) - 'Effects of Menstruation on School-girls' weekly work' British Medical Journal Vol. 1. pp. 326-28

DALTON K. (1968) - 'Menstruation and Examinations' Lancet Vol. 2. pp. 1386-138

DAVIS F.B. (1964) - Educational Measurements and their interpretation Belmont, Calif: Wadsworth Pub. Co. Inc.

DEFRIES J., ASHTON G., JOHNSON R., KUSE A., McCLEARN G., Mi M,, RASHAD M., VANDENBERG S. & WILSON J. (1976) - 'Parent offspring resemblance for specific abilities in two ethnic groups' Nature Vol. 261, pp. 131-33

DENTON C. & POSTLETHWAITE K. (1982) - The identification of more able pupils in Comprehensive Schools Parts I and II Oxford Educational Research Group for D.E.S.

DEPARTMENT OF EDUCATION AND SCIENCE (1977) - Gifted Children in Middle and Comprehensive Schools London: H.M.S.O.

DIELISLE J. & RENZULLI J.S. (1982) - 'The revolving door identification and programming correlates of creative production' Gifted Child Quarterly Vol. 26 part 2. pp. 89-95

DONALDSON M. (1978) - Children's Minds London: Fontana/Groom Helm

DOWNS R.M. & STEA D. (1977) - Maps in the Mind-Reflections on Cognitive Mapping
New York: Harper & Row Pubs.

DRESSEL P.L. & NELSON C.H. (1958) - Questions and problems in science Test item folio No. 1.
Princeton: Educational Testing Services

DRIVER R. in response to SHAYER M. in ARCHENBOLD W.F., DRIVER R.H., ORTON A., AND WOOD-ROBINSON C. (eds) (1980) - Cognitive Development Research in Science and Mathematics
Leeds: University of Leeds Press

ELIOT J. (1980) - 'Classification of figural spatial tests'
Perceptual and motor skills
Vol. 51 pp. 847-851

ELIOT J. & SMITH I.M. (1982) - Handbook of figural and spatial tests
(in press) London: N.F.E.R. Pub. Co.

EPSTEIN H.T. (1979) - 'Correlated brain and intelligence development in humans' in Hahn, M.E. Jensen, C. and Dudek, B.C. (Eds)
Development and Evolution of Brain Size Behavioural Implications
New York: Academic Press.

ESTEAM (1980) - Report of the 1979 Summer School for Warwickshire Schools
Institute of Education - University of Warwick

ESTEAM PROJECT (1983) - Science Enrichment materials published by Warwick University Department of Science Education

EVANS B. AND WAITES B. (1981) - I.Q. and Mental testing (an unnatural science and its social history)
London Basingstoke - The Macmillan Press Ltd.

EYSENCK H.J. (1967) - The Biological Basis of Personality
Springfield, Illinois, U.S.A.: Charles C. Thomas

EYSENCK H.J. & KAMIN L. (1981) - Intelligence: The Battle for the Mind
London: Macmillan Press Ltd.

FAIRWEATHER H. & HUTT C.J. (1972) - 'Sex differences in perceptual motor skills in children' in Ounstead T. and Taylor D.C. (Eds.) (1972)
Gender Differences: Their Ontogeny and Significance
Edinburgh & London: Churchill Livingstone

FAIRWEATHER H. (1976) - 'Sex differences in cognition' Cognition Vol. 4 pp. 231-280

FEHR H.F. (1953) - 'General ways to identify students with Science and Mathematical potential' Mathematics Teacher Vol. 46 pp. 230-234

FLAVELL J.H. (1963) - The Development Psychology of Jean Piaget Princeton, New Jersey: D. Van Nostrand Co. Inc.

FLEXER B.K. & ROBERGE J.J. (1983) - A longitudinal investigation of field dependance - Independance and the development of formal operational thought British Journal of Educational Psychology Vol. 53 pp. 195-204

FRALLEY J.S., ELIOT J. & DAYTON C.M. (1978) - 'Further study of the X linked Recessive Gene hypothesis for inheritance of spatial abilities' Perceptual and Motor Skills Vol. 47 pp. 1023-1029

FRENCH A.P. (Ed.) (1979) - Einstein - A centenary volume London: Heinemann

FURTH H.G. (1970) - Piaget for Teachers Englewood Cliffs, New Jersey: Prentice Hall Inc.

GEORGE W.C. (1979) - 'The talent search concept: an identification of the intellectually gifted.' Journal of Special Education Vol. 13 part 3, pp 221-37

GEORGE F.H. (1980) - Problem Solving London: Gerald Duckworth & Co. Ltd.

GETZELS J.W. & JACKSON P.W. (1962) - Creativity and Intelligence (Explorations with Gifted Students) London: J. Wiley & Sons Inc.

GOLDBERG J. & MEREDITH W. (1975) - 'A longitudinal study of spatial ability Behavioural Genetics Vol. 5 pp 127-135

GOLDSTEIN A.G. & CHANCE J.E. (1965) - 'Effects of practice on sex related differences in performance on embedded figures task' Psychonomic Science Vol. 3 pp 361-362

GOURLAY N. (1979) - 'Heredity v Environment and Integrative Analysis' Psychological Bulletin Vol. 86 part 3 pp. 596-615

GRAYBILL L. (1975) - 'Sex differences in formal thought' Journal of Research in Science Teaching Vol. 12 part 4 pp 341-346

GUILFORD J.P. (1950) - 'Creativity' American Psychologist Vol. 5 pp 444-454

GUILFORD J.P. (1967) - The Nature of Human Intelligence New York: McGraw Hill Book Co.

GUILFORD J.P. (1969) - 'The three faces of intellect' in Wolfe D. (Ed) Discovery of Talent Cambridge, Mass: Harvard Univ. Press

GUILFORD J.P. & FRUCHTER B. (1973) - Fundamental Statistics in Psychology and Education New York: McGraw Hill Book Co.

HARRISON G.A., WEINER J.S., TANNER J.M., & BARNICUT N.A. (1977) - Human Biology Oxford University Press

HARTLAGE L.C. (1970) - 'Sexed linked inheritance of spatial ability' Perceptual and Motor Skills Vol. 31 pp. 610

HAUSER S.T. (1976) - 'Loevinger's model and measure of ego development: a critical view' Psychological Bulletin Vol. 83 part 5 pp 928-955

HEAD J. (1980) - 'Personality and Cognitive Development in Archenbold W.F., Driver R.H., Orton A. and Wood Robinson c. (Eds.) (1980) Cognitive Development Research: Science and Mathematics Leeds: University of Leeds Press

HEAD J. (1982) - 'What can Psychology contribute to Science Education?' School Science Review Vol. 63 No.225 pp 631-42

HEIMS A.W., WATTS K.P., & SIMMONDS V. (1972) - Shapes Analysis Test Manual High Wycombe: The Test Agency

HENDRICKSON E. & HENDRICKSON A. (1982) - B.B.C. Transcript of interview with Hugh Sykes on Today Programme 19.2.1982

HIER D.B. & CROWLEY W.F. (1982) 'Spatial ability in Androgen deficient men' The New England Journal of Medicine Vol. 1306 Part 20 pp 1202-1205

HOWS R. (1980) - 'Spatial Ability, Sexism and Society' Physics Education Vol. 15 pp 229-30

HUGHES E.R. (1979) - Conceptual powers of children: An approach through Mathematics and Science London & Basingstoke: Macmillan Education Ltd.

HUGHES A.G. & HUGHES E.H. (1937) - Learning and Teaching London, N.Y., Toronto Longmans Green & Co.

INHELDER B. & PIAGET J. (1958) - The Growth of Logical thinking from childhood to adolescence New York: Basic Books

KAIL R., PELLEGRINO J. & CARTER P. (1980) - Developmental stages in mental rotation Journal of Experimental Child Psychology Vol 29 pp 102-116

KARPLUS R., KARPLUS E. FORMISORO M., & PAULSEN A.A. (1977) - 'A survey of proportional reasoning and control of variables in seven countries' Journal of Research in Science Teaching Vol 14 part 5 pp 411-17

KATCHADOURIAN H. (1977) - Biology in Adolescence San Francisco: W.H. Freeman & Co.

KING L. (1981) - An investigation into the characteristics and progress of a group of first year comprehensive school children identified on transfer as intellectually more able Unpublished M.Sc. Thesis. Department of Science Education - University of Warwick.

KOHEN-RAZ R. (1977) - Psychobiological aspects of cognitive growth New York: Academic Press

KRAFT R.H. (1977) - 'An E.E.G. Study of hemispheric brain functioning of 6 to 8 year old children during Piagetian and curriculum tasks with variation in presentation mode' (Temple University 1977) Dissertation Abstracts International 37 (9) 5587-8 March 1977

KUHN T. (1970) - The Structure of Scientific Revolutions University of Chicago Press

LAKATOS I. (1978) - Methodology of Scientific research programmes
Eds. Worral J. and Currie G.
Cambridge University Press

LANGER J. (1969) - Theories of Development
New York: Holt, Reinhart & Winston Inc.

LAWSON A.E. & NORLAND F.H. (1976) - 'The factorial structure of Piagetian Tasks'
Journal of Research in Science Teaching
Vol 13 part 5, pp 461-466

LAWSON A.E., KARPLUS R. & ADI H. (1978) - 'The acquisition of propositional thought and formal operational schemata during secondary school years'
Journal of Research in Science Teaching Vol 15 part 6 pp 465-478

LAWSON A.E. (1980) - Relation among level of development, Cognitive style and grades in College Biology Course.
Science Education Vol 64 No 11 pp 95-102

LESSER G.S., DAVIS F.B. & NAHEMOW L. (1962) - 'The identification of gifted elementary school children with exceptional scientific talent'
Educational and Psychological Measurement Vol 22 part 2 pp 349-364

LEVINE D.I. & LINN M.C. (1977) 'Scientific reasoning ability in adolescence: Theoretical viewpoints and educational implications'
Journal of Research in Science Teaching Vol 14 No 4 pp 371-384

LEVY J. (1976) - 'Cerebral lateralisation and spatial ability'
Behaviour Genetics Vol 6 part 2 pp 171-188

LEWIS D.G. (1967) - 'Ability in science at ordinary level of the general certificate of education' British Journal of Educational Psychology Vol 37 pp 361-370

LINDSAY A.D. (1911) - The Philosophy of Bergson
London: J.M. Dent & Son Ltd.

LINN M.C. & THIER H.D. (1975) - 'The effects of experimental science on the development of logical thinking' Journal of Research in Science Teaching Vol 12 pp 49-62

LINN M.C. (1980) - When do adolescents reason? European Journal of Science Education Vol 2 No. 4 pp 429-440

LJUNG BERGT-OLOV (1965) - The Adolescent Spurt in Mental Growth Stockholm: Almquist & Wiksell

LLOYD B. & ARCHER J. (1976) - Exploring Sex Differences London, New York, San Francisco: Academic Press

LOVELL K. & BUTTERWORTH I.B. (1966) - 'Abilities underlying the understanding of proportionality Mathematics Teaching Vol 37 pp 5-9

LOVELL K. & SHIELDS J.B. (1967) - Some aspects of a study of the gifted child' British Journal of Educational Psychology Vol. 37 pp 201-208

LOWES J.S. (1961) - 'The construction and validity of a spatial test British Journal of Educational Psychology Vol 31 pp 297-299

LUNZER E.A. (1968) - 'Formal Reasoning' in Lunzer, E.A. and Morris J.F. (Eds) (1968) Developments in Human Learning London: Staple Press

MACCOBY E.M. & JACKLIN C.N. (1975) - The Psychology of Sex Differences London: Oxford University Press

MACQUARRIE J. (1973) - An Existentialist Theology Pelican Press

MANDELL M.M. & ADAMS S. (1948) 'Selection of physical scientists' Educational and Psychological Measurement Vol 8 pp 575-581

MARTINSON R.A. (1975) - The identification of the gifted and talented Reston Virginia - Council for exceptional children

McGEE M. (1979a) - Human Spatial Abilities, Sources of Sex Differences London: Praeger Pubs.

McGEE M. (1979b) - 'Human Spatial Abilities: Psychometric Studies and environmental, genetic and neurological influences' Psychological Bulletin Vol 86 pp 889-918

MESSANT R.R. (1976) - 'Female hormones and behaviour' in Lloyd, B. and Archer, (1976) Exploring Sex Differences New York: Academic Press

MILLER G.A. (1956) - 'The magical number seven plus or minus two some limits on our capacity for processing information' Psychological Review Vol 63 pp 81-97

MONEY J. & ALEXANDER D. (1966) 'Turner's syndrome: Further demonstration of the presence of special cognitive deficiencies' Journal of Medical Genetics Vol 3 pp 42-48

MULAIK S.A. (1972) - The Foundations of Factor Analysis New York: McGraw-Hill

NEIMARK E.D. (1975) - 'Intellectual development during adolescence' in Horowitz F.D. (Ed.) Review of Child Development Research Vol 4. University of Chicago Press

NEIMARK E. (1979) - Confounding with cognitive factors: an artefact explanation for the apparent non universal incidence of formal operations thought

in

SIGAL I., GOLLINKOFF R. & BRODZINSKY D. (Eds.) - Piagetian theory and research: New Directions and Applications

Hillside N.J.: Erlbaum

NEWCOMBE N. & BANDURA M.N. (1983) - Effect of Age at puberty on spatial ability in girls: A question of Mechanism Developmental Psychology Vol 19 No. 2 pp 215-224

N.F.E.R. (1981) - Differential Aptitude Tests Windsor: NFER Publishing Co. Ltd.

OLSON M.B. (1977) - 'Right or left hemispheric information processing in gifted students' Gifted Child Quarterly Vol. 21 part 1. pp 116-121

ORMEROD M.B. AND WOOD C. (1983) - A comprehensive study of three methods of measuring attitudes to science. 10-11 year old pupils. European Journal of Science Education Vol 5 No. 1 pp 77-86

PASCUAL-LEONE J. & GOODMAN D. (1979) - 'Intelligence and experience: A neopiagetian approach' Instructional Science Vol 8 pp 301-367

PEEL A.E. (1967) - The Pupil's Thinking London : Old Browne Book

PERRYMAN M.R. AND PURCELL C.J. (1983a) Testing Ability in Science Department of Science Education University of Warwick - Esteam Project

PERRYMAN M.R. AND PURCELL C.J. (1983b) Report on 1982 Science Summer School - Science and Agriculture Department of Science Education University of Warwick - Esteam Project

PETERSON A.C. (1976) - 'Physical Androgeny and cognitive functioning in Adolescence' - Developmental Psychology Vol 12, part 6 pp 524-533

PIAGET J. (1950) - The Psychology of Intelligence London: Routledge & Kegan Paul

PIAGET J. & INHELDER B. (1956) - The Child's Conception of Space London: Routledge & Kegan Paul

PIAGET J. & INHELDER B. (1958) - The Growth of Logical Thinking London: Routledge & Kegan Paul

PIAGET J. (1971) - Structuralism London - Routledge and Kegan Paul

PIAGET J. (1972) - Insights and Illusions of Philosophy London: Routledge & Kegan Paul

PIBURN M. (1980) - 'Spatial reasoning as a correlate of formal thought and science achievement for New Zealand students' Journal of Research in Science Teaching Vol 17 part 5 pp 443-448

POPPER K.R. (1959) - The logic of Scientific Discovery London: Hutchinson

REDGROVE J.A. (1971) - 'Menstrual cycles' in Colquhoun W.P. Biological Rhythms and Human Performance New York: Academic Press

RENNER J.W. & LAWSON A.E. (1975) - 'Relationships of science subject matter and developmental level of learners' Journal of Research into Science Teaching Vol 12 part 4 pp 347-358

RENNER J.W. (1976) - 'Significant physics content and intellectual cognitive development as a result of interacting with physics content' American Journal of Physics Vol. 44 pp 218-222

RICHARDSON M.W. AND KUDER G.F. (1939) - The calculation of test reliability coefficients based upon the method of rational equivalence Journal of Educational Psychology Vol 30 pp 681-687

ROE A. (1952) - 'A psychologist examines sixty-four eminent scientists' Scientific American Vol 187 pp 21-25

ROSE S. (1973) - The Conscious Brain London: Winfield & Nicolson

ROVET J. & NETLEY C. (1982) - Processing deficiencies in Turner's Syndrome Developmental Psychology Vol 18 No. 1 pp 77-94

ROTMAN B. (1977) - Jean Piaget: Psychologist of the Real London: Harvester Press Ltd.

RUBENZER R. (1979) - The role of the right hemisphere in learning Gifted child quarterly Vol 23 Pt 1 pp 78 - 100

SALKIND N.J. (1981) - Theories of Human Development New York: Van Nostrand Co.

SCIENCE EDUCATION (1979) - Summary of Research in Science Education 1977 by PETERSON R.W. and CARLSON G.R. Science Education Vol 73 pp 425-548

SCIENCE EDUCATION (1980) - Summary of Research in Science Education 1978 by GABEL D.L. KAGAN M.H. and SHERWOOD R.D. Science Education Vol 74 pp 429-568

SCREEN P.A. (1981) - ESTEAM a curriculum development for exceptionally able children in science School Science Review Vol 63 No. 223 pp 350-355

SHAYER M. (1970) - 'How to Assess Science Courses' Education in Chemistry Vol. 7 No. 5 pp 182-184

SHAYER M. (1972) - 'Conceptual demands in the Nuffield O-level physics course' School Science Review No 186 pp 26-34

SHAYER M. (1976) - 'Development of thinking of middle and secondary school pupils' School Science Review No. 200 pp 568-571

SHAYER M. (1979) - 'Has Piaget's construct of formal operational thinking any utility?' British Journal of Educational Psychology Vol 49 pp 265-276

SHAYER M., WYLAN H., ADEY P. & KUCHEMANN D. (1981) - CSMS Science Reasoning Tasks Windsor: Test Dept. NFER Publishing Co.

SHELDON R., BAKER S.R. & TOLLEY L. (1972) - 'The relation of visualisation skills to achievement in freshman chemistry' Journal of Chemical Education Vol 49 part 11 pp 775-776

SHEPARD R.N. & METZLER J. (1971) - 'Mental rotations of 3-D objects' Science Vol 171 pp 701-703

SHEPARD R.N. & FENG C.A. (1972) 'Chronometric studies of mental paper folding' Cognitive Psychology Vol 3 pp 228-243

SIEMANKOWSKI F.T. & MACKNIGHT D. (1971) - 'Spatial cognition: Success prognosticator in college science courses' Journal of Research in Science Teaching Vol 8 part 1 pp 56-59

SINGER G. & MONTGOMERY R.B. (1969) - 'Comments on Roles of Activation and Inhibition in sex differences in cognitive abilities' Psychological Review Vol 76 part 3 pp 325-327

SKEMP R.R. (1979) - Intelligence, Learning and Action Chichester: John Wiley & Sons

SMITH I.M. (1964) - Spatial Ability - its Educational and Social Significance University of London Press

SMITH I.M. (1972) in Heim A.W., Watts K.P. and Simmons V. Shapes Analysis Test: Manual of Instructions High Wycombe, The test Agency

SMITH W.S. & LITMAN C.I. (1979) - 'Early adolescent boys and girls learning of a spatial visualisation skill' Science Education Vol 63 part 5 pp 671-676

SPRINGER S.P. & DEUTSCH G. (1981) - Left Brain Right Brain San Francisco: W.H. Freeman & Co.

STAFFORD R.E. (1961) - 'Sex differences in spatial visualisation as evidence of sex linked inheritance' Perceptual and Motor Skills Vol 13 p. 428

STILLMAN A.B. (1982) - 'The Rationale for abilities testing and specifically the development of a scientific ability test' School Science Review Vol 63 No.224 pp 423-433

SUPER D.E. & BACHRACH P.B. (1957) - Scientific careers and vocational development theory New York: Teachers' College, Columbia University

TANNER J.M. (1962) - Growth at Adolescence Oxford & Edinburgh: Blackwell Scientific Publications

TAYLOR F.A. (1962) - Colour Technology London: Oxford University Press

TAYLOR-PARKER S. (1978) - 'Critique of Brainerd' Behavioural and Brain Sciences Vol 2 pp 199-200

TERMAN L.M. & ODEN M.H. (1947) - The Gifted Child Grows Up: 25 year follow-up of a superior group
Stanford, California:
Stanford University Press

THOMSON G.H. (1939) - The Factorial Analysis of Human Ability
London: London University Press

THURSTONE L.L. (1947) - Multiple Factor Analyses
Chicago & London:
University of Chicago Press

TYLER L.E. (1963) - Tests and Measurements
Englewood Cliffs, New Jersey:
Prentice Hall Inc.

VERNON P.E. (1950) - The Structure of Human Abilities
London: Methuen

VERNON P.E. (1972) - 'The Distinctiveness of field independance'
Journal of Personality
Vol 40 pp 366-391

VAN DER VLUGT H. (1979) - 'Aspects of normal and abnormal neurophysiological development'
in Gazzaniga M.S. (Ed.)
Handbook of Behavioural Neurobiology Vol. 2
Neuropsychology
New York:

VUYK R. (1981) - An Overview and Critique of Piaget's Genetic Epistemology 1965-80 Vols 1 & 2
London: Academic Press (Inc.)
London Ltd.

VYGOTSKY L.S. (1962) - Thought and Language
Massachusetts: M.I.T. Press

WABER D.P. (1976) - 'Sex differences in cognition: a function of maturation'
Science Vol 192 pp 572-574

WABER D.P. (1977) - 'Sex differences in mental abilities, Hemispheric lateralisation and rate of physical growth in adolescence
Developmental Psychology
Vol 13 part 1 pp 28-38

WABER D.P. (1980) - 'Maturation: Thoughts on renewing an old acquaintanceship' in
Caplan D. (Ed.) Biological Studies of Mental Processes
Massachusetts: M.I.T. Press

WEBB R.A. (1974) - 'Concrete and formal operations in very bright 6-11 year olds' Human Development, Vol 17 pp 292-300

WISC (1976) - WECHSLER D. - WISC-R. Manual Wechsler Intelligence Scale for children, revised, containing British amendments to the administration and scoring criteria sections by NFER 1976 Anglicised version 1974 American

WEST L.H.T. & FERSHAM P.J. (1974) - 'Prior knowledge and the learning of science. A review of Ausubel's theory of this process' Studies in Science Education 1 pp 61-81

WILSON J.R. DE FRIES J.C. McLEARN G.E. VANDENBERG S.G. JOHNSON R.C. & RASHAD M.N. (1975) - 'Cognitive abilities: Use of family data as a control to assess sex and age differences in two ethnic groups' International Journal of Aging and Human Development Vol 6 part 3, pp 261-276

WITKIN H.A. DYK R.B. PATERSON H.F. GOODENOUGH D.R. & KARP S.A. (1962) - Psychological Differentiation: Studies of development New York: John Wiley

WITTROCK M.C. (1977) - The Human Brain New York: Prentice Hall

WOOD R. (1976) - 'Sex differences in Mathematics at G.C.E. ordinary level Educational Studies 2:2 June pages 141-160

WOOD R. (1977) - 'Cable's comparison factor: is this where girls' trouble starts' Maths in School 6:4 Sept. pp 18-21

YATES A. & PIDGEON D. (1957) - Admission to Grammar Schools London: Newnes for N.F.E.R.

YOUNG J.Z. (1978) - Programmes of the Brain Oxford University Press

ZIV A. (1977) - Counselling the Intellectually Gifted Child Toronto: Guidance Centre Faculty of Education, Toronto University

APPENDIX A

TURNABOUT

In the groups of 5 figures shown below one is the mirror image of the others.   This is the odd one out.

Select the odd figure and put its letter in the square on the answer sheet.

Two examples are given

a)

A B C D E

All except D will be  if they are turned round.

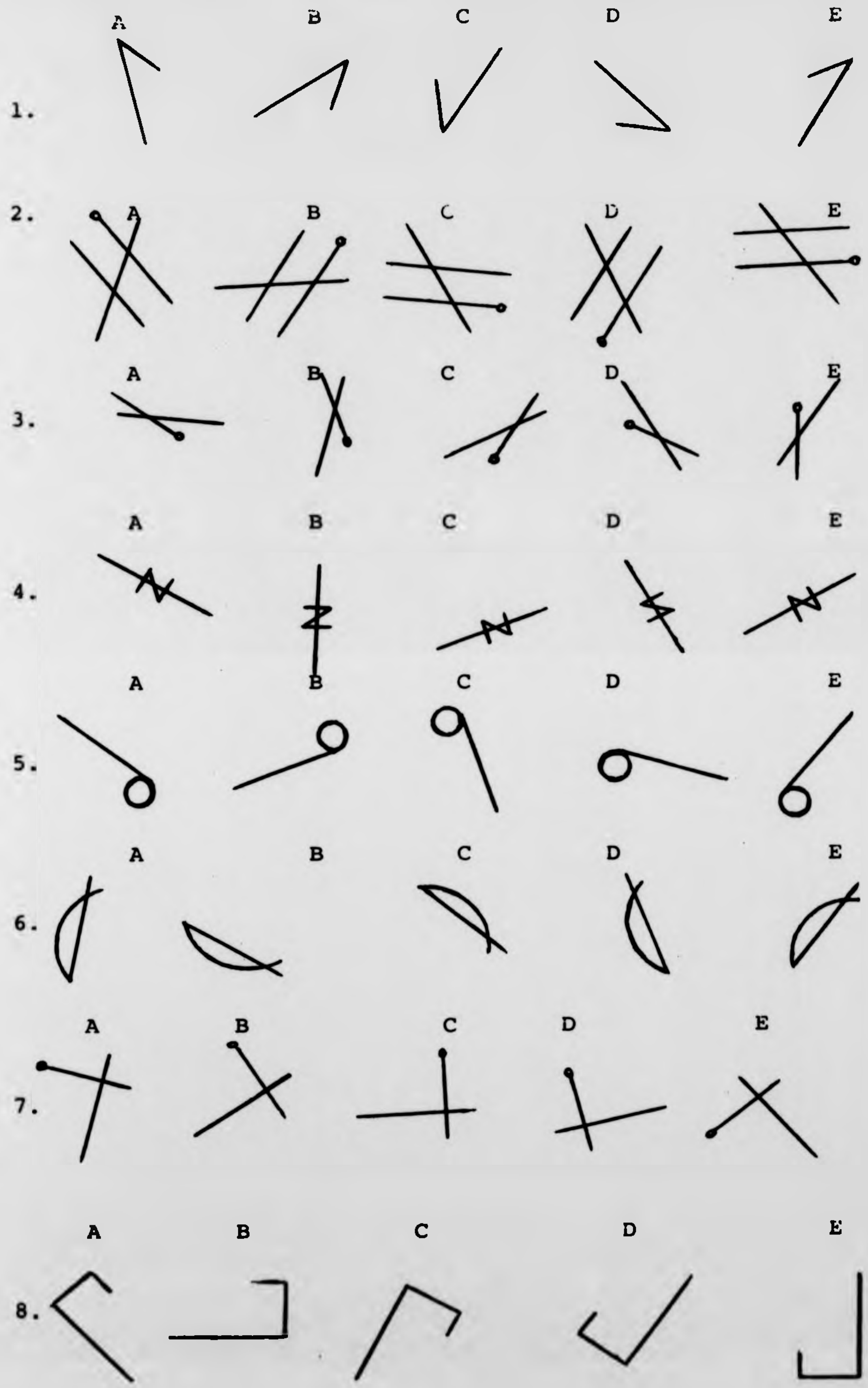
This will be 

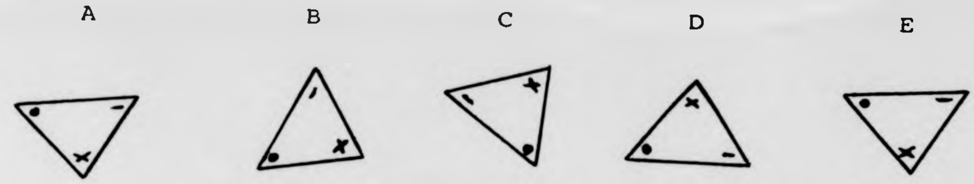
b) Another example

A B C D E

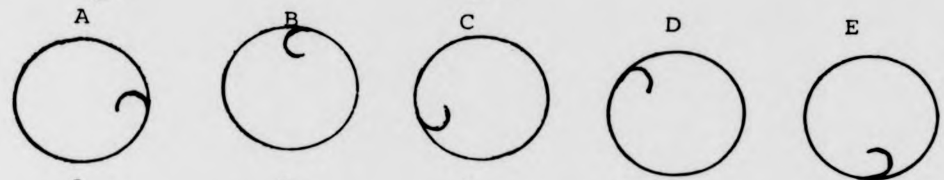
Answer

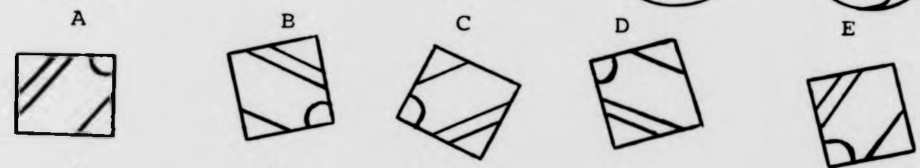
a	D
b	


A B C D E
1.
2. A B C D E
A B C D E
3.
A B C D E
4.
A B C D E
5.
A B C D E
6.
A B C D E
7.
A B C D E
8.

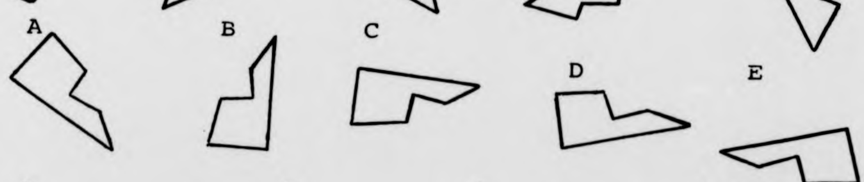
9. 

10. 

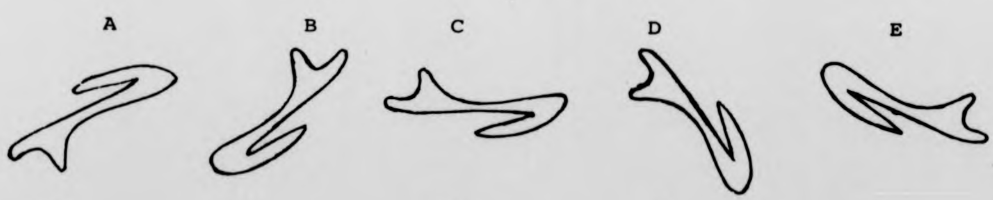
11. 

12. 

13. 

14. 

15. 

16. 

17. 

End of Turnabout - <u>DO NOT TURN OVER</u>

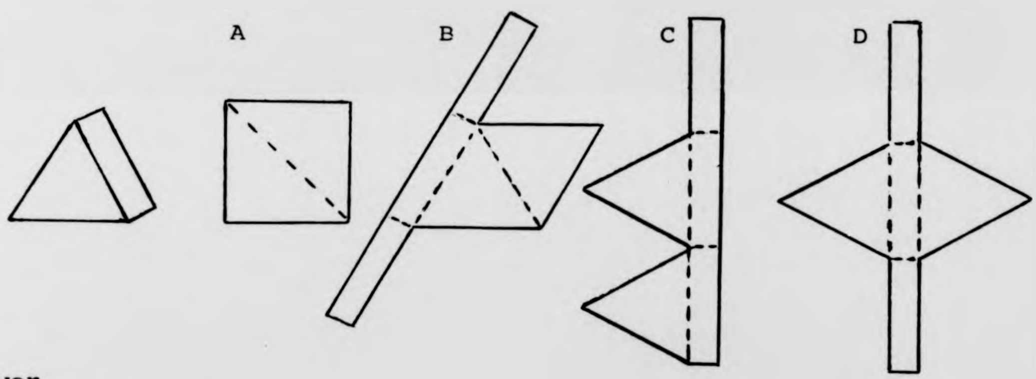
BLANK PAGE IN ORIGINAL

CUTOUTS

The first figure is a picture - a three dimensional figure made of cardboard. The four drawings on the right are drawings of the figure before the cardboard was folded along the dotted lines. Only one will make the first figure. Select the letter and put it into the square in the box on the answer sheet.

An example is given below

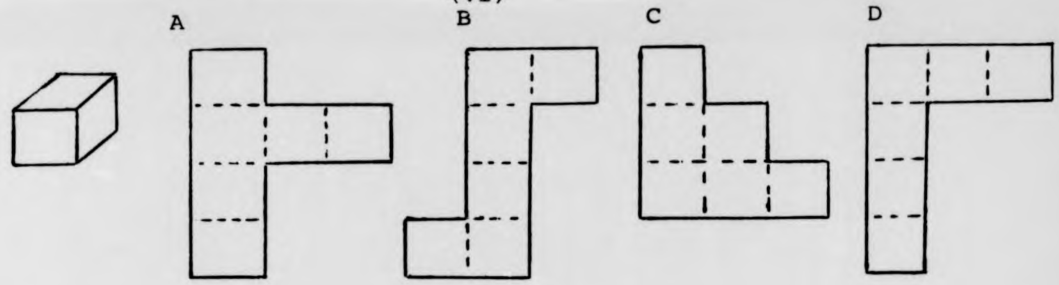
a) A B C D



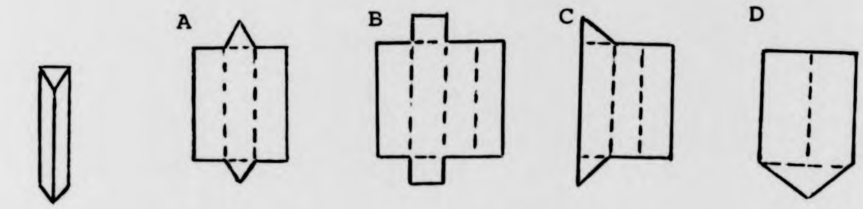
Answer

a

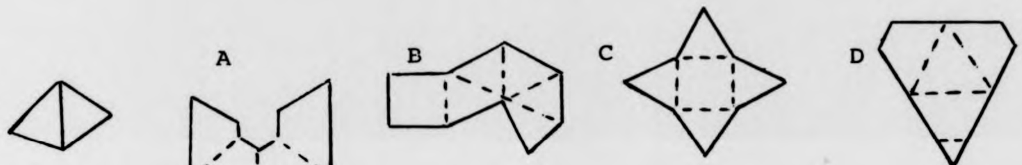
1.



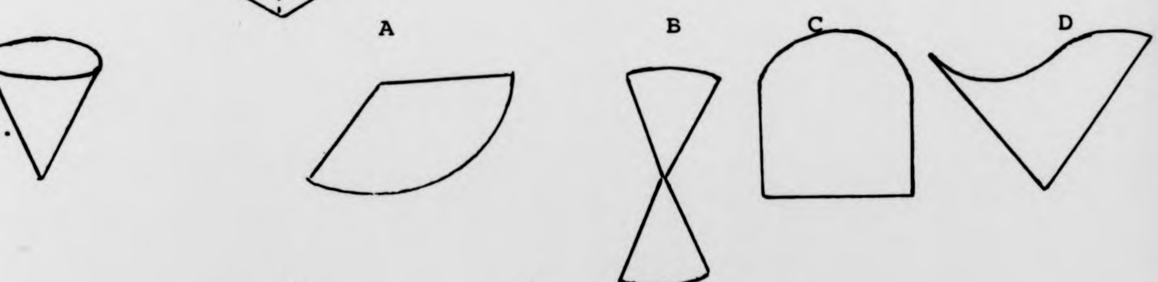
2.



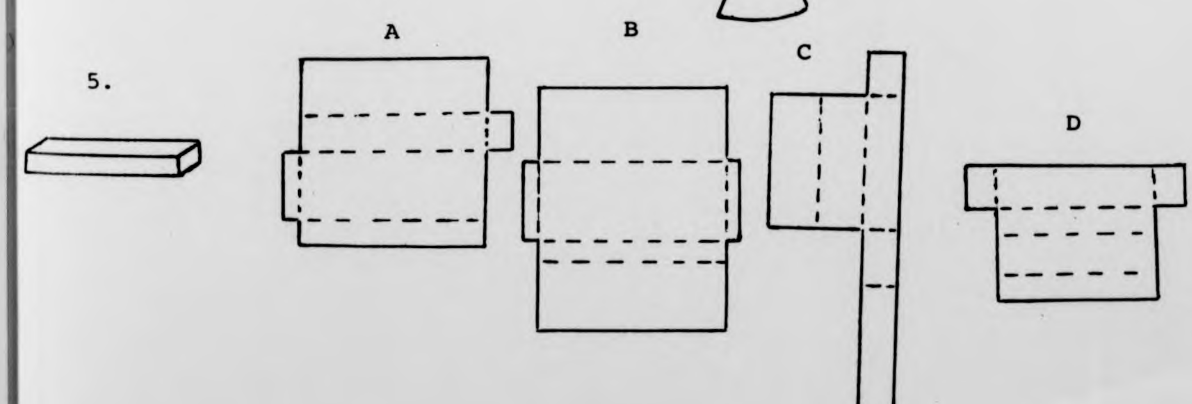
3.



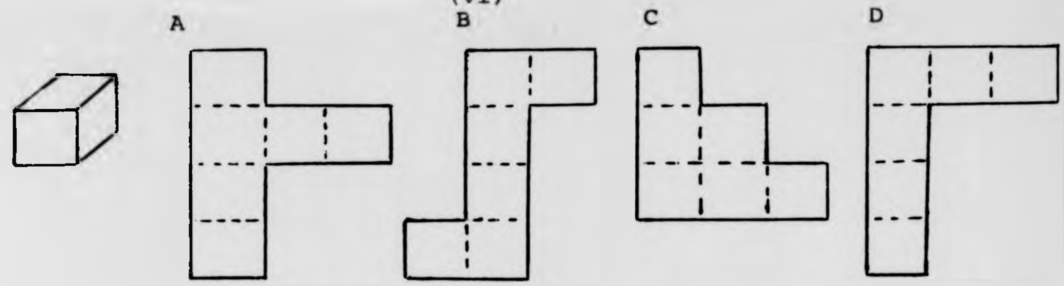
4.



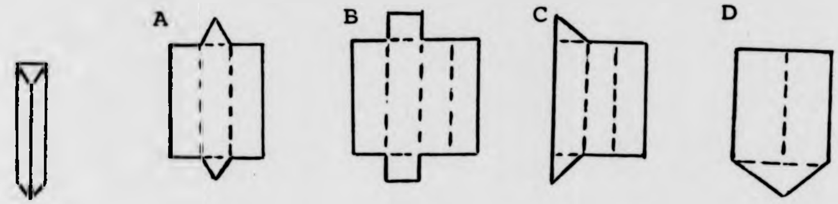
5.



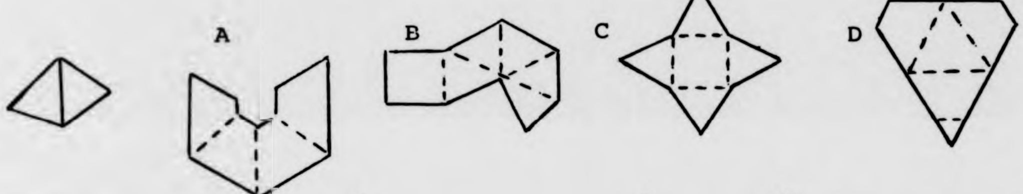
1.



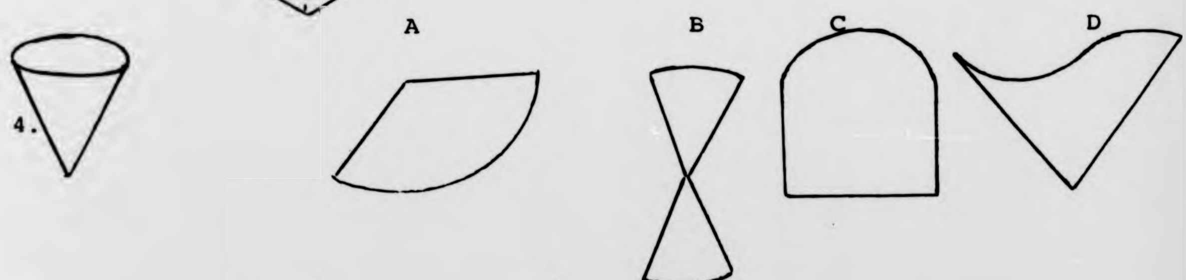
2.



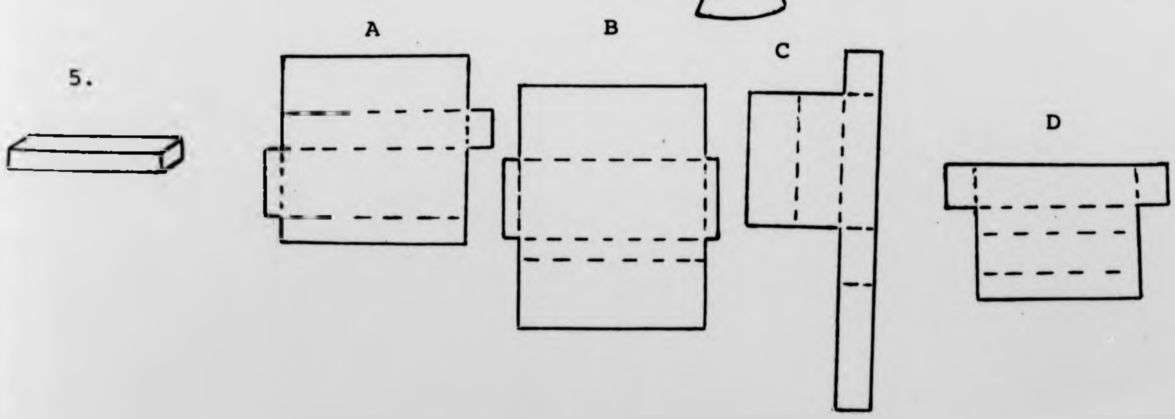
3.



4.



5.



BLANK PAGE IN ORIGINAL

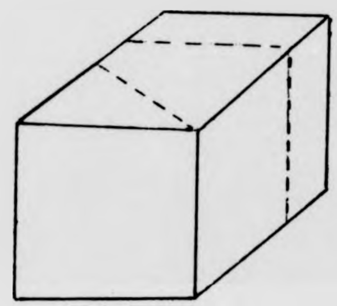
JIG CUBES

The drawings A and B show two cubes which can be built from the smaller pieces shown by the dotted lines. Below A and B are the pieces that can be used to build the cubes. Use all the pieces, some to make A and some to make B.

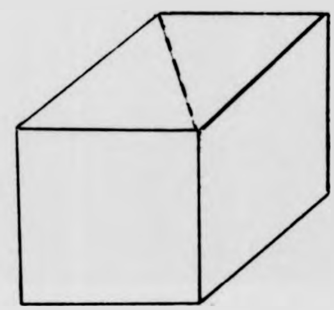
In the box on the answer sheet put the numbers of the pieces alongside the letter of the cube they will make.

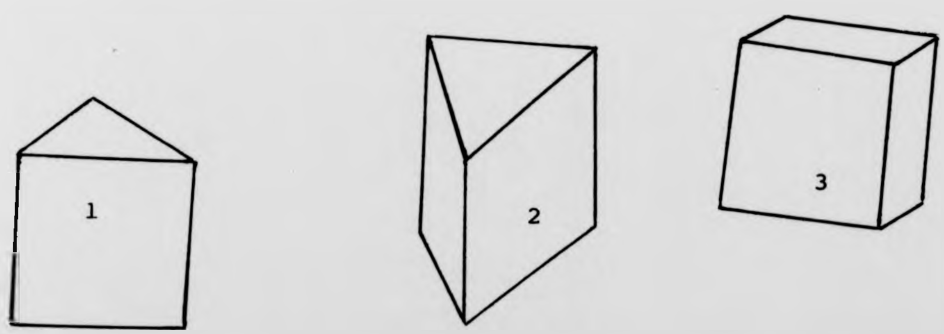
An example is given.

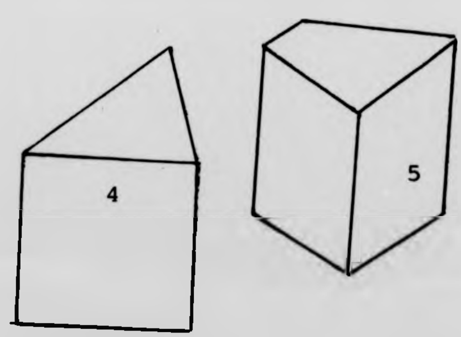
A



B



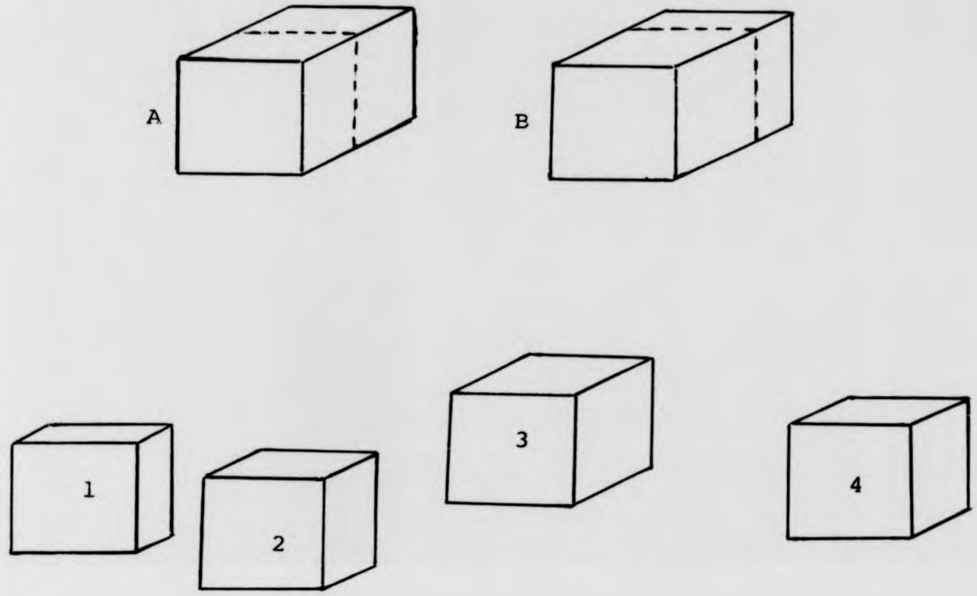




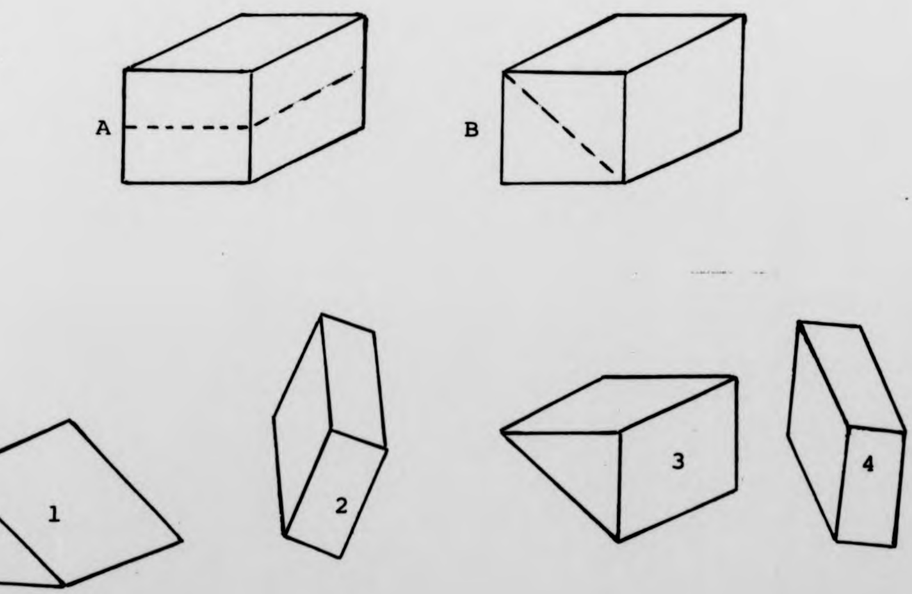
Answer

A	
B	

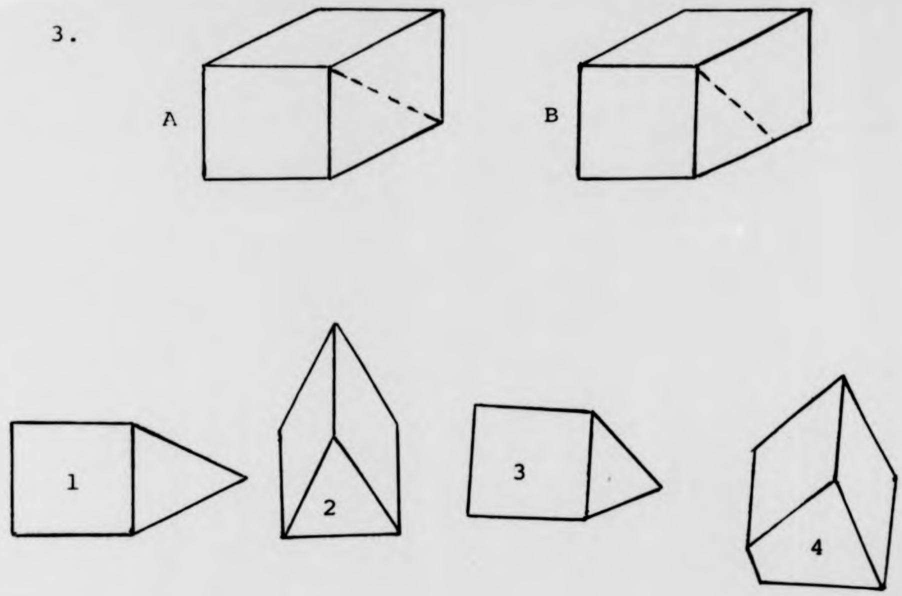
1.



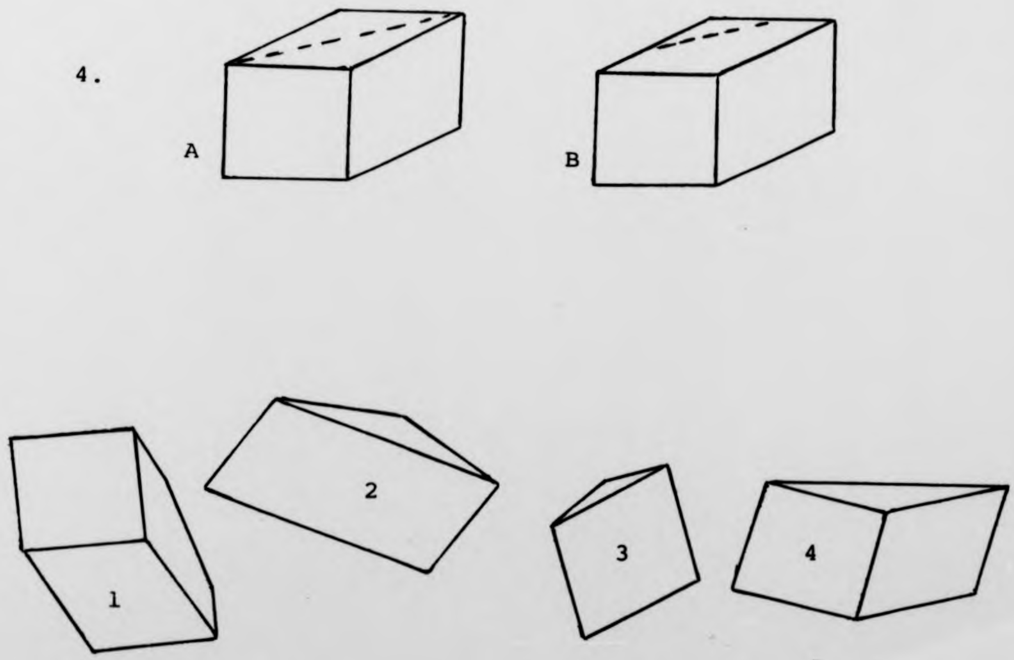
2.



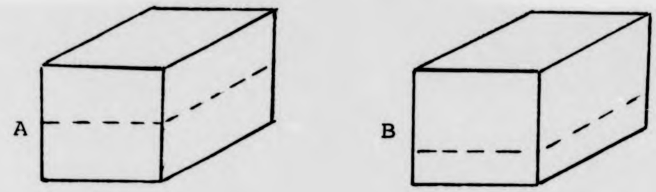
3.

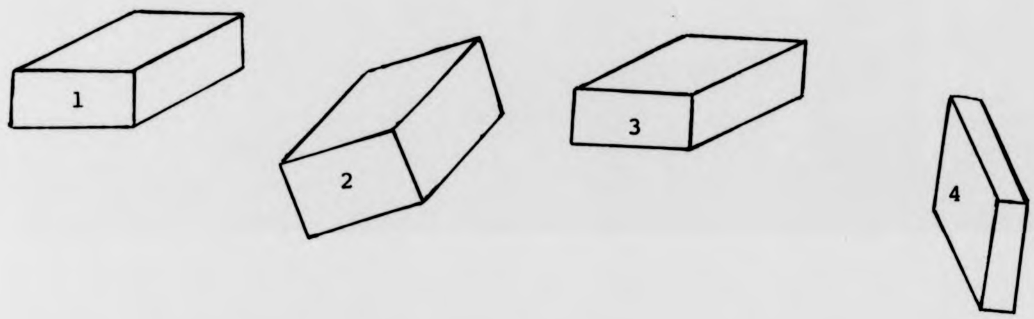


4.

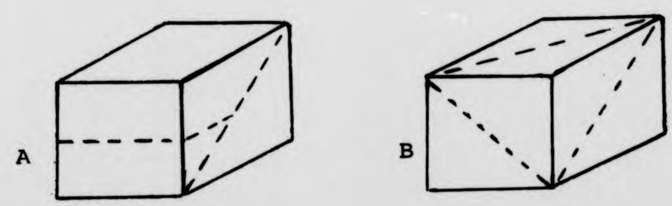


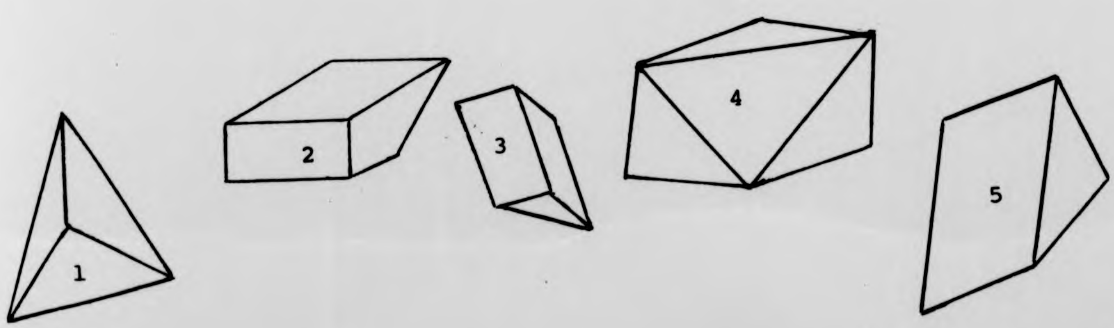
5.



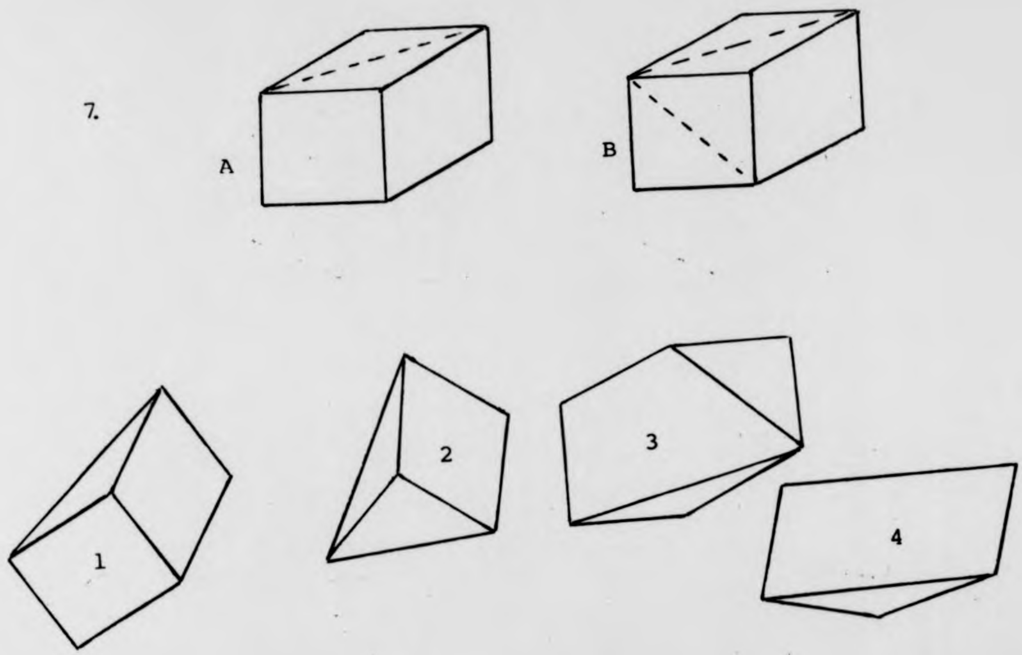


6.

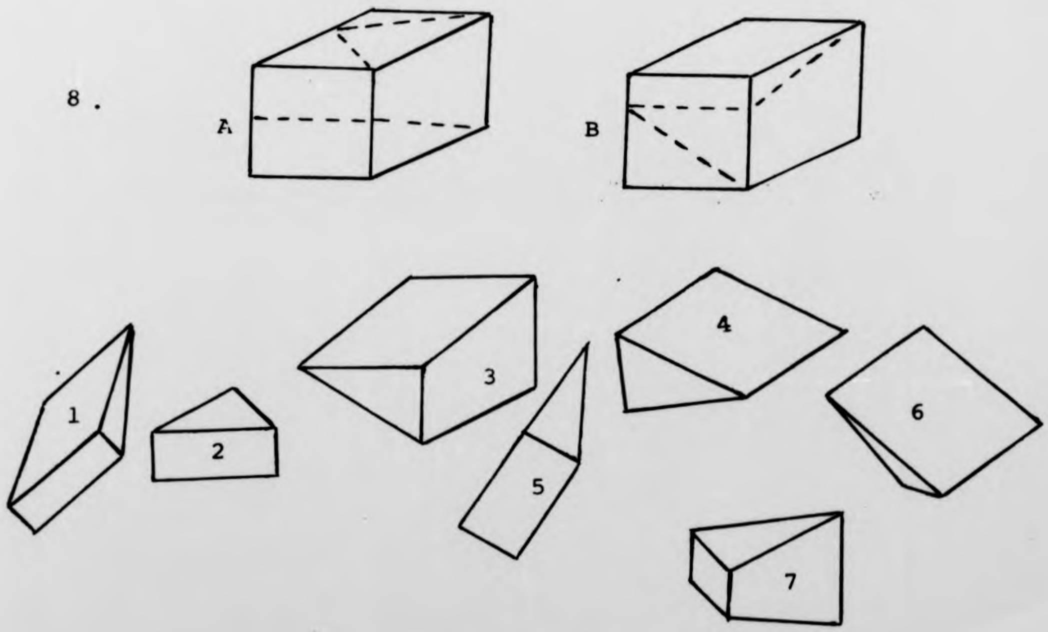




7.

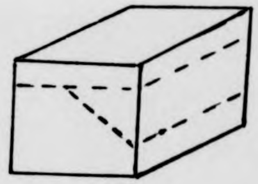


8 .

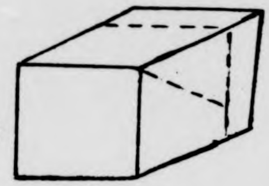


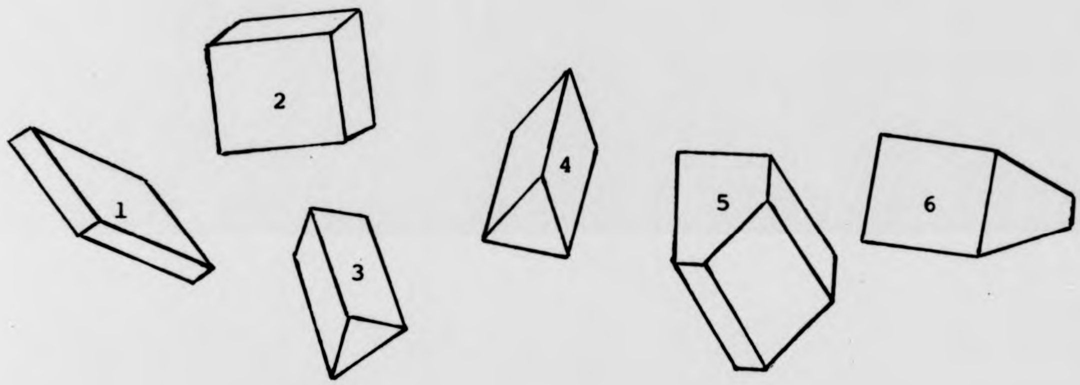
9.

A



B





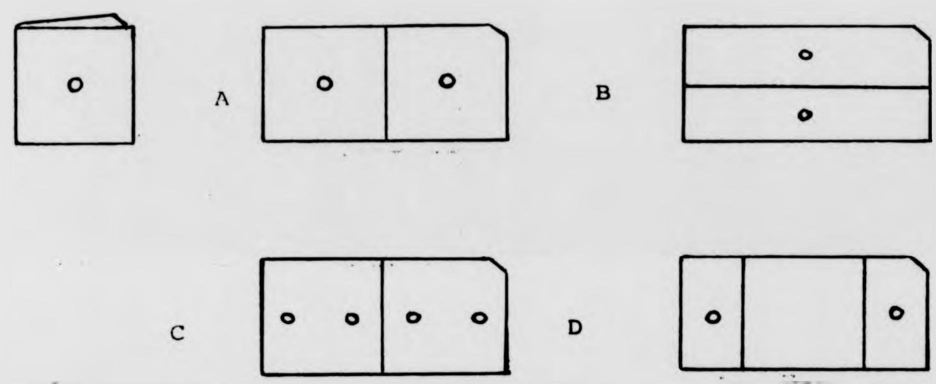
End of Jig cubes - <u>DO NOT TURN OVER</u>

FOLDS

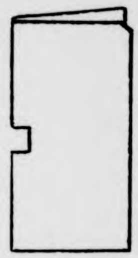
Here are some diagrams of folded paper and alongside each are four possible views of the inside of the paper when opened out.

Select the letter of the diagram you think is the correct view. The diagram of the folded sheet has a corner cut away so you can match it with the flat sheet.

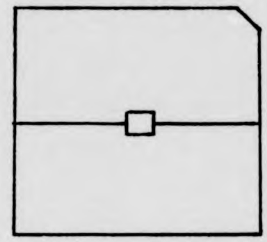
Example



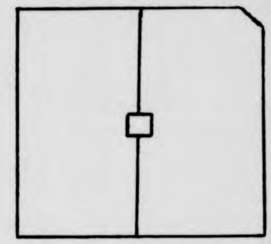
1



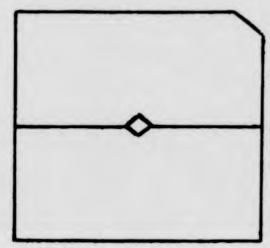
A



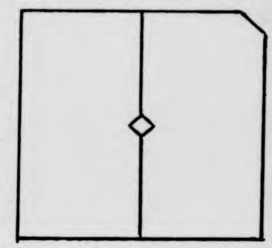
B



C

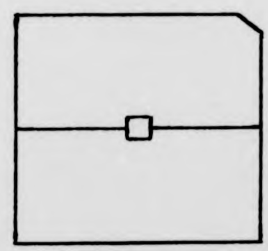


D

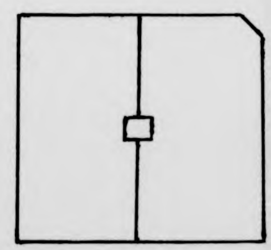


2

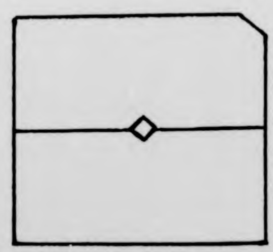
A



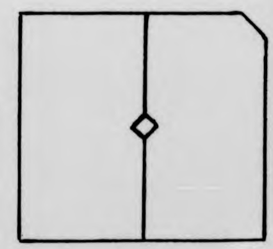
B



C

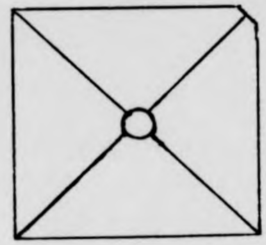


D

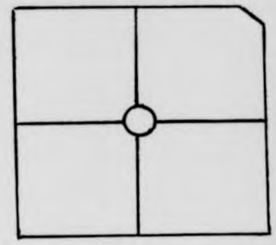


3

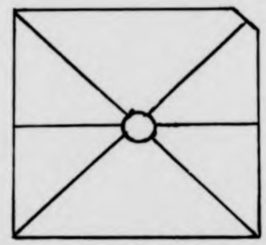
A



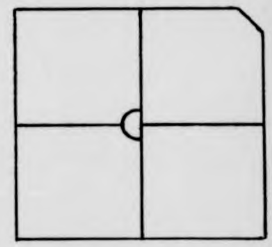
B



C

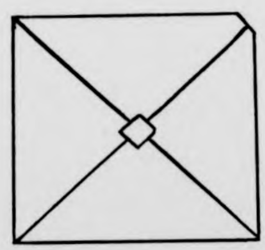


D

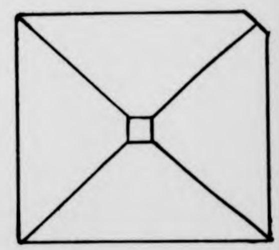


4

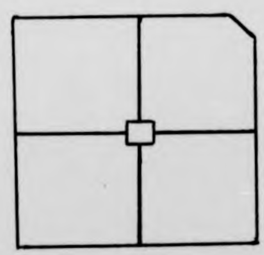
A



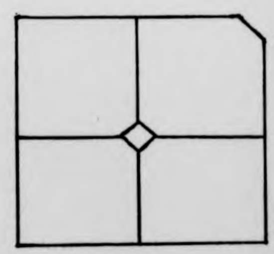
B



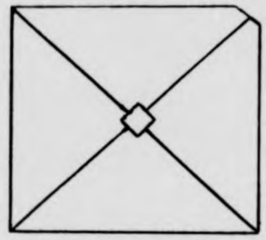
C

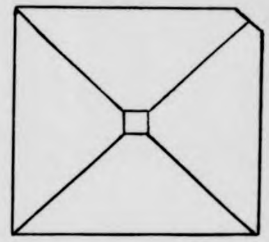


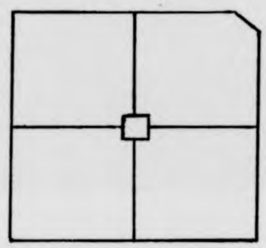
D

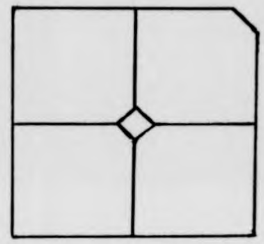


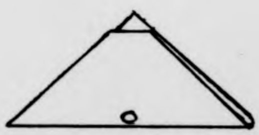
5. 

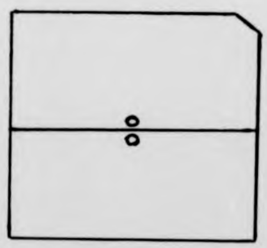
A 

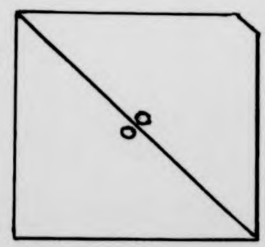
B 

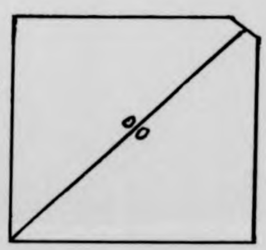
C 

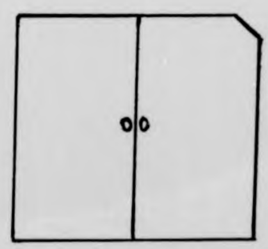
D 

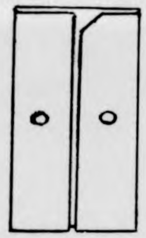
6. 

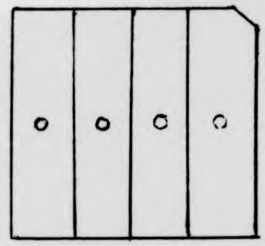
A 

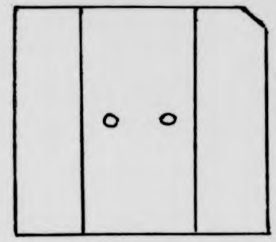
B 

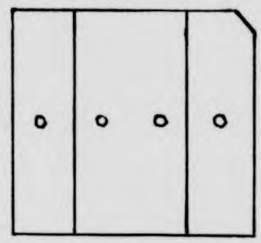
C 

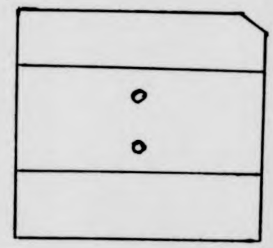
D 

7. 

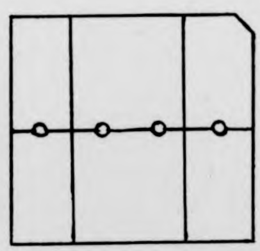
A 

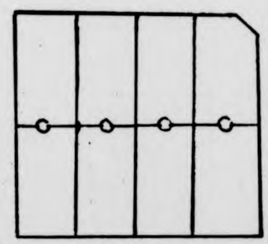
B 

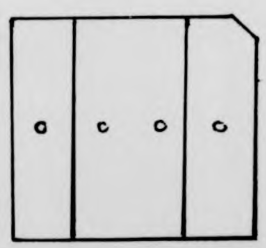
C 

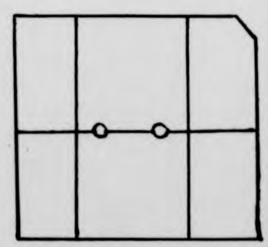
D 

8. 

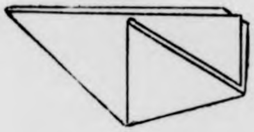
A 

B 

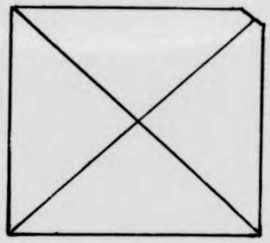
C 

D 

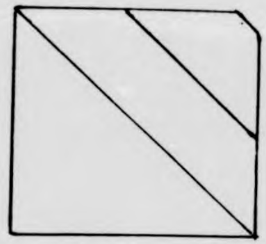
9.



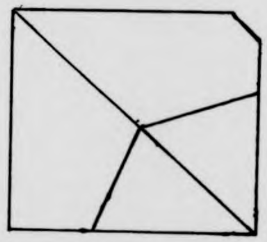
A



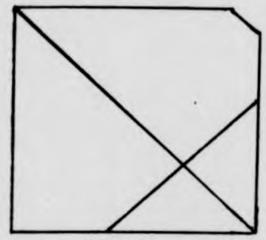
B



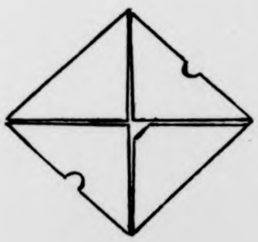
C



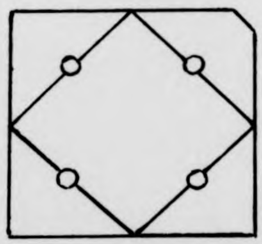
D



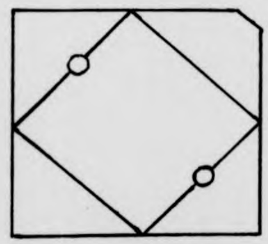
10.



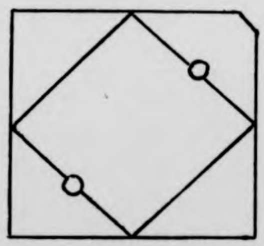
A



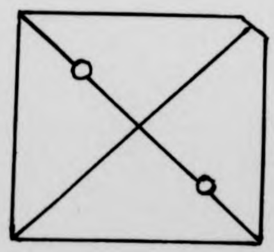
B

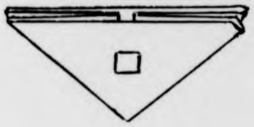


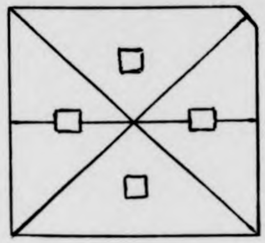
C

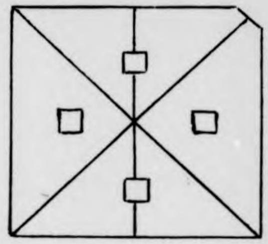


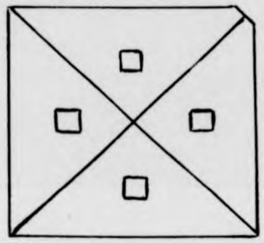
D

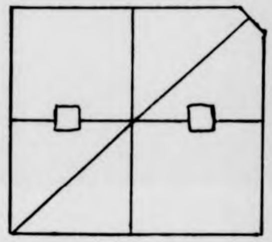


11. 

A 

B 

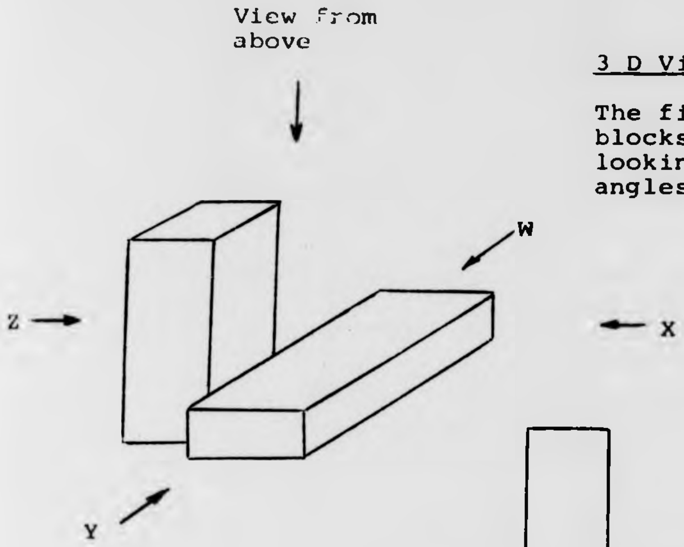
C 

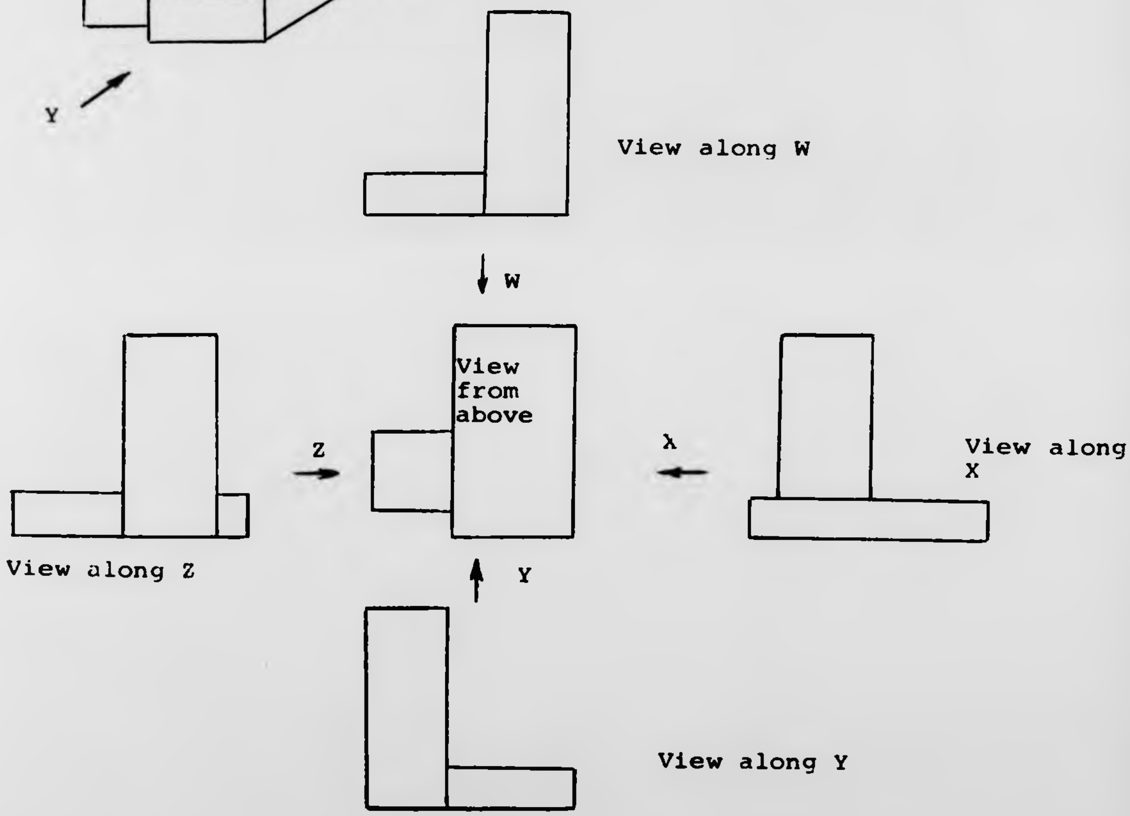
D 

END OF FOLDS

3 D Views

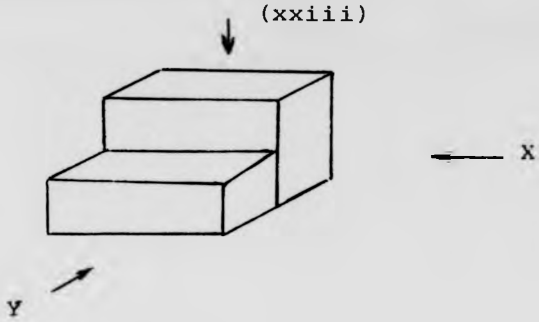
The first drawing is of some wooden blocks. The views you would get looking at the blocks from different angles is shown below.



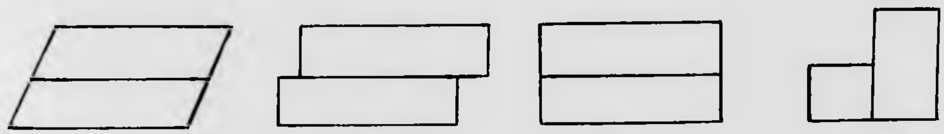


In the questions that follow you are given 3 dimensional drawings and four possible views you would get looking in the direction stated. Select the correct one and put its letter in the correct box on the answer sheet. An example is done for you.

(xxiii)

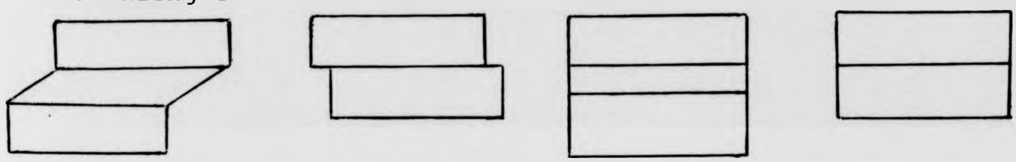


View from above



A B C D

View along Y



A B C D

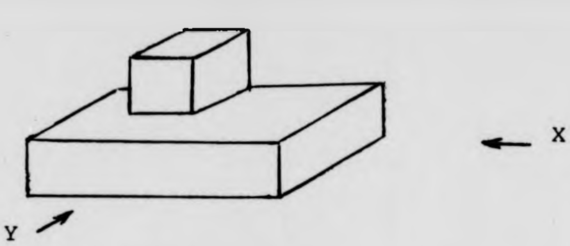
View along X

A B C D

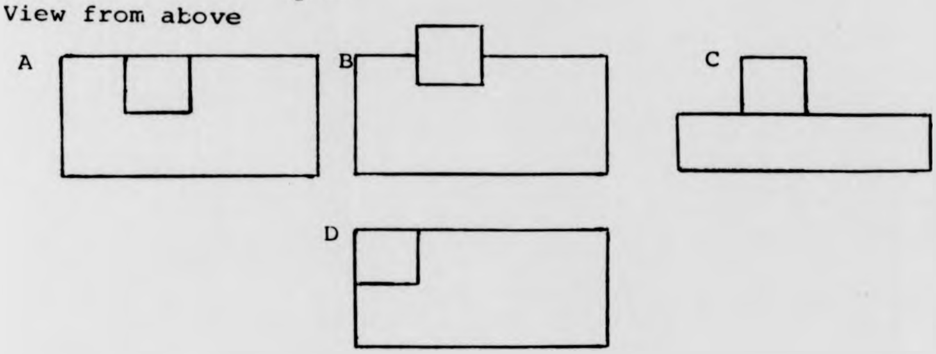
Answer

View from above	
View along Y	
View along X	

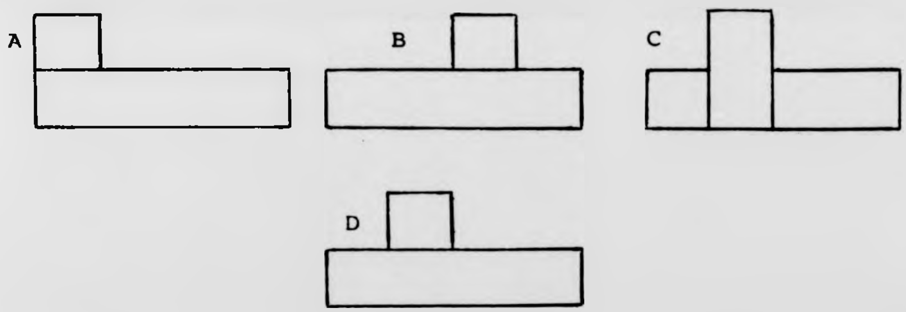
1.



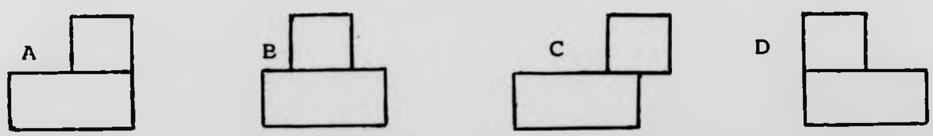
View from above



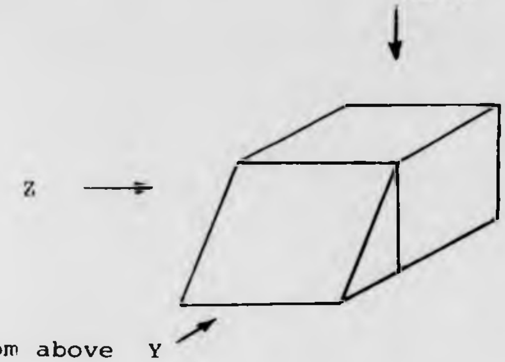
View along Y



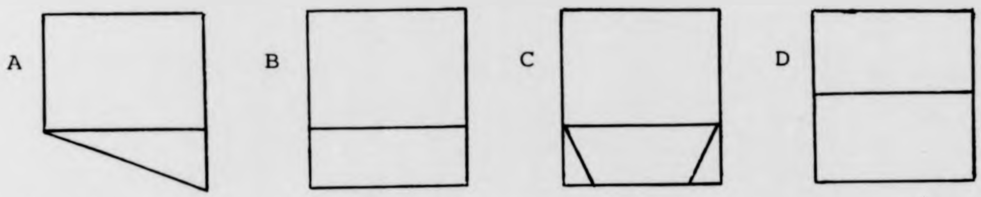
View along X



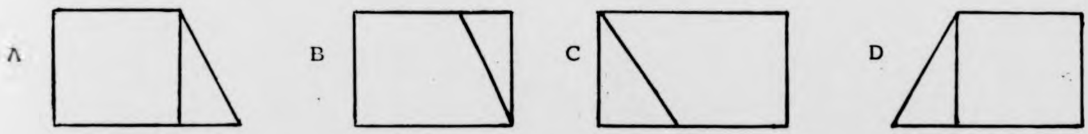
2.



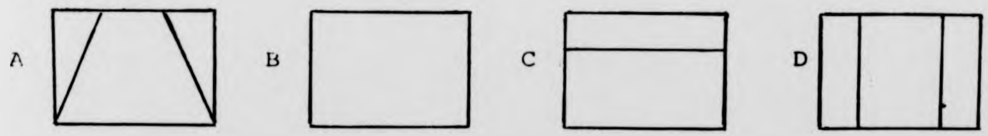
View from above Y



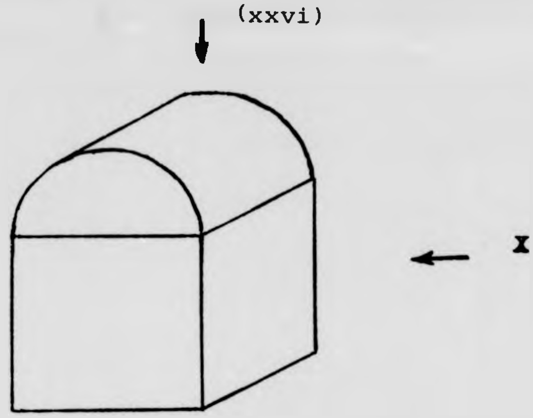
View along Z



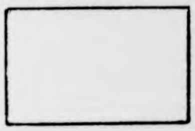
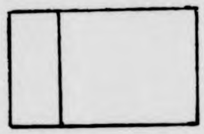
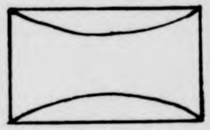
View along Y



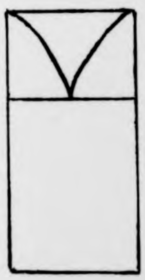
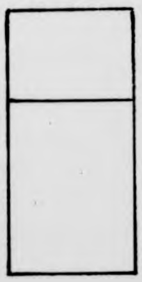
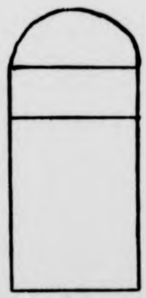
3.



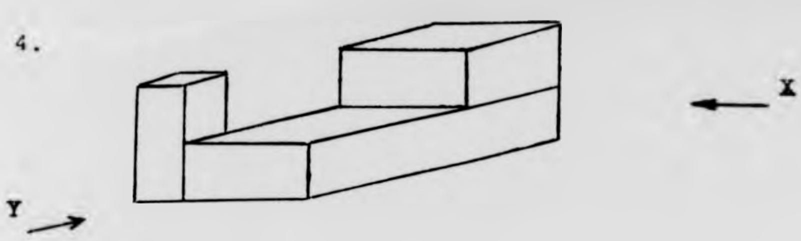
View from above

A  B  C  D 

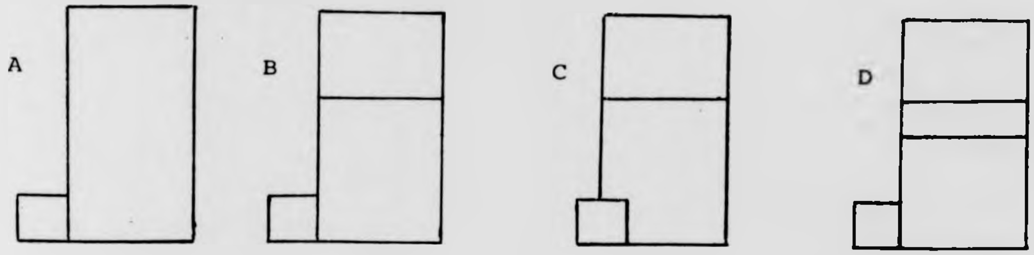
View along X

A  B  C  D 

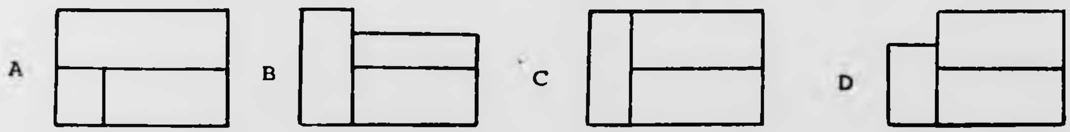
4.



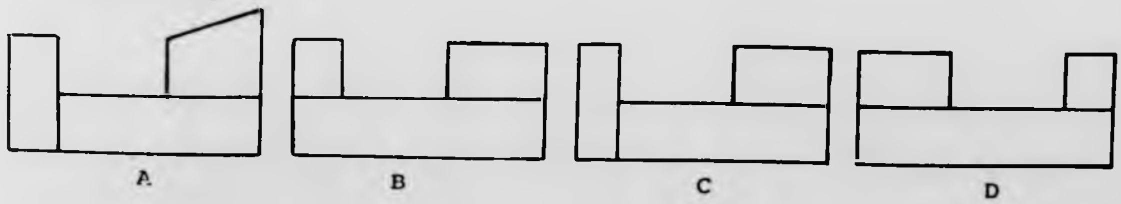
View from above



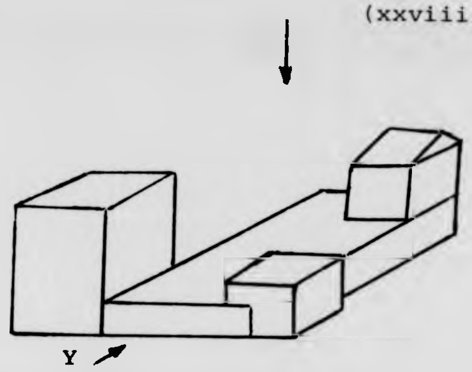
View along Y



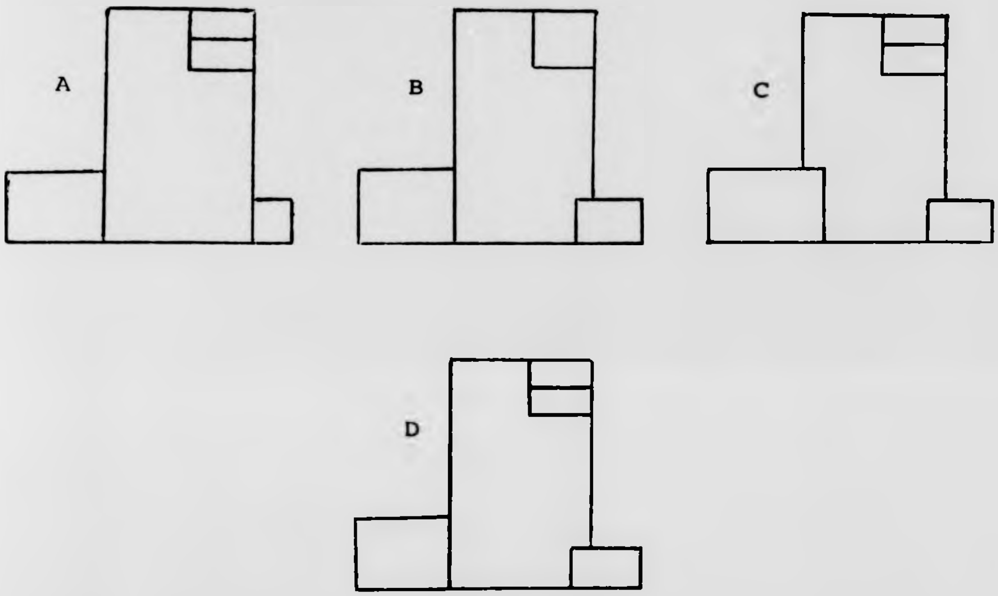
View along X



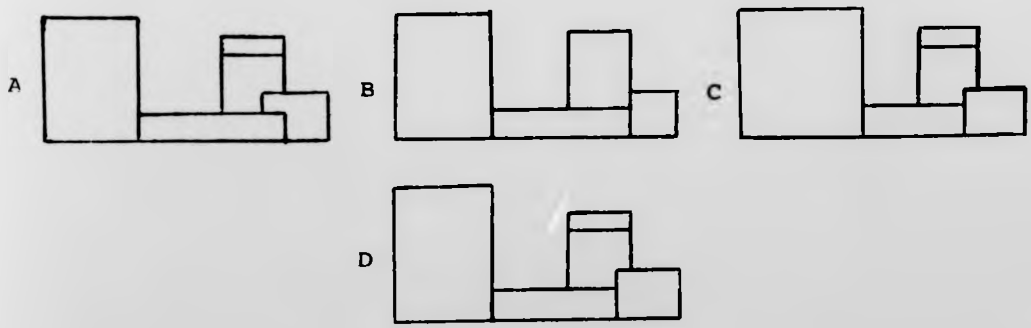
5.



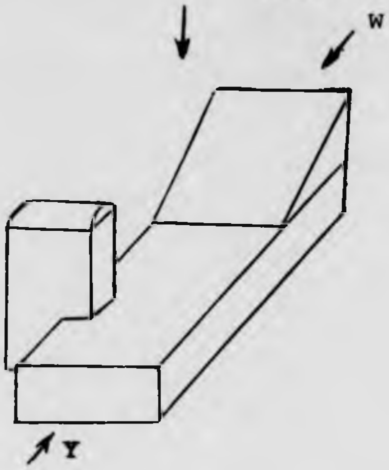
View from Above



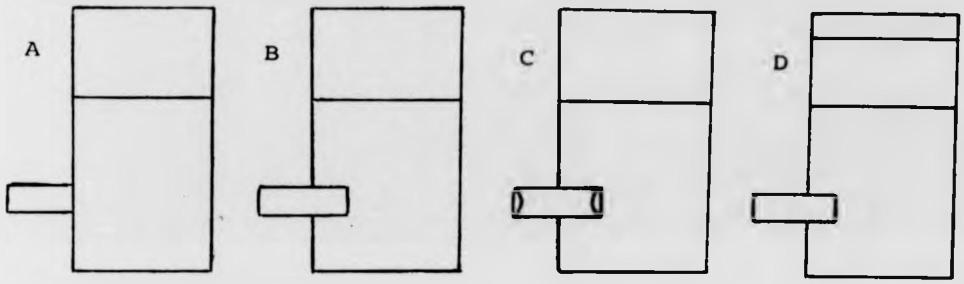
View along Y



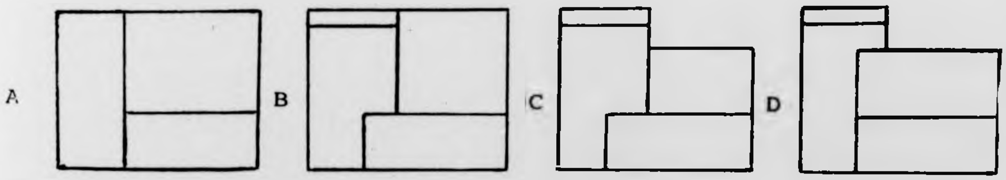
6.



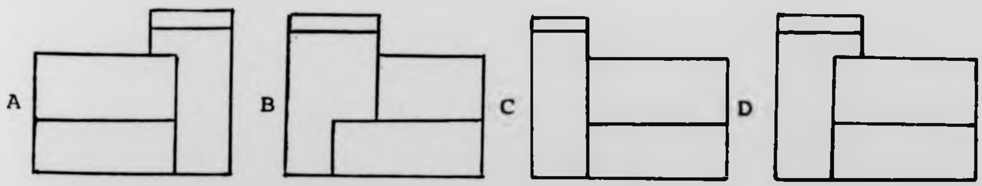
View from above



View along Y



View along W



END OF 3 D Views

ANSWER SHEET

N.B. These timings have been arrived at for the 11-13 year age group and give good distribution. Timings for other ages have not yet been determined.

TURN ABOUT TIME 4 Mins. EXAMPLE A

1	E	5	A	9	D	13	A
2	D	6	B	10	A	14	C
3	A	7	D	11	D	15	D
4	D	8	B	12	E	16	C
						17	A

CUTOUTS TIME 2¾ Mins. EXAMPLE D

1	B	4	A	7	B	10	D
2	A	5	D	8	A		
3	C	6	C	9	C		

JIGCUBES TIME 6½ Mins. EXAMPLE A

1)	A 2.4 / B 1.3	4)	A 2.4 / B 1.3	7)	A 1.4 / B 3.2		
2)	A 2.4 / B 1.3	5)	A 1.3 / B 2.4	8)	A 2.5.6.7 / B 1.3.4		
3)	A 1.2 / B 3.4	6)	A 2.3.5 / B 4.1	9)	A 1.3.5 / B 6.4.2		

FOLDS TIME 4 Mins. EXAMPLE A

1	A	5	A	9	C
2	C	6	B	10	B
3	B	7	C	11	B
4	D	8	A		

Turn over

3-D VIEWS TIME 3½ Mins. EXAMPLE: View from above C

View along Y D

View along X B

1)

VIEW FROM ABOVE	A
VIEW ALONG Y	D
VIEW ALONG X	A

2)

VIEW FROM ABOVE	B
VIEW ALONG Z	A
VIEW ALONG Y	B

3)

VIEW FROM ABOVE	A
VIEW ALONG X	B

4)

VIEW FROM ABOVE	B
VIEW ALONG Y	C
VIEW ALONG X	B

5)

VIEW FROM ABOVE	D
VIEW ALONG Y	A

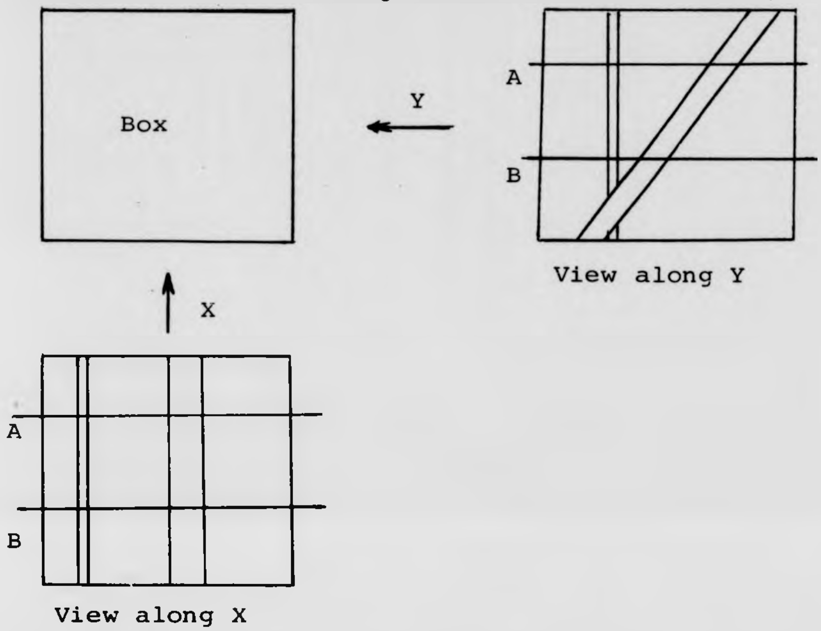
6)

VIEW FROM ABOVE	B
VIEW ALONG Y	C
VIEW ALONG W	A

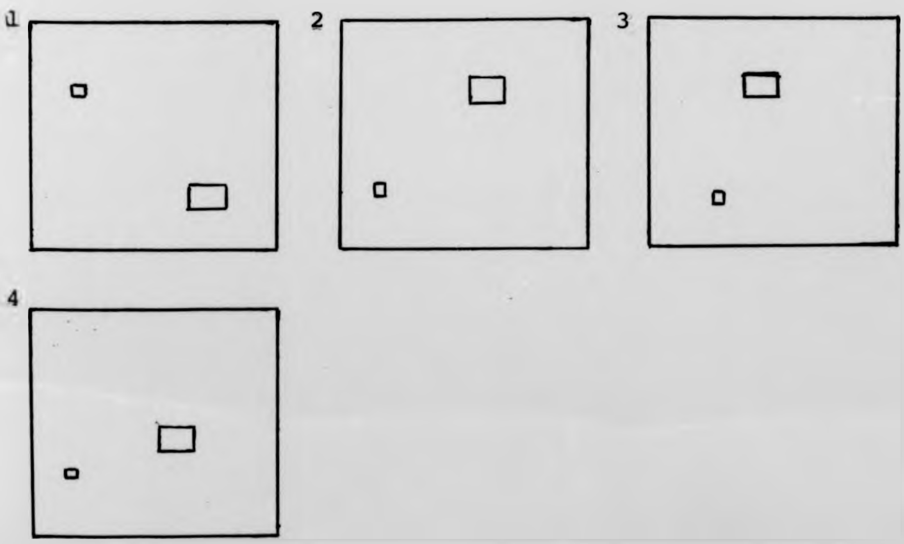
APPENDIX B

Cross sections I

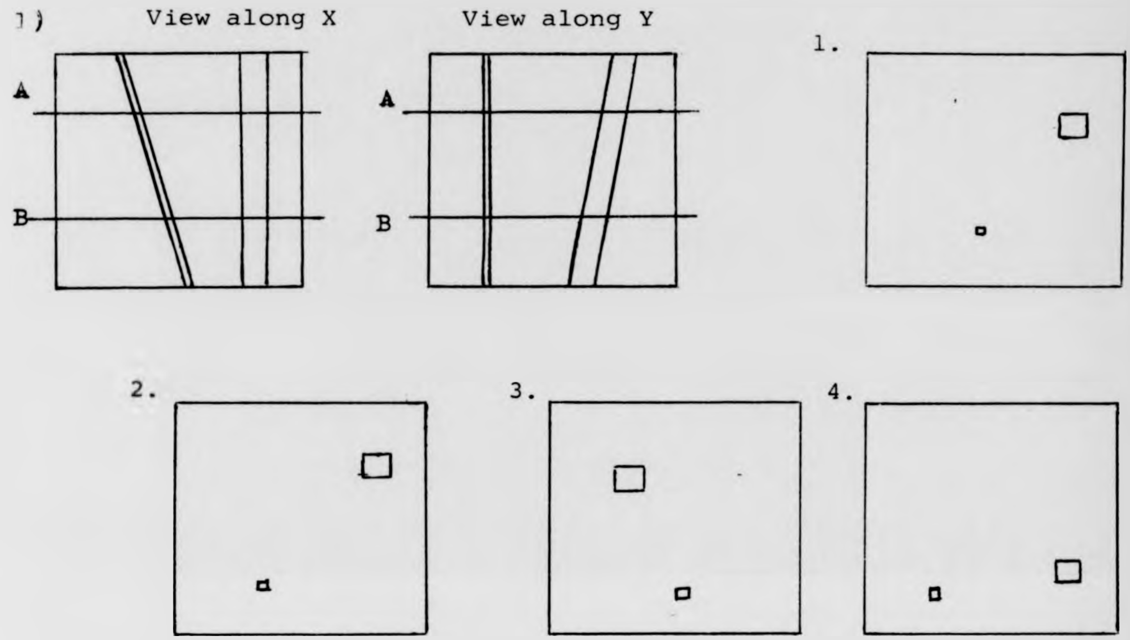
The views X and Y are of a glass sided box in which two bars have been placed. The drawing below shows the box and the two directions X and Y along which it can be viewed.



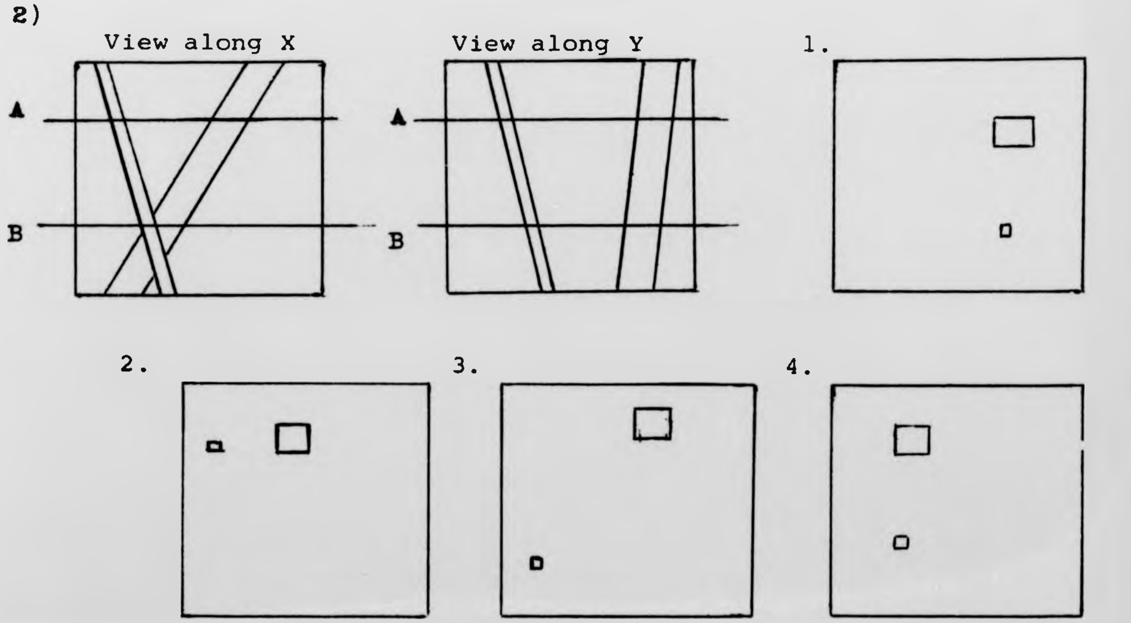
Of the four squares shown below two show the view you would get looking down on slices cut through A and B. (They are the cross sections through A and B). Select which box is the section through A and which through B and put the number in the box on the answer sheet.



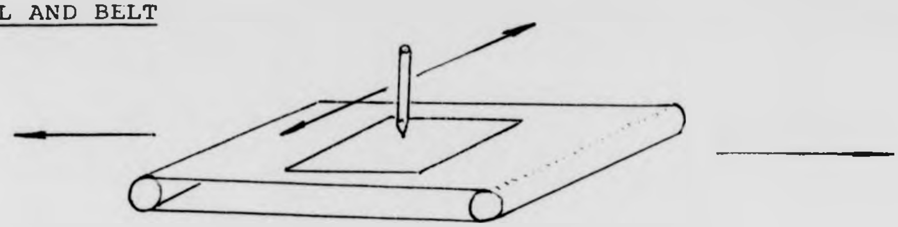
1)


View along X
View along Y
A
B
A
B
1.
2.
3.
4.

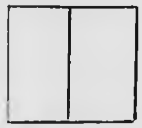
2)


View along X
View along Y
A
B
A
B
1.
2.
3.
4.

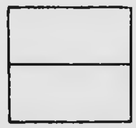
PENCIL AND BELT



The pencil shown above is free to move backwards and forwards across the belt but not along it. If a piece of paper is placed under the pencil and the pencil moves from centre backwards and forwards it will draw a line as shown.



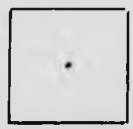
If the pencil is kept fixed in the centre of the paper and the belt moves right and left it can draw a line as shown.

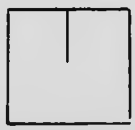


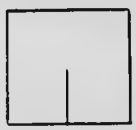
Below you are shown two drawings.

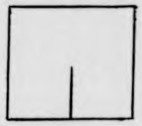
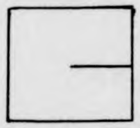
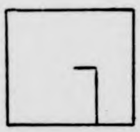
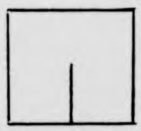
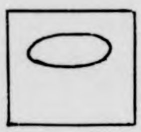
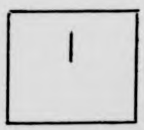
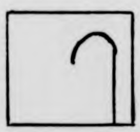
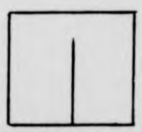
A) is made by the movement of the belt and pencil TOGETHER
B) is made by the movement of the pencil ALONE

By each two drawings state how the belt has moved.
Some examples are given. In some cases the two drawings cannot go together (impossible).

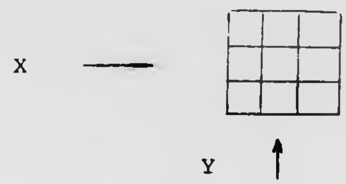
A  B  Belt moves to right

  Belt moved to the left

  Impossible

	Belt and Pencil	Pencil alone	Answer
1	A 	B 	
2			
3			
4			
5			
6			
7			
8			

You are asked to construct a pile of bricks using the following instructions. Here is a view from above of a base on which you are to build a pile of bricks.



A full base will contain 9 bricks in one layer

Construct the block as follows.

1) (i) On the bottom layer place blocks in all positions except the top right hand corner as seen from above.

(ii) On the next layer place bricks on the left hand and centre rows leaving the right hand row empty.

(iii) On the next layer put bricks in the bottom left hand and top left hand corner.

How many bricks have you used? ....

Was it possible to build the pile? ....

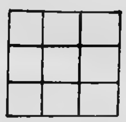
If not explain why ....

....

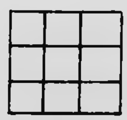
....

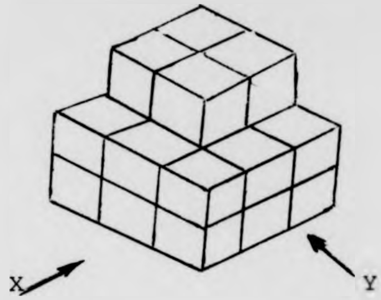
If you could build the pile what would you see looking along X and Y? Again cross out bricks you would not see.

View along X



View along Y





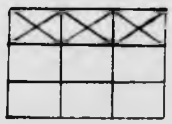
View along Y

View along X

In the views above put crosses in the squares that you would remove to show the view along X and Y.

e.g. if you think you would see

then you would put



How many bricks are in the pile shown above?

Turn the pile of bricks through 90° in a clockwise direction when viewed from above. What view would you now get along X and Y? Again cross out the bricks you would not see.

View along X

View along Y

2) (i) fill the bottom layer completely
(ii) fill the second layer except for the centre brick
(iii) fill the top layer completely.

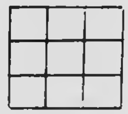
How many bricks have you used? ....

Was it possible to build the pile? ....

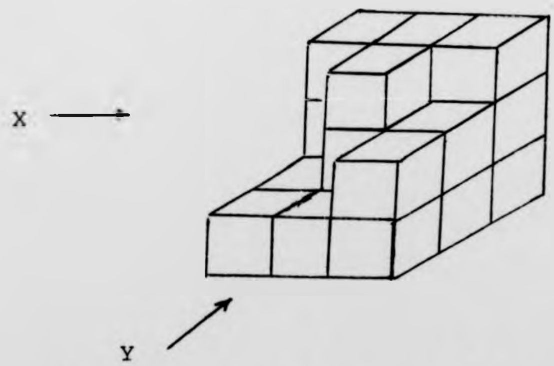
If not explain why ....

....

If you could build the pile what would you see looking along X and Y?

View along X  View along Y 

Using the same method give instructions for building the pile of bricks shown below



(i) ....

....

(ii) ....

....

(iii) ....

....