Spatial Ability of Students in Selected Addis Ababa Schools

Temechegn Engida*

The Nature and Measurement of Spatial Ability

Spatial Ability as a Separate Factor of Intelligence

Two theories have dominated the psychometric conception of intelligence. The first view was advanced by Spearman. Psychometrics was essentially the invention of Spearman, who, in 1904, published a paper that contained all the basic psychometric reasoning about the nature of intelligence. As Kline (1991) put the point:

> ... the question Spearman asked was this: why is it that human abilities are positively correlated, that is, why is there a general tendency for those who are good at one thing to be good at others? ... The simple answer ... was that this correlation could be accounted for by general intelligence or g. This g was common to all tasks requiring ability. ... For example, ability at ... physics depended upon g plus a physics component and so on (p.3).

Thus, Spearman argued that each test of intelligence could be analyzed into a general ability factor, g, that accounts for most of the variance and into several specific or group factors.

Other British psychometricians such as Vernon held the view that intelligence consists of a hierarchy of abilities with g at the base. Vernon's analysis of intelligence tests in 1950 revealed two main groups of tests. *The major group factors described by Vernon are v:ed, or verbal educational abilities, and k:m, spatial mechanical*

* Assistant Professor, Department of Curriculum and Instruction, Faculty of Education, Addis Ababa University.

Spatial Ability of Students in Selected Addis Ababa Schools

Temechegn Engida*

The Nature and Measurement of Spatial Ability

Spatial Ability as a Separate Factor of Intelligence

Two theories have dominated the psychometric conception of intelligence. The first view was advanced by Spearman. Psychometrics was essentially the invention of Spearman, who, in 1904, published a paper that contained all the basic psychometric reasoning about the nature of intelligence. As Kline (1991) put the point:

> ... the question Spearman asked was this: why is it that human abilities are positively correlated, that is, why is there a general tendency for those who are good at one thing to be good at others? ... The simple answer ... was that this correlation could be accounted for by general intelligence or g. This g was common to all tasks requiring ability. ... For example, ability at ... physics depended upon g plus a physics component and so on (p.3).

Thus, Spearman argued that each test of intelligence could be analyzed into a general ability factor, g, that accounts for most of the variance and into several specific or group factors.

Other British psychometricians such as Vernon held the view that intelligence consists of a hierarchy of abilities with g at the base. Vernon's analysis of intelligence tests in 1950 revealed two main groups of tests. *The major group factors described by Vernon are v:ed, or verbal educational abilities, and k:m, spatial mechanical*

* Assistant Professor, Department of Curriculum and Instruction, Faculty of Education, Addis Ababa University.

abilities. These groupings can be thought of as indicative of more specific abilities, one level up from g in the psychometric hierarchy of abilities (Anderson, 1992: 23).

Thurstone (in Anderson, 1992) held the other view on the other side of the Atlantic. To him performance on a test was not a reflection of general ability *per se*, but was dependent upon a limited number of independent abilities. He identified eight factors in all, giving them the name 'Primary Mental Abilities': verbal, word fluency, number, spatial visualization, rote memory, inductive reasoning, deductive reasoning, and perceptual speed.

According to Thurstone; (in Anderson, 1992) human intelligence was composed of more primary factors than a general ability factor. Later Guilford claimed 120 ability factors, while Catell dealt with five large factors of which the spatial factor is one of the primary factors (Kline, 1991). Analysis of the earlier research works in this area also revealed that studies of mechanical aptitude and practical ability were, historically, the roots of the identification of the spatial factor. McFarlane, for instance, found in her study of *practical ability* evidence for a group factor over and above general intelligence -- g (McGee, 1979).

Since 1925, numerous factor analytic studies have been conducted to further differentiate spatial ability into two or more sub-abilities. Extensive reviews and analysis of such works have also been conducted in the last two decades (Lohmann, 1979; McGee, 1979; Pellegrino and Kail, 1982; Guttman, et al., 1990). However, there is little agreement among these studies as to the number of distinct spatial abilities that may exist and how best to characterize each one. Before 1930, the debate over the existence versus non-existence of a spatial factor was dominant. However, *a plethora of factor studies since that date have provided strong and consistent support for the existence of two distinct spatial abilities: visualization and orientation* (McGee, 1979: 892). In the same year Lohman's (1979) analysis revealed three major spatial factors (and some minor ones) all of which require mental transformation. They are Spatial Relation (SR), Spatial Orientation (SO), and Visualization (Vz).

Linn and Petersen (1985) employed a meta-analysis technique that focused on similarities in the processes that respondents use for individual items. Accordingly, they identified three factors or categories of spatial ability: Spatial Perception, Mental Rotation, and Spatial Visualization.

Generally, many factor analytic studies on the definition and structure of spatial ability revealed the complex nature of spatial ability. The above breakdown of spatial ability into various categories is not exhaustive. Although such factors as visualization and orientation reappear in several different studies, there is no consistency among the various authors in assigning definitions to the same names. These studies have also *not resulted in a theoretical framework for the structure of spatial abilities because the factor names are often for vague and overlapping concepts* (Guttman et al., 1990: p. 220).

Developmental Level and Spatial Ability

Developmental level is usually taken to be synonymous with mental age. Many perspectives of intelligence recognize that cognitive abilities increase with development. The general observation is that, when children are measured for their developmental level using cognitive tests, they become more proficient at these tests as they become older. They do that in a similar way for all cognitive tests. Mental age is thus regarded as the appropriate measure of intelligence during development (Anderson, 1992).

Spatial ability as one of the specific factors of intelligence has been investigated using various tests to identify, directly or indirectly, such a developmental trend. Of course, the study of age-related differences in spatial ability has its roots in Piaget's recognition of the developmental nature of spatial representation. The ability of children between 9 and 11 years of age to maintain the relative positions of parts of a figure, of figures relative to one another, and of whole configurations relative to different points of view makes them different from younger children (Piaget and Inhelder, 1967).

In Piagetian theory, the critical developmental changes in spatial processing of children occur during middle childhood. Pellegrino and Kail (1982) argue, recognizing that children may acquire the rudiments of spatial processing by middle childhood, that development in this regard continues with individuals ultimately attaining different levels of skill. In a research designed to investigate the development of spatial processing in late childhood and adolescence, samples of 8 and 19-year-olds children were studied (Pellegrino and Kail, 1982). The research showed that there was a regular developmental change in rate of mental rotation. It is then argued that age differences may reflect the greater speed with which older individuals construct internal representations of unfamiliar stimuli.

Barke (1978), using a paper-and-pencil spatial ability test, observed such a developmental trend for students of grade 7 through 9 in German gymnasiums. He then argued that grade 8 would be the minimum stage at which students begin correctly visualizing spatial objects presented as two-dimensional diagrams. Ben-Chaim et al. (1986) also observed similar developmental trend using a paper-and-pencil spatial visualization test for middle school students (grade 5 through 8) with fifth graders performing at the bottom. Such a trend in spatial ability is thus consistent with common findings of grade effects for cognitive variables. Munroe et al. (1985) concluded from their research on children that there is continuity in the development of spatial skills from early to middle childhood.

Gender and Spatial Ability

The study of sex-related differences in spatial ability has been the subject of many researchers for the last decades of the last century. All attempted, directly or indirectly, to answer the question: Do sex-related differences in spatial abilities exist? For some researchers and reviewers the answer seems 'yes, of course'; for others 'no'; still some others find themselves between the 'yes-no' boundary. This part of the study discusses the views of all these groups. For convenience, the review is presented from the earliest to the latest (i.e. arranged chronologically).

Early in the 1970s Macoby and Jacklin (1974:351) concluded from their analysis of research works in the field that *boys excel in visual-spatial ability*. This male-superiority was found in adolescence and adulthood and the advantage increases through the high school years. This difference, however, was not detected in childhood.

Late in the 1970s two researchers tackled the question of sex difference in spatial ability. Barke's (1978) empirical research in a German gymnasium has shown that there is a significant difference ($p<0.05$) in spatial ability of grades 7 through 9, in each case favoring the boys. McGee's (1979) review, on the other hand, showed that the sex difference on spatial tasks does not reliably appear until puberty.

Newcombe (1982) was of the opinion that sex differences in spatial ability could have been detected even during childhood had it not been for the use of less sensitive methodologies and small samples. According to Newcombe, the small sex-related differences present in childhood increase somewhat in size with age. In contrast, Munroe's et al. (1985) follow-up study of 6 to 12-year-old Kenyan children indicated that a sex difference in spatial ability was not displayed in later childhood although a difference was found earlier for the same sample.

Linn and Peterson (1985) employed meta-analysis using studies on spatial ability published between 1974 (Macoby and Jacklin' review) and 1982. Their analysis has shown that sex difference on spatial ability is not uniform across all the factors of spatial ability. They contradicted the conclusion that the difference is first detected in adolescence.

The year 1985 seems to be a time when a great challenge to the conventionally accepted male superiority in spatial ability has been greatly questioned. Caplan et al. (1985) strongly criticized the research reports on male superiority on spatial ability. Considering the fact that sex differences on spatial tasks are modest and inconsistent across different tasks, they concluded that *sex differences in spatial ability do not exist or at least it is by no means clear as yet* (p. 797). Since there is a great deal of definitional dilemma on spatial ability, *it is at least premature and indeed inappropriate to ask about sex differences in spatial ability* (Ibid.). These conclusions of Caplan and others have somewhat upset some of the other authorities in the field and initiated a further debate on the issue. Many of the researchers (Burnett, 1986; Eliot, 1986; Hiscock, 1986; Halpern, 1986; Sanders et al., 1986) forwarded their reactions to the Caplan et al. article. Although many of them acknowledged the definitional dilemma reported by Caplan and others, they questioned the validity of the conclusions on sex differences in spatial ability. Failure to consider age, to account for the various factors of spatial ability, to sample the appropriate empirical studies for the analysis were mentioned as shortcomings of the study by Caplan *et al.*

Johnson and Meade (1987) argued, using public school students of grades K--12 or ages 6--18, that a male advantage in spatial ability appears reliably by age 10 (4th grade) and that the magnitude of the advantage remains constant through age 18. They also noted the largest sex difference occurring at grades 10-12. On the contrary, towards the end of the 80s, two studies (Methen, 1988 and Klainin, 1989 quoted in Colemann and Gotch, 1998) showed using non-Western populations that females performed better than males on tasks requiring spatial ability.

At the beginning of the 1990s Barke and Kuhrke (1992) conducted a research entitled *Girls in Science and Technology* and reported sex differences in spatial ability of pupils. Voyer et al. (1995), paying due attention to the Caplan et al. (1985) conclusion, concluded that *the present meta-analysis reports further evidence for the existence of sex* *differences in favor of males in tests that assess mental rotation and spatial perception skills* (p. 262). Their analysis further showed that there is little or no sex difference on the spatial visualization category. McArthur and Wellner (1996) pointed out, using a Piagetian framework, that *gender differences, when present, are found to be small, and [that] similarities between performance are greater than differences between performance* (p. 1076).

The above review of research on spatial ability shows us that in each decade there was no clear consensus among researchers on the issue of sex-related differences in spatial ability.

Research Design and Methodology

The Research Method

The research was conducted by administering a translated version of an already developed spatial ability paper-and-pencil test to students. Although highly valid and reliable instruments can be available in the original language, the target language version of the test might suffer from problems arising during the translation process. These issues were considered in this research and the following paragraphs discuss how such issues are controlled to a reasonable extent.

Barke (1978) developed the original version of the Spatial Ability Test (SAT) for use in German schools. The writer of this paper then translated the present version of the test into Amharic in order to conduct the research work using this test in Ethiopian schools. This was a first challenging task since translation of research instruments for use in another language is, by its very nature, a complex and difficult process. The test consists of 40 items classified under five categories.

Research work on translation of research instruments for cross-cultural studies suggest four different types of translation methods for ensuring quality in the translated version: back translation, bilingual technique, committee approach, and pretest procedures (Brislin, 1980). They also recognize that the quality of translation depends on which of the following purposes the translator (researcher) wants to emphasize: accuracy of information, the affect and feelings of the original version, cultural context of the source and the second language versions, and finally the linguistic (grammatical) form of the second language.

Most of the four translation methods mentioned above require the use of bilingual experts, in this case German and Amharic speakers. Such experts are rare, or at least they were not accessible to the researcher. Moreover, the back translation technique would be important particularly when the researcher does not know the target language (Amharic). In this case, the researcher (and translator) is a native speaker of the target language. Thus, as a means of quality assurance in the test, the pretest procedures were employed. Brislin (1980) argues that this method is as good as the interviewers (examiners) doing the pre-testing. Considering this, the researcher (translator) himself administered both the pre-testing and the actual testing in Ethiopia.

In most cases, the test has many diagrams. These diagrams are common to everybody. Nevertheless, since it has written questions corresponding to each diagram, the cultural context of the words used in the questions was considered. This was found important because it was difficult to find single-word Amharic equivalents for some words in the German version of the test. A case in point is the following.

The German word *Kugel* (sphere) might be translated into the Amharic words *kib*, or *dibulbul*, or *kuas*. However, the researcher believes that the first term could be confused with something *circular* and the second might imply anything that is *rounded*. The third word is equivalent to the German word *Ball* (ball). This term is used in the Amharic version as equivalent to the German *Kugel* (sphere) because, for the purpose of this research, it is the least ambiguous of the three words and because it is the most familiar to every test taker.

Before the test was administered to the students, the researcher translated the German version into English for possible comparison purposes. Both the English and Amharic versions were then given to four Ethiopian teachers in Addis Ababa so that they could forward some suggestions. None of them disagreed with the equivalence of the two versions. Another hint to 'unfamiliarity' of the words in a research instrument can be obtained directly from the test takers. No complaint or question in this regard was detected during both the pre-testing phase and the actual testing phase in grades 7 – 12 (about ages 12 – 17).

Purposes of the Research

The first focus of this research is to identify the major factors in explaining difference and/ or similarities in students' ability to imagine three-dimensional structures represented as two-dimensional diagrams. In other words, is there a clear trend in the spatial ability of students indicating that schools encourage the ability of their children to a different extent? What are the likely factors for the difference and/ or similarities?

Developmental trend in spatial ability, the second focus of this study, has been an issue for some researchers. However, there is little or no evidence as to whether pupils of the same age, enrolled in different types of schools, show significant increase in spatial ability at the same grade level. Although a sex difference in spatial ability has been the focus of many researchers, there is little consensus on the commonly accepted male superiority in spatial tasks. The third focus of this research is thus to investigate this issue further with the intention of identifying the trend of sex differences in spatial ability of boys and girls in schools.

Subjects of the Study

The first stage in the sampling process was to identify two different types of schools in Addis Ababa: Government (G) and Non-Government (NG) schools. The majority of the children join the G schools because attendance in these schools is, up to now, free of charge. The NGs are financed by some sort of private or non-governmental organizations and hence are not free of charge. The only children who can join these schools are those whose parents can afford the relatively higher school fees. It is thus logical to expect the NGs to be relatively more equipped than the Gs. However, both types of schools should use the same type of school curricula developed centrally. The researcher used these two school types to obtain data directly from 762 students of grades 7 – 12, as shown in Table 1.

The Research Hypotheses

Three major null hypotheses are formulated to make the comparison between the spatial ability of students of the two school types. The three major null hypotheses are the following:

H_o1. There are no significant differences in spatial ability between the government and non-government school students.

H_o2. There are no significant differences in spatial ability of students of the various grade levels in each of the school types.

H_o3. There are no significant differences in spatial ability between boys and girls of the various grade levels in each of the school types.

The hypotheses are formulated in their null form primarily because there is no previous research of such a kind on Ethiopian students. The limited cross-cultural research work are also not sufficient to invite the formulation of any directional (or alternative) hypothesis.

Table 1. The Sample by School Type, Grade and Sex

Grade	Category	Ethiopia		Total
7	**Total**	99	137	236
	Boys	48	83	131
	Girls	51	54	105
8	**Total**	89	94	183
	Boys	44	52	96
	Girls	45	42	87
9	**Total**	60	47	107
	Boys	32	24	56
	Girls	28	23	51
10	**Total**	73	60	133
	Boys	45	39	84
	Girls	28	21	49
11	**Total**	41	38	79
	Boys	22	24	46
	Girls	19	14	33
12	**Total**	24	-	24
	Boys	11	-	11
	Girls	13	-	13
Grand	**Total**	386	376	762
	Boys	202	222	424
	Girls	184	154	338

G = Government Schools
NG = Non-Government Schools
-- = .No data collected

Evaluation of the Spatial Ability Test (SAT)

Evaluation here refers to item analysis. In this research, 42 grade 10 students were used for the item analysis purpose. Of these, 26 were boys and 16 girls. Descriptive statistics about these samples were calculated. The statistics reveal that there is a range of 20 scores and a mode of 10. The distribution curve, plotted as histograms and the corresponding normal distribution curves, has shown a moderately normal distribution. Given the smallness of the sample size, the distribution is encouraging.

Quantitative analysis of the distribution of scores also revealed that the curve does not deviate significantly from normal distributions. This is confirmed by carrying out the Kolmogorov-Smirnov (K-S) test. It is thus safe to use the grade 10 sample for further test evaluation purposes.

It is a common practice to measure item difficulty in terms of the percentage (or proportion) of test takers who answer it correctly. Although there are different techniques of determining item discrimination index, this research employed the technique available in SPSS 9.0 for windows. Each of the 40 items in the test is analyzed in terms of its difficulty and discrimination index. In addition, the difficulty indices were determined for each of the boys and girls sub-sample. In all samples (total, boys, girls) there are very few items with extreme values of difficulty indices. In other words, there are very few 'very difficult' and 'very easy' items. Most of the items have Ps ranging from 25 to 75.

One of the common features in the samples of students is that the discrimination indices of the items do not have any negative values. This means that, irrespective of the extent of discrimination, the proportion of better students who answered all the items correctly is always greater than that of weak students on the test as a whole. Qualitatively speaking, this is what is expected from a good quality test. The majority of the items in the test have discrimination indices of around 0.7. Finally, a reliability index of $r = .8844$ was obtained using the 42 students in grade 10.

Validity of a test is concerned with what the test measures and how well it does so. It tells us what can be inferred from test scores. Although there are different types of validity, construct validity was considered in this research. Construct validity refers to the extent to which the test may be said to measure a theoretical construct or trait. Such validation requires the gradual accumulation of information from a variety of sources (Anastasi and Urbina, 1997). Two techniques that contribute to construct identification are used in this research.

1. Factor Analysis. This is a particularly relevant technique for construct validation procedures. Factorial validity is essentially the correlation of the test with whatever is common to a group of tests or other indices of behavior. Larger samples are used to determine whether the 5 different subtests of SAT measure different or the same type of spatial ability. The results of the factor analysis have shown only one factor with an eigenvalue greater than 1. Note that an eigenvalue is a measure of standardized variance with a mean of 0 and a standard deviation of 1 (Kinnear and Gray, 1994). The variance that each standardized variable contributes to a principal components extraction is 1. Hence, a component with an eigenvalue of less than 1 is less important than an observed variable, and can therefore be ignored.

Because only one factor is found in the test, one does not need to consider, as far as SAT is considered, the various sub-abilities or factors of spatial ability that are identified in other research work. We can thus say that the whole test measures the spatial ability of the test takers, which is in this case the ability to imagine three-dimensional structures represented as two-dimensional diagrams.

2. Internal Consistency. In this technique, the criterion is none other than the total score on the test itself. In one of the applications of the criterion of internal consistency sub-test scores are correlated with total score (Anastasi and Urbina, 1997). Accordingly, the grade 10 sub-test scores are correlated with their respective total scores. The results have shown that each of the five sub-tests scores correlate highly significantly and positively with their respective total score, evidencing the internal consistency of the entire SAT. The results also tell us the homogeneity of the test.

In general, one can say that the Spatial Ability Test (SAT) has reasonable levels of objectivity, reliability and validity -- of course, for the specified population of students. It can thus be argued that the major research findings obtained using this test and to be reported in subsequent pages are also reliable and valid outcomes (Temechegn, 2000).

Analysis and Interpretation of Data

One of the aims of this research is to compare the spatial ability of students of grade 7—12 in the two different types of schools: Government and Non-government. To this end, data were collected from 762 students, as shown in Table 1, using the spatial ability test (SAT). The data were analyzed using SPSS and the results are presented as follows.

To begin with, the data were analyzed at school type and grade level and the results are presented in Table 2.

Table 2: Descriptive Statistics by School Type and Grade

TYPE	STAT	7	8	9	10	11	12
G	N	99	89	60	73	41	24
	Mean	7.52	8.38	9.33	11.64	14.34	19.04
	St. dev.	2.94	2.83	3.44	4.05	7.98	9.20
	Minimum	2	3	3	3	6	6
	Maximum	16	18	18	27	35	37
NG	N	137	94	47	60	38	
	Mean	10.88	11.19	12.15	13.93	20.47	
	St. dev..	3.66	4.25	4.21	4.52	7.06	
	Minimum	3	4	4	6	10	
	Maximum	21	23	21	26	38	

The research has revealed that the mean scores of NG students are better than that of G students at each grade level. It is also found out, using the Mann-Whitney-U (non-parametric) test at the 0.05 level, that the differences between the two school-types are significant for the total sample at each grade level.

One can argue that if the so called *visual domain* is realized through the use of building blocks, mechanical toys and the like during children's play, the students who likely had access to such playing materials during childhood would perform significantly better than the ones who had not. Building blocks, mechanical toys, drawing materials play such a small part in Ethiopian traditional culture in general. However, the children of the few elite groups are likely to have access to such materials that are available in modern toyshops in Addis Ababa. This might have contributed to the better performance of students of the NG schools as compared to that of the G schools. This generalization, however, needs further investigation. One thing that can be said at this level is that further analysis in terms of boys and girls has confirmed the above generalization only between the boys of NG and that of G schools. The girls in the two school types did not show significant differences in most of the cases.

Based on the above analysis one can say that the first hypothesis, which states *there are no significant differences in spatial ability between the Government and Non-government school students*, can be rejected at the .05 level of significance. That means, the spatial abilities of the total sample of students in the non-government (NG) and in government (G) schools are found to differ significantly, in favor of the former. While the hypothesis is also rejected for similar comparison between boys of the two school types, the girls in the two school types did not differ significantly in their spatial ability and hence the hypothesis is accepted for comparison between girls in the G and in the NG schools.

The second aim of the research is to identify whether the students' spatial abilities show significant differences between two successive grades. Whether the grade levels showing significant differences are dependent on the school types is also a point of investigation. Identifying the developmental nature of spatial ability is another issue to be considered under this type of analysis.

The research has shown that the increase in the spatial ability of the students from grade 7 to 12 is not linear. Rather, a sharp increase in spatial ability is observed at some grade (age) levels. Quantitatively, significant differences between two successive grades are observed only after grade 9. For this analysis, the mean scores for each of the lower grades are subtracted from that of the immediate higher grade. For instance the mean of grade 7 students is subtracted from that of grade 8 students (7/8), and the results are presented in Table 3.

As can be seen from Table 3, the mean scores of the lower grades are always lower than that of the higher grades in both school-types. In other words, the ability tested shows a developmental trend from grade 7 through 12. Such a trend is in agreement with the findings of other researchers reviewed in section 1.2 of this paper.

Table 3: Total Mean Differences between Two Successive Grades by School Type

School Type	7/8	8/9	9/10	10/11	11/12
Government (G)	0.86	0.95	2.31	2.70	4.70*
Non-Government (NG)	0.31	0.96	1.78	6.54*	--

-- = No data in one of the two grades.
* = Significant transitions

Table 3, however, shows that the maximum mean difference appears at different grade levels between the two school types. The highest significant difference appears one grade (age) level earlier in the Non-Government (NG) schools than in the Government (G) schools. The boys and the girls, when analyzed separately, also show the same developmental trend as the total sample does. The Mann-Whitney-U test has also shown that the transitions from 11 to 12 in G and from 10 to 11 in NG are significant at the 0.05 level.

Based on the above analysis the second hypothesis, which states *there are no significant differences in spatial ability of students of the various grade levels in each of the school types,* should be accepted for most of the grade levels and be rejected for specific grades. The grades at which significant difference in spatial ability are observed are different in the different school types.

The third main objective of this study is to identify the difference and/ or similarity between the spatial ability of boys and girls of the selected schools as measured by the spatial ability test (SAT). Are there significant differences between the boys' and the girls' abilities of imagining three-dimensional objects represented as diagrams in each of the two school types?

The research has clearly shown that the sex difference in the total sample is very negligible except in the case of grade 12. Statistical tests using the Mann-Whitney-U at the 0.05 significance level has confirmed the above qualitative observations.

Some three decades earlier the influence of social pressures in bringing about differential skills between boys' and girls' spatial abilities was forwarded. This view argues that the development of spatial abilities is shaped by culturally determined sex-role patterns leading to differential experiences with relevant materials. Accordingly, social pressures will result in boys spending more time on such activities as block building or model construction (Shermann, 1967). In other words, those cultures which tend to be stricter on their female than their male children may inhibit the development of spatial ability more in girls than in boys. This view is also known as the social conformity model.

It is true that in Ethiopia, girls should mostly stay at home and be involved in household chores while the boys go away from home and engage themselves in out-of-home activities. Thus, if the above social conformity model were true, we should have significant sex differences throughout the investigated grades. However, one might argue that significant sex differences are not observed because the data are analyzed by combining the scores of the students who are attending two strictly different school types. This might be true. Nevertheless, it has to be proved by the results of the analysis at school type level, which is the next point of discussion.

In order to examine the extent of sex differences in spatial ability at school type level the data are grouped according to the school types and grades, each grade level in turn is classified into boys and girls. The major descriptive statistics resulting from this type of categorization are presented in Table 4.

Table 4: Means (Standard Deviations) by Grade, School Type, and Sex

Grade Type	7	8	9	10	11	12
G Boys	8.35 (2.84)	8.36 (2.84)	9.56 (3.77)	11.76 (3.81)	14.64 (7.76)	22.09 (8.30)
Girls	6.73 (2.84)	8.40 (2.86)	9.07 (3.07)	11.46 (4.48)	14.00 (8.42)	16.46 (9.43)
NG Boys	11.52 (4.14)	11.83 (3.84)	12.71(4.02)	14.28 (4.35)	20.67 (8.15)	--
Girls	9.89 (2.48)	10.48 (4.63)	11.57 (4.42)	13.29 (4.86)	20.14 (4.91)	

Table 4 shows that the boys score better on the spatial ability test than the girls do in both school types and all grades. However, there is a difference in extent of sex-difference in the two school-type samples. These qualitative descriptions can be made more precise by showing the mean differences (the mean of the boys minus the mean of the girls) and the statistical test of significance of differences (in this case the Mann-Whitney-U test).

The difference between boys and girls in both samples decreases across the grade levels till grade 10. Starting from grade 11, the difference begins to increase in the Government (G) schools, with the grade 12 students showing the maximum sex difference. Both the boys and girls of the grade 12 students in G have greater mean scores than the boys and girls of the grade 11 students. However, the transition from grade 11 to 12 in G is accompanied by a mean difference of 7.45 for the boys but only of 2.46 for the girls. This sharp transition for the boys, coupled with only smooth transition for the girls, has brought about the large sex difference in the G.

The difference in the NG sample of grade 11 decreases a little bit as compared to grade 10. The statistical tests have confirmed the significance of the difference in grade 7 of both G and NG schools at the 0.05 level of significance. However, this significant difference can be attributed to chance factors owing to the smallness of the magnitude of the difference in grade 7 of both samples and also to the absence of significant differences in other grades of both G and NG schools. It should also be noted that, comparatively speaking, the sex differences in the NG are greater than in the G. It can thus be said that the conformity model has no validity at all at least in the Ethiopian samples since little or no sex difference is observed in both types of analysis: by total sample and by school type.

Based on the above analysis of sex differences one can say that the third hypothesis, which states *there are no significant differences in spatial ability between boys and girls of the various grade levels in each of the school types*, is accepted at the 0.05 level of significance.

Discussions, Conclusions and Recommendations

This research has shown that students' spatial abilities in the selected schools are generally low. This means that the average spatial ability of students in the schools has not yet sufficiently developed to interpret 2D diagrams as 3D objects.

It is conventionally held that experiences with geometrical figures and objects have positive contributions to the development of spatial ability. It is probably for this reason that almost all tests of spatial ability use items consisting of geometrical figures and objects. One can thus argue that the absence of well-constructed geometrical figures and shapes in our school textbooks would be one of the factors for the lower spatial ability of the students. Mere presence of the topics and the spatial diagrams, however, would not be sufficient to explain the differences between the two school types. That means, the extent to which these diagrams are illustrated and visualized in school lessons by teachers plays an important role. Experience tells us that little is done in this regard, at least in our government schools. This, in turn, implies that spatial ability is a trainable skill and hence can be improved by strengthening the spatial component of school curricula and lessons using illustrative diagrams and models.

The analysis has shown that there is a general increase in the spatial ability of the students with an increase in grade level or age. This might be due to increased mental maturity across the span of school life. However, significant differences between two successive grades are observed only at some levels. A brief look at these grade levels tells us that there are variations between both government and non-government schools. The NG show the significant transition between two successive grade levels earlier (grade 10 to 11) than the G do (grade 11 to 12). Note that all the students in the two school types have the same years of school experiences and the same average age. In spite of this, the Government schools are the ones which reach this level very late. The girls of these G schools have not at all shown the significant jump at the indicated grades.

The question, however, is why such differences are observed between the two school types. Little has been investigated in this regard. However, discussions with science teachers in the investigated schools give a clue that the NG schools are using self-made and manufactured models in teaching certain science lessons while the teachers in G schools do not usually do this. The contents of a specific science subject for these grade levels were also investigated for this purpose (Temechegn, 2000). Based on these studies, one can say that whenever spatial concepts are taught at a given grade level to a greater depth than at the grade level below it, the students in the upper level perform significantly better on spatial tasks than those in the lower grade level. This is particularly true when the concepts are taught with the help of models.

The analysis of data from a differential perspective in the performance of the two sexes (boys and girls) has shown little or no sex difference in the two school types. It has been argued earlier that the social conformity model does not fit to the observed situation in the schools. The model predicts greater sex differences between boys' and girls' spatial ability in those cultures which tend to be stricter on their females than their male children (Shermann, 1967). The Ethiopian culture is, one can say, conservative towards its female children.

Females are supposed to marry as early as possible. They should stay at home with their mothers to work on domestic tasks. On the contrary, the boys can stay away from home and engage themselves in out-of-home activities. Unlike the prediction of the model, however, the students have shown non-significant differences, except in grade 7. Of course, the magnitudes of the sex differences are a little bit larger in the Non-Government (NG) sample than in the Government (G) samples. The G grade 12 boys have also performed better than the grade 12 girls.

It is thus argued that sex differences in spatial tasks at upper primary and secondary school levels (grades 7-12) are the product of a number of factors in a given society. These factors can be classified as a bit of social conformity, the availability of spatial tasks in the culture and in the school experiences. Mere rigid socialization and conformity is not a sufficient condition for sex differences in spatial ability. Rather, those (the boys) who are free to go away from home for playing should be at the same time provided with spatial tasks by the culture. The spatial experiences in schools should also be presented in a male-biased manner for significant sex differences. These three factors interact with each other. The relative importance of one or the other for students during cross-cultural studies should be weighted in terms of the cultural group and the specific school type students are involved in.

Although the Ethiopian culture demands conformism from its female children, it has little input specially designed to its male children concerning spatial tasks. For instance, the boys play during childhood by constructing differently shaped objects such as cars, animals, etc. with mud. The girls as well play by constructing differently shaped household objects with mud. As they grow older, the boys concentrate more on making dams for small flowing water, playing with bottle tops and marbles by arranging them in order on the ground and trying to hit any one or more of the ordered objects according to the rule. The girls, on the other hand, play more of gymnastic-oriented games by drawing lines and squares on the ground and jumping over the lines according to the rule, jumping with/ over ropes, playing with specific number of small and rounded stones by throwing them one by one upwards and catching them before they fall on the ground. Almost no parents of the Ethiopian children have the financial capacity to control and direct the kind of play their children may be engaged in. Building blocks and mechanical toys also play such a small part in the Ethiopian traditional culture. The Ethiopian schools also have little or no means of providing spatial tasks to their students. The high school curricula are dominated by factual-information in the form of short notes. Even if there are some spatial tasks, the large class sizes do not seem favorable to promote the development of spatial ability in general, and bring sex differences in particular. These facts about the samples seem to agree with the argument forwarded above and can explain why the boys and the girls performed almost similarly on the spatial ability test.

Nevertheless, one thing needs further examination. Although the G and NG students' performances on SAT are significantly different from each other, both school types did not show significant sex differences. It was also pointed out earlier that, relatively speaking, the NG students are from the well-to-do family that can afford paying the school fees in these schools. These students can thus have access at home to some western type toys available in toyshops. It is probably for this reason that the mean differences between the boys and the girls at each grade level in these schools (NG) are relatively larger than in the G. However, the effect of the additional western type toys to the children of the NG might have been masked and could not bring significant sex differences because of their non-typical nature to the Ethiopian culture. The children of both G and NG usually play together and interact with the typical Ethiopian culture, and the additional advantage of the NG students is limited to quantity and type homes and bed rooms and probably what they are exposed to here.

There is little doubt that the ability to construct and mentally manipulate three-dimensional images of drawings is important to students. Many present textbooks and more and more teachers expect students to 'read' the subject rather than visualize models, interpret data displayed on charts and mentally manipulate three-dimensional images. Although school curricula are full of concepts that can be understood only through the application of the ability to form mental images, these concepts are presented without means of visualizing them. Curriculum developers, textbook writers and teachers usually assume that students are able to learn meaningful materials that are presented in verbal form and in diagrams and charts. Implicit in their assumption is that students are able to form mental images of 3D objects presented as 2D diagrams. In other words, they assume that the students' spatial abilities have sufficiently developed to interpret the 2D diagrams as 3D objects.

However, this research on spatial ability of grade 7-12 students has revealed that most of the students have low level of spatial ability. The degree of performance on the spatial tasks is dependent more on the kind of school experiences. If the mathematics, science, geography and other subjects which have spatial components and the use of models contribute to higher spatial ability, then we can reasonably say that spatial ability is a skill that can be improved or developed by educational interventions.

The above results and arguments have important educational implications. If students who have not yet sufficiently developed their spatial ability are introduced through verbal and theoretical means to concepts that require formation of mental images (such as solids and their properties), then they may (1) resort to rote memorization as much as possible, (2) develop misconceptions, or (3) try to avoid such concepts and subjects at all. In any case, the students will be at their disadvantages. Materials memorized for the sake of passing exams would not have any lasting effect; they would easily be forgotten and hence their learning would not have any educational significance. The development of misconceptions would interfere with true understanding of further concepts. Usually, the information being learned by students and the sources, such as teachers and textbooks, which mediate this information, are supposed to be the major factors for such a difficulty (Temechegn, 1997). If misconceptions persist longer, it is difficult to erase them from students' minds.

On the other hand, if students try to avoid such school subjects as science, they are unknowingly minimizing their choice of possible future career. What is worse is that if they are trying to avoid those school subjects which involve mental imaginations, they would be deprived of the ticket for joining the scientific and technical world of work and the most respected areas of study such as engineering, anatomy, solid state chemistry, metallurgy, information technology, etc.

Thus, the questions facing particularly the Ethiopian science educators and teachers are:

- Should our curricula be free of spatial concepts and principles which require a certain level of spatial ability because, at present, our pupils do not have these skills to a sufficient extent?
- Should we, in spite of the lower level of our students' spatial abilities, introduce the spatial concepts to our students?

Because science is to a large extent abstraction and hence requires the students' ability to form mental images, alternative (1) is not plausible. Mental imagination is necessary in science and, one can say, there is no true science without it. On the other hand, answering alternative (2) affirmatively poses the questions 'how' and 'when'.

In general, the overall poor performances of the students tell us that spatial ability is not that much a component of these schools' curricula. Given the social significance of spatial ability in various fields of work, particularly for those students who do not necessarily join higher education institutions, it is recommended that schools pay due attention to help their students develop these skills. However, one should keep in mind that a student with poorly developed spatial ability should not be taught primarily by verbal means. Rather, he/she should be provided with activities that will help improve spatial skills. There may be several ways of doing this depending on the particular type of subject area (for details, see Temechegn, 2000).

Teachers should be trained in how to construct and appropriately use models for specific purposes. Such skills, however, are not likely to be developed in the traditional lecture courses for the would-be teachers. That means, teachers need to get additional in-service training in model-use. Models can be used as demonstration means and be constructed by students as part of the specific subject lessons. Although it is time consuming, the latter approach could help develop a better imagination of the spatial substances and their component parts.

The training of teachers should take into account the importance of teachers' knowledge of the potential uses and limitations of the various models. This is important because all models do not convey the same information about a given substance and because students may wrongly attach meanings to the various parts of the model. Students may extend a model beyond its usefulness and may not know that models have certain drawbacks. It is thus important not only to discuss the models but also to criticize them during specific lessons. The instructional process should therefore explicitly address aspects of the parts that are supported by the model. Moreover, because such an emphasis on introduction and use of models is not part of the University subject-area courses, teachers need to receive additional training on the subject.

Special attention should be given to the Ethiopian conditions. Materials that are available in the developed nations are mostly not available in Africa. Even if they are available, buying them is beyond the financial means of almost all schools, particularly the Government (G) schools. Thus, further research in the coming decade should primarily concentrate on exploring the locally available materials. Some suggestions for chemistry are available in Temechegn (2000).

According to Bruner (in Ausubel, 1968), the task of teaching a subject to a child at any particular age is one of representing the structure of that subject in terms of the child's way of viewing things. Any subject can be taught effectively in some intellectually honest form to any child at any stage of development. The philosophy behind such curriculum organization is the so called the *spiral curriculum*, in which *intuitively learned content may serve as anchoring ideas or as general background for the latter learning of the same content at a higher level of abstraction, thereby increasing its potential meaningfulness* (p. 197). Others also argued that the age at which children can learn a given intellectual task is not an absolute, but is always relative, in part, to the method of instruction employed (p. 182).

However, Ausubel warned us that even if all abstract concepts could be restructured on an intuitive basis, it would still be unreasonable to expect that they could be made comprehensible to children at any grade level. He, instead, argues that the conditions of meaningful learning depend on the material to be learned and the cognitive structure of the learner.

> It is apparent, therefore, that insofar as meaningful learning outcomes in the classroom are concerned, the availability, and other significant properties, of relevant content in different learners' cognitive structures constitute the most crucial and variable determinants of potential meaningfulness (p. 40).

What can be said, in the Ethiopian context, is that owing to the scarcity of mediating materials (such as models) and the low level spatial ability of the students, it is not pedagogically sound to discuss the very abstract concepts at the very beginning (grade 7) of the Ethiopian curricula. Let the grade 7 students observe various phenomena and interpret them in terms of their understandings (Temechegn, 2002).

References

Anderson, M. (1992). **Intelligence and Development: A Cognitive Theory**. UK: Blackwell Publishers.

Anastasi, A. and Ubrina, S. (1997). **Psychological Testing (7th ed.).** New Jersy: Prentice Hall.

Ausubel, D. P. (1968). **Educational Psychology: A Cognitive View.** New York: Holt, Rinehart and Winston, Inc.

Barke, H. D. (1978). *Ein an räumlichen Strukturen orientierter Weg zur Einführung von Grundbegriffen der Chemie im Unterricht*, dissertation at the University of Hannover.

_____ and Kuhrke, R. (1992). *Einführung in die Chemie: Von Alltagsvorstellungen zu ersten Strukturmodellen der Materie, Mädchen in Naturwissenschaften und Technik (MiNT)*, **Band** 4, Frankfurt am Main: Peter Lang.

Ben-Chaim, D., Lappan, G. and Houang, R. T. (1986). *Development and Analysis of a Spatial Visualization Test for Middle School Boys and Girls*. **Perceptual and Motor Skills**, 63, 659-669.

Brislin, R. W. (1980). *Translation and Content Analysis of Oral and Written Materials*, in H. C. Triandis and J. W. Berry (eds.), **Handbook of Cross-Cultural Psychology**, Volume 2: Methodology. Boston: Allyn and Bacon, Inc.

Burnett, S. A. (1986). *Sex-Related Differences in Spatial Ability: Are They Trivial?* **American Psychologist**, 41, 1012-1013.

Caplan, P. J., MacPherson, G. M. and Tobin, P. (1985). Do Sex-Related Differences in Spatial Abilities Exist? A Multilevel Critique with New Data. **American Psychologist**, 40(7), 786-799.

Coleman, S. L. and Gotch, A. J. (1998). *Spatial Perception Skills of Chemistry Students*. J. Chem. **Educ.**, 75 (2), 206-209.

Eliot, J. (1986). Comment on Caplan, MacPherson, and Tobin. **American Psychologist**, 41, 1011.

Guttman, R., Epstein, E. E., Amir, M. and Guttman, L. (1990). *A Structural Theory of Spatial Abilities*. **Applied Psychological Measurement**, 14(3), 217-236.

Hiscock, M. (1986). *On Sex Differences in Spatial Abilities.* **American Psychologist**, 41,1011.

Halpern, D. F. (1986). *A Different Answer to Question, Do Sex-Related Differences in Spatial Abilities Exist?* **American Psychologist**, 41, 1014-1015.

Johnson, E. S. and Meade, A. C. (1987). *Developmental Patterns of Spatial Ability: An Early Sex Difference.* **Child Development**, 58, 725-740.

Kinnear, P. R. and Gray, C. D. (1994). **SPSS for Windows Made Simple**. Hove (UK): Lawrence Erlbaum Associates, Publishers.

Kline, P. (1991). **Intelligence: The Psychometric View**. London: Routledge.

Linn, M. C. and Peterson, A. C. (1985). *Emergence and Characterization of Sex Differences in Spatial Ability: A Meta-Analysis.* **Child Development**, 56, 1479-1498.

Lohman, D. F. (1979). Spatial Ability: A Review and Reanalysis of the Correlational Literature. **Technical Report No. 8, Aptitude Research Project**. School of Education, Stanford University.

Macoby, E. E. and Jacklin, C. N. (1974). **The Psychology of Sex Differences**. Stanford: Stanford University Press.

McArthur, J. M. and Wellner, K. L. (1996). *Reexamining Spatial Ability within a Piagetian Framework.* **Journal of Research in Science Teaching**, 33(10), 1065-1082.0

McGee, M. G. (1979). *Human Spatial Abilities: Psychometric Studies and Environmental, Genetic, Hormonal, and Neurological Influences.* **Psychological Bulletin**, 86(5), 889-918.

Munroe, R. H. , Munroe, R. L. and Brasher, A. (1985). *Precursors of Spatial Ability: A Longitudinal Study among the Logoli of Kenya.* **The Journal of Social Psychology**, 125(1), 23-33.

Newcombe, N. (1982). *Sex-Related Differences in Spatial Ability: Problems and Gaps in Current Approaches*, in Potegal, M. (ed.), Spatial Abilities:

Developmental and Physiological Foundations. New York: Academic Press.

Pellegrino, J. W. and Kail, R., Jr. (1982). *Process Analysis in Spatial Aptitude*, in R. J. Stenberg (ed), **Advances in the Psychology of Intelligence**: Volume 1. New York.

Piaget, J. and Inhelder, B. (1967). **The Child's Conception of Space**. New York: Norton.

Sanders, B., Cohen, M. R. and Soares, M. P. (1986). *The Sex Difference in Spatial Ability: A Rejoinder.* **American Psychologist**, 41, 1015-1016.

Sherman, J. A. (1967). *Problems of Sex Differences in Space Perception and Aspects of Intellectual Functioning.* **Psychological Review**, 74, 290-299.

Temechegn Engida (1997). *The Teaching of Higher Order Thinking Skills in Chemistry.* **Educational Journal**, 3(5), 74-94.

_____ (2000). *Structural Chemistry and Spatial Ability in Chemical Education: A Case of Selected German and Ethiopian Schools.* Unpublished Dissertation, University of Muenster, Germany.

_____ (2002). *Reflections on African Science Education for the New Millennium: The Case of the Ethiopian Chemistry Curriculum for Beginners.* **International Journal of Science Education**, 24(9), 941-953.

Voyer, D., Voyer, S. and Bryden, M. P. (1995). *Magnitude of Sex Differences in Spatial Abilities: A Meta-Analysis and Consideration of Critical Variables.* **Psychological Bulletin,** 117(2), 250-270.