Reports from

THE INSTITUTE OF EDUCATION UNIVERSITY OF GÖTEBORG

AN ECOLOGICAL ANALYSIS OF TEST SCORE CHANGES OVER TIME

Kjell Härnqvist and Gun Stahle

No 64 December 1977

INSTITUTE OF EDUCATION, UNIVERSITY OF GÖTEBORG Fack S-431 20 MÖLNDAL SWEDEN

AN ECOLOGICAL ANALYSIS OF TEST SCORE CHANGE OVER TIME

Kjell Härnqvist and Gun Stahle
Göteborg University, Sweden

This article, based on a thesis by Stahle, was drafted at the Center for Advanced Study in the Behavioral Sciences, where Härnqvist was a Spencer Fellow during the academic year 1976-1977.

An Ecological Analysis of Test Score Changes over Time

Abstract

Average test scores in mental ability tests increased between two comparable national samples of thirteen-year olds in Sweden tested with a five-year interval. The increases were greater for girls than for boys. The changes occurred simultaneously with several changes in social and educational conditions. In order to analyze the relations between test score changes and changes in the environment the samples were divided in ecologically more homogeneous subgoups and test scores were correlated with a number of ecological indices. Test scores were positively related to a global ecological index (density), more so for boys than girls, and more so at the first than the second occasion. The latter finding indicated that an equalization between different areas had taken place during the interval. Changes for boys were positively related to ecological changes in the direction of urbanization, while changes for girls seemed to have more to do with changes in the educational system stressing equality of treatment of boys and girls. This impact was most striking in the spatial visualization test.

An Ecological Analysis of Test Score Changes over Time

In recent years a decline in achievement test scores at pre-college level has been observed in the United States. A detailed analysis of different time series by Harnischfeger and Wiley (1975, 1976) indicated that this decline was real and not an artifact due to changes in test composition, norms, or tested samples. The decline began in the mid-sixties after a continuous upward trend. It involved grades 5 through 12 and almost all tested achievement areas.

A similar decline from 1960 to 1975 in achievement scores was found by Flanagan (1976) who compared 9-11 graders in a sample of Project Talent schools on identical tests, i.e. with stronger control on sample and content than regular testing programs can afford. Some ability tests, however, showed opposite trends, e.g. Abstract Reasoning for male and female students, and Mechanical Reasoning and Visualization for females. In general scores on so-called IQ tests have been increasing over a long period (Thorndike, 1975; a comprehensive survey of earlier research findings is given by Stahle, 1973).

Harnischfeger and Wiley (op. cit.) related the decline in test scores to concomitant changes in several school and out-of-school variables that might have influenced the decline. They made probable that, for instance, tendencies among the high-school students to elect less academic programs were responsible for part of the decline. Available statistical series, however, did not make it possible for them to analyze the relations so closely as they wanted and they indicated a number of additional checks that were needed. Among other things an analysis of factors contributing to increasing achievement, like that observed until the mid-sixties, should be helpful also for understanding causes of decline.

Also from a general development point of view it is of interest to study changes in external conditions as possible change agents for individual traits. Nesselroade and Baltes (1974) stress the importance of time/cohort effects on adolescent personality, but from their review it is also evident that so far very few attempts have been made to systematically analyze the impact of different kinds of historical changes.

The present paper presents one such analysis made in a Swedish study of changes in ability test scores between 1961 and 1966 in national samples from the age cohorts born in 1948 and 1953 and tested at thirteen years of age. The overall results showed a substantial increase between the two cohorts. At the same time several changes in society took place. The present analysis, which is based on a thesis by Stahle (1973), tries to establish links between test score changes and external conditions. Ecological variables are used for the analysis of change factors.

Methods of analysis

A comparison of average test scores between two points in time gives just one difference that can be related to initial states or concomitant changes in other variables. The difference can be specified in various ways, for instance, for tests with different contents, or for males and females, but still the possibilities for a causal analysis remain very weak.If there had been many points in time, a time series for test scores could have been compared to time series for other variables. In the present case it is necessary to find other ways for comparative analysis. One such possibility is to break down the national average into averages for different geographical areas and relate test scores and test score changes to characteristics of the areas.

Even though test scores are available for just two points in time, area characteristics, in principle, could be collected on a time series basis. This would make it possible to study the time lag of test score changes in relation to changes in external conditions, i.e. the time it takes for a certain external change to make an impact on test scores. A full model could look like Figure 1.

Insert Figure 1 about here

It is important to note that this model can be applied only on an aggregate level because that is the only level for which identity can be established between observational units at different points in time. On the individual level the design is, of necessity, cross-sectional: different individuals were tested at t_0 and t_1. This is a rather paradoxical situation because what we really would like to study is an individual level variable (ability) in relation to the individual characteristics of the families in the areas as contrasted to changes in the collective characteristics of the areas. In order to make such a distinction between compositional and contextual effects (cf., e.g., Hauser, Sewell & Alwin, 1976; Alwin, 1976), or in other terms demographic and group-caused effects (cf. Cronbach, 1976), a simultaneous analysis has to be performed at individual and ecological levels, but this possibility is excluded, at least in a strict form, by the nature of the data.

Even the utilization at an aggregate level of such a path model, however, met with several difficulties in the present case. Area characteristics were available only for t_0 and t_1, and this is not information enough for a study of time lag. Furthermore the number of observational units became rather limited in spite of test information for large samples. As will be shown 46 units could be used, and even though test averages for these units, with boys and girls analyzed separately, are based on about 100 pupils on average for these units, some units got rather small. Finally most ecological characteristics of the units turned out to be highly intercorrelated. The small number of observational units and multi-collinearity of the independent variables restrict the number of predictors that can profitably be used in a path analysis (cf. Hanushek & Jackson, 1977, pp. 86-93). Therefore not even a part of the model has been used in its standard form, but rather modifications based on successive findings in the empirical analysis of data. These will be described in direct connection with the reporting of results.

Samples and tests

In 1961 all pupils in Swedish schools who were born on the 5th, 15th or 25th in any month of 1948 took part in the basic data collection for a longitudinal study, the Individual Statistics Project. The same procedure was repeated in 1966 for pupils born in 1953. This means that the project has available data for two nationwide ten percent samples of thirteen year olds. Details about the design of the project and results from other studies have been reported, i.e. by Härnqvist (1966, 1968), Svensson (1971), and Härnqvist and Svensson (1973).

The basic data collections included administration by the classroom teachers of a written intelligence test which was constructed especially for the project and not used in any other connection between the two administrations in 1961 and 1966. The test had three subscales:

Verbal: Opposites (40 items, K-R 20 reliability in subsample .87)

Reasoning: Number Series (40 items, K-R 20 .92)

Spatial: Metal Folding (40 items, K-R 20 .88)

Table 1 shows the number of cases for which complete test data were available in the two samples. These numbers correspond to 87 percent of the expected sample size in 1961 and 88 percent in 1966:

Insert Table 1 about here

Construction of ecological units

Different principles can be used for the construction of ecological units. It is possible to use administrative subdivisions of the country, e.g., counties, municipalities, or other aggregations of adjacent areas. It is also possible to aggregate areas in different parts of the country that have similar characteristics according to population density or other relevant variables. Use of administrative units facilitates the collection of statistics for other variables to be used in the analysis; often they can be taken directly from published official statistics. On the other hand most administrative units have a rather heterogeneous composition which keeps much of the interesting variance in other characteristics within them and leaves relatively little between-variance to be analyzed. Use of population density makes the units fairly homogeneous in different respects but may cause difficulties in the collection of explanatory variables.

In the present study both principles were used. Sweden was divided in nine areas according to location, each consisting of some adjacent provinces ("län"). This subdivision will be called location. The city of Stockholm was one of these nine areas.

The measure of population density used was the number of inhabitants that according to the census of 1960 had their residence within a circle with 20 kilometers' radius around the midpoint of each primary administrative unit ("kommun") in the country. (For a general discussion of this type of index, see Öberg, 1976, pp. 69-71.) There were 1029 such units and the number of inhabitants within their 20 km circle varied from 1 to 1126 thousands. The distribution was heavily skewed with 896 units below 100 thousands. Natural logarithms were used for making the scale more symmetric about the mean and it was divided in fifteen classes of density, where 15 stands for the most densely populated areas. Each class comprised one half unit on the logarithmic scale. The density index had an average of 9.4 and a standard deviation of 3.1, almost exactly the same for males and females in the two cohorts.

In order to get a sufficient number of sufficiently homogeneous units for the ecological analyses the two classifications were used in combination. Each observational unit then became fairly homogeneous in respect of both population density and location in the country, which in turn should have maximized differences between units. When the two indices were tabulated against each other it was found that only 68 of the possible 135 cells contained any of the 1029 primary administrative units.

When the samples of pupils (cf. Table 1) were distributed over the 68 cells according to location and density in combination, 22 of these cells had less than 30 cases of either sex in either year. These cells, altogether containing 329 pupils in 1961 and 358 in 1966, were excluded from the ecological part of the study. This leaves 46 cells with an average cell size of 111 in 1961 and 99 in 1966 when test scores for boys and girls are analyzed separately.

Ecological variables

In addition to the location and density indices already described ecological variables were drawn from two sources:

1) From the project's own data in 1961 and 1966. This category includes measures of family background (e.g., percent of the pupils in a cell whose fathers graduated from secondary education), and school organization (percent of the pupils in a cell that attend a comprehensive school). These indices are based on male and female pupils together, i.e., on twice as many cases as the test score averages. For each variable one index was computed for 1961, one for 1966, and in addition a difference between the two measures. This resulted in 23 x 3 variables based on data from the project. After the exclusion of indices with very small variation between cells or very skewed distributions 13 initial state measures and 14 change measures remained for further analysis. The variables kept are listed in Appendix 1.

2) From official statistics available at the level of primary administrative units ("kommuner"). With one exception these indices refer only to 1960 or 1961. The variables cover such things as trends in population development, average taxable income of economically active inhabitants, housing standards, and occupational distribution of economically active men and women. Altogether these categories include 14 variables. In addition the number of television licenses per thousand inhabitants was recorded for 1961 and 1966 as well as the difference. After exclusion of variables on the basis of their distributions 12 initial state measures and 1 change measure remained for analysis. The variables kept are listed in Appendix 2.

The two kinds of ecological variables differ in principle. The first kind, in most instances, refers to the characteristics of the samples studied in the project, i.e. to families in age groups that have thirteen year old children. The other kind refers to characteristics of the whole population in the administrative unit and the environment they live in. Unfortunately almost all change measures are of the first kind which complicates their interpretation. Does a change mean a change in the community characteristics, or a difference between families from different age cohorts, or a difference due to sampling fluctuations? These alternatives have to be considered in the interpretation of results.

One indication that measures derived from the project data are characteristic also for the whole communities is found in the correlations between initial state measures. The average correlations within and between different categories are shown in Table 2.

Insert Table 2 about here

Project-based indices (P) correlate nearly as much with those based on official statistics (S) as between themselves, and nearly as much with density as the official statistical measures do.

For the whole set of 25 initial state measures a first principal component accounted for 63 percent of the variance, and only three more components had eigenvalues exceeding one. All four together covered 82 percent of the variance.

The first varimax rotated factor had its highest loadings in the indices based on percent of fathers and mothers with secondary school graduation (P), and percent of men in white-collar occupations (S). The 14 variables that loaded .70 or more on the first factor had an average correlation of .84 with density.

The next three factors are best characterized by the following indices:

No. 2 - Percent of men in non-agricultural working-class occupations (S).

No. 3 - Housing units built in 1951-1960 in percent of all housing units in 1960 (S).

No. 4 - Percent of fathers in commercial occupations (P).

The change measures derived from project data were less highly intercorrelated than the initial measures. In this case the first principal component accounted for only 24 percent of the variance. Five more components had eigenvalues above one, and the six together covered 83 percent of the variance in the 14 measures. Only four of the variables correlated significantly with density. Together they indicated that there was a comparatively larger transfer from agricultural to non-agricultural working-class occupations, and a larger increase in the occupational activity of the mothers, in the less densely populated areas, i.e., in those where there were substantial groups in the initial state to draw from.

Results I: Overall changes between 1961 and 1966

Table 3 reports means and standard deviations based on the individual scores of pupils tested in the two cohorts.

Insert Table 3 about here

All average scores increased between the two years and most markedly so for the girls for whom the increase amounted to a fifth of the initial standard deviations of total scores. The group differences found in the table were checked in a two-way analysis of variance (unweighted means) and the results were expressed in terms of contributions to overall variation in Table 4.

Insert Table 4 about here

Cohort differences contributed most to the variation in the verbal, reasoning and total test scores, while sex differences had a stronger impact in the spatial test. In that test, a significant interaction was found as an expression of a considerable reduction of the sex difference between the cohorts. The same held for the total score.

The changes in test scores over the five-year period were simultaneous to overall changes in some of the ecological indices. Table 5 shows the size of the changes that were significant when analyzed on the Density x Location level.

Insert Table 5 about here

During the five-year interval a major part of the implementation of two innovations took place at the national level. Compulsory education was extended from eight to nine years, a selective lower-secondary school that existed parallell to grades 7-8 was integrated in a comprehensive compulsory school, and the curriculum was revised for all grades. In 1961 32 percent of the sample attended an experimental version of the comprehensive school, in 1966 81 percent were in the new system which was fully implemented in grade 6 in 1970, through grade 9 in 1973.

The other innovation was the spread of television over the country. During the interval satisfactory receiving conditions were achieved in remote areas, and the frequency of licenses required for using a TV set nearly doubled. The increase was negatively correlated with density (-.70).

The other changes recorded in Table 5 are of the order around 30 percent and, with one exception, in the positive direction. The two highest socioeconomic groups increased their share of the families; so did graduation from secondary education among fathers and mothers. Increased participation by mothers in the labor market was shown in two indices. The only decrease over the interval was for the agricultural socioeconomic group. As already was mentioned increased occupational activity among mothers and transfer from agriculture occurred more often in the sparsely populated areas. This location of the change suggests that the ecological changes shown in Table 4 are likely to be more general than just referring to the characteristics of two groups of parents with a five-year interval.

Results II: Test scores related to density

The density index was highly correlated with most of the initial ecological variables and to a moderate extent also with some of the changes. Therefore we will begin with showing the relation between density, taken as a global ecological index, and test scores in both cohorts. These relations then will be specified with the help of other variables in the two following sections.

Regression slopes for test scores in relation to density are reported in Table 6.

Insert Table 6 about here

One step upwards on the density scale means roughly 50 percent increase of the population found within 20 km radius. In 1961 such a step led to an average increase in the total test score of .79 unit for boys and .51 for girls. In 1966 these increases were signficantly smaller (.45 and .20). Similar decreases in regression coefficients were found in all subtests even though they were not significant in Number Series and Metal Folding for girls. The scores of the boys were significantly more related to density than those of the girls, except in Opposites both years and Metal Folding in 1966.

When differences in means and in regression slopes were combined an interesting overall pattern of the changes appeared, and it is shown in Figures 2-4.

Insert Figures 2-4 about here

In the verbal test both boys and girls changed upwards all along the density scale but more so in the sparsely populated areas than at the top of the density scale.

In the reasoning test and the spatial test the girls changed upwards all along the density scale, but test scores for boys increased only in the lower parts of the density scale. Boys excelled girls in the densely populated areas; much less so, or not at all, in the sparsely populated areas.

With this we have established (1) that there are relations between a global ecological variable and test scores and and test score changes, and (2) that their relations differ both between boys and girls and between different abilities.

Next we will see if the other variables can contribute to an understanding of these relationships.

Results III: Initial test scores related to ecological variables

Multicollinearity, in combination with the limited number of observations for each variable (N=46), would make straightforward multiple prediction of test score averages from the whole matrix of ecological variables rather misleading. Therefore three alternative approaches have been used -- all of them in order to detect more specific relations between area characteristics and test scores than the density index can indicate.

1) Test score averages were correlated with the density index and with each one of the 25 variables derived from project data (P) and official statistics (S). Then density was partialled out of the other ecological variables and the remainder was correlated with test scores (part, not partial, correlation).

2) One characteristic variable from each of the first four factors derived from the ecological indices was used for predicting test score averages, for one subtest at a time, by means of multiple regression.

3) The same predictor variables were used also in a canonical correlation with the three subtests simultaneously on the criterion side.

Table 7 shows the results of the first approach. Only part correlations significant at 5 percent level are recorded. Ecological variables are listed in order of their loadings on the first factor.

Insert Table 7 about here

The correlations with the density index is another way to demonstrate what already has been shown by the regression coefficients in Table 6. The relations are stronger for boys than for girls, and for the verbal than for the reasoning and spatial tests.

Significant part correlations appear mainly for variables that have high loadings on the first factor. This strengthens the impression gained from the results for density that test scores on an ecological level vary positively with characteristics that distinguish urban from rural areas -- more people with education above the compulsory level, more white-collar jobs, higher taxable incomes, and better housing.

Table 8 shows the results of the multiple and canonical correlation approaches, together with the zero order correlations for the predictor variables.

Insert Table 8 about here

The addition of three more dimensions increases the variances explained with 3-6 percentage units for boys and 4-9 for girls. The larger increment for girls is due to higher correlations for new-built housing and negative correlations for commercial occupations.

The canonical weighting of the tests results in a slight increase over the multiple correlation for unweighted totals for boys, and in a large increase for girls. The difference is due to the larger variation in correlations between tests for girls.

In the optimal case the ecological indices explain 55 percent of the variance in test score means for boys and 47 percent for girls. More could hardly be expected since the test score averages have a considerable sampling variation also. With the group sizes involved the expected sampling variances amount to between 50 and 60 percent, on average, of the observed variances for test score means.

Results IV: Test score increases related to ecological variables

The correlations between test score changes and density are much smaller than those reported in the previous section and thus it does not seem necessary to partial out density from other ecological correlations with changes. Neither are the correlations between ecological changes high enough to require a selection among the explanatory variables. Therefore zero order correlations will be used in this section.

Table 9 reports the correlation between test score increases, density and ecological indices in the initial state. Except in relation to density, only correlations significant at the 5 percent level will be listed.

Insert Table 9 about here

Density has small negative correlations with all test score increases except Metal Folding for girls. Only the one with Opposites is significant, however. The general picture is one of decreasing variation between levels of density over the period -- a trend that was more easily observed in the regressions of individual scores in Table 6 and Figures 2-4.

The correlations between initial state measures and test score increments support this notion of equalization between areas. Variables that are related to middle-class and urban living conditions are negatively correlated with increments; those related to working-class and agricultural occupations positively correlated. The negative correlation for "Workers among mothers" does not really mean an exception to this generalization since this variable has to be seen in contrast to the incidence of occupationally non-active women rather than to that of white-collar occupations. Thus it seems that areas which initially lagged behind in intellectual performance became somewhat better off during the five-year period.

A comparison between the variables in Table 9 and those that had significant part correlations with initial test scores (Table 7) shows that several of the variables are the same in both tables. This is especially true for the girls for whom all five variables in Table 7 appear also in Table 9. In all these cases positive part correlations in Table 6 correspond to negative correlations of roughly the same size in Table 8. Areas which initially scored higher than expected from their density lost this surplus over the period. This looks like an interesting substantive result but there is reason to suspect that it, to some extent, depends on sampling variation. If the sample from a certain district that took the test in 1961 scored above the mean for all pupils in the age cohort in the district, then the sample mean for the same district in 1966 is likely to have regressed towards its population mean resulting in a negative difference over the period. Such mechanisms could give rise to the pattern shown by Tables 7 and 9 in combination. If so, the correlation between initial measures and test score changes should not be interpreted too specifically. Still the indications of an equalization between more and less densely populated areas remain from the analysis of individual scores.

Finally, in Table 10, significant correlations between test score increases and concomitant changes in ecological variables are reported.

Insert Table 10 about here

The significant correlations are few but positive which means that improvements covary in a consistent manner. Most of the correlations have to do with increase in the proportion of secondary education and white-collar occupations -- three of the correlations interestingly enough with increases among the mothers. For these correlations three different interpretations have to be considered. One is that the increase in the ecological index meant an increase for the areas as such, which was accompanied by test score increases. The other two have to do with the nature of the indices as derived from project data. One alternative is that the increases are real but refer only to cohort differences between families having thirteen-year-old children in 1961 and 1966. Another alternative is that the increases are spurious and reflect simultaneous sampling fluctuations in test scores and background characteristics (correlated errors).

For two of the variables, however, the last two explanations cannot reasonably hold. A positive correlation between increase in socioeconomic group E (non-agricultural working-class occupations) and test score increases is not likely to result from cohort increments of this group among the parents, neither from over-sampling of working-class pupils, because pupils with this background tend to score below average -- not above. The increase in comprehensive school coverage refers to all pupils in the district and not only to the tenth of the age cohort that was sampled.

In the first instance the ecological change is likely to indicate an increase in urbanization and industrialization of the district, similar to that registered along the density scale and its correlation with change measures. The other educational and occupational indices in Table 10 fit into the same pattern.

The increased coverage of the comprehensive school could, in principle, be an indication of some similar trend but since it is rather unrelated to other ecological variables it might be regarded as a specific change agent. Primarly it is related to the increase in spatial test scores among the girls -- the most striking change in relation to traditional sex differences. Figure 5 shows the relation between increases in the two variables at each level of the density scale.

Insert Figure 5 about here

The two sets of averages are very little related to the density scale but follow each other closely. In the discussion we will try to interpret this in the context of the other findings.

Discussion and conclusions

We began the presentation of results by showing that an overall increase in test scores occurred between national samples in 1961 and 1966. The increase was larger for girls than for boys. During the same interval several changes in the environment took place. School curriculum was revised. Television reached a broader audience. Families with teenagers moved upwards along the educational and occupational ladders. Occupational activity increased among mothers. Agricultural occupations became less frequent.

In trying to establish causal relations between environmental and test score changes we divided the country in smaller areas characterized by a global ecological variable -- density -- a location index,and a number of more specific ecological indices for each Density x Location group based on project data and official statistics.

Initial test scores were rather strongly related to density -- for boys generally and for girls at least in verbal ability. The ecological variables that were chosen to reflect more specific characteristics of the areas, or the families in the areas, were in general highly correlated between themselves and with density, and being so they were not able to very much strengthen or specify the relationship between performance and environment.

Over the five-year period the relation between test scores and density was reduced and test scores increased more in sparsely than in densely populated areas. Quite conceivably this could have to do with changes in the density of the administrative units between the census of 1960 on which the index was based, and 1966 for which the same index was used. Not many units, however, can have changed places along the scale during a five-year period. A change of one step on the density scale takes, on average, 50 percent population increase or 33 percent population decrease. The average population growth in one year (1960-1961) was 6.3 per mill with a standard deviation of 18.4 per mill among the units. With five such consecutive increments, all of them two standard deviations above the mean, the expected increase would be 23 percent. With five such consecutive decrements, all of them two standard deviations below the mean, the expected decrease would be 6 percent. Both cases are highly unlikely and would not result in a change in density except for units near the borderlines of the density classes. According to this, lack of constancy in the density classification is not likely to be responsible for the reduced relation between density and test scores. It rather seems that the character of the Density x Location units changed and more so in the sparsely populated areas. It would mean that a real equalization took place between more and less densely populated areas of the country.

One potential equalizing influence was the spread of TV over the country which affected sparsely populated areas more than average (-.70). TV increase, however, was not found among the variables that had significant correlations with test score increments. Still there was a slight indication of influence on gains in verbal as contrasted to spatial test scores among girls (Table 11).

Insert Table 11 about here

Another potential equalizer on the national level was the implementation of the comprehensive school. It had positive correlations with all test scores, some of them even relatively large (Table 11).

Even though comprehensive school growth was unrelated to density (-.03) it might have had different impact in rural than in urban areas. In rural areas it created a possibility of continuing with an academic type of lower-secondary education in the home community which had not been available in the segregated system. This certainly increased the aspirations for academic choices (cf. Härnqvist, 1966) and maybe indirectly also test scores. Furthermore the new curriculum stressed equality between boys and girls in various ways, perhaps most concretely so in handicraft instruction. Traditionally only the girls had done needlework and only the boys woodwork. The new curriculum required a minimum of 20 hours of woodwork for girls and 20 hours of needlework for boys. Although this is a minor part of the handicraft instruction, let alone the total program, it should have introduced girls to interpreting blueprints of three-dimensional objects -- very similar to the task in the spatial test. Even more important, however, this and other curriculum changes might have influenced the expectations as to what girls should be able to accomplish. This is our interpretation of the "surplus" correlation for the spatial test.

This result has somewhat of a parallell in Flanagan's study (1976) and he describes his findings in the following terms.

Both of the remaining two tests have been found predictive of success in mechanical work. Interestingly, both show modest gains for the females and very slight losses for the males. -- This appears to reflect a change away from the sex stereotyping that characterized the 1960 group.

Another parallell is found in a study by Fennema and Sherman (1977) where sex differences in spatial visualization disappeared when the number of "space related" courses was controlled by means of analysis of covariance.

Also findings in Figure 4 can be interpreted in terms of sex roles and varying similarity in the experiences of boys and girls. In rural areas boys and girls are likely to have had more similar everyday experiences of practical things and activities, and this could explain why sex differences in the non-verbal tests are so much smaller in rural than in urban areas. It is tempting to guess that they might have been even smaller if the increased exposure to TV had not had a differential influence on abilities in the opposite direction but admittedly these correlations fail to reach significance.

Seven more indices based on project data changed significantly between the two years (cf. Table 5). In the initial state all of them were highly correlated with density (mean .82, range .73 - .88) and all had high loadings on the first ecological factor (mean .80, range .67 - .90). The overall development in the Density x Location units went in the direction of urbanization and industrialization. This trend was stronger in the less densely populated units as shown by the correlations between density and change in ecological indices (more non-agricultural workers, higher occupational activity among mothers).

Four of the seven indices were represented by significant test change correlations in Table 10 and these supported the interpretation that gains were larger in areas developing in the direction of urbanization-industrialization. Taking all test change correlations for the seven ecological change measures into account makes the pattern more confusing, however, as the distributions of correlations for boys and girls show (Table 12).

Insert Table 12 about here

The correlations for boys support the interpretation that the changes registered in the ecological variables to some extent have been instrumental in improving the conditions for intellectual growth. Whether it is on the family level or the area level remains inconclusive and so does the question whether they are "truly environmental" or not. The correlations for girls, on the contrary, do not support this hypothesis, but instead they are compatible with the earlier findings that both initial and final test scores, in general, are less related to density for girls than for boys - a finding that may indicate that urban and rural environments function more equally in relation to girls than to boys. Still the girls' test scores have increased and in some respects even more so than those of the boys. Looking back upon the results for the two "national innovations" -- TV and comprehensive school - we find the opposite pattern. On the whole these were more related to the girls' than to the boys' test scores.

Thus it seems that both initial level and gains for boys were more influenced by conditions in their immediate environment represented by the educational and occupational level of the family or in the area. Macrolevel conditions had more impact on the girls. Interestingly, the largest gain in the verbal test was for girls in working-class families in communities where upper-secondary education was not available (Svensson, 1971, Appendix 2), and this is not a group where one expects the most effective stimulation from the immediate environment. Rather a compensatory influence is needed and this would have to come through more universal channels -- mass media, school, and contacts with other social groups. It does not even have to be exclusively an intellectual stimulation. A change in expectations and self-concepts -- like the one aimed at in the comprehensive school and propagated in contemporary society -- might also have an impact on girls' performances. As Maccoby and Jacklin (1974, p. 133) conclude:

> --- the studies on personality correlates of intellectual performance have continued to suggest that intellectual development in girls is fostered by their being assertive and active, and having a sense that they can control, by their own actions, the events that affect their lives. These factors appear to be less important in the intellectual development of boys -- perhaps because they are already sufficiently assertive and have a sufficient sense of personal control over events, so that other issues (e.g. how well they can control aggressive impulses) become more important in how successfully they can exploit their intellectual potential.

Some further support for the generalization about girls is provided by Wernersson (1977, Ch. 7) who found that sex-typing of activities and behavior was less frequent among girls with a more conceptualizing cognitive style than among boys generally and among more "concrete" girls. The relation between intellectual performance and personality correlates is even likely to be circular: a more assertive outlook fostering intellectual development, and a more conceptual approach reducing sex differentiation in attitudes and behavior, and so on.

The "other issues" relevant for boys according to Maccoby and Jacklin might possibly be construed to include also the following observations. In general, boys are less well adjusted than girls to the institutional demands set up by school (Svensson, 1971, pp. 122-124). They also are more dependent than girls on occupational goals vs. academic interests in their choice of an education above the compulsory (Härnqvist, in press). In the "status attainment process" (Sewell & Hauser, e.g. 1976), where these choices play an important role, socioeconomic background, parental encouragement and the educational aspirations of friends are highly influential. These findings together suggest that, in order to channel boys' activities in a cognitive and academic direction, educational and occupational goals need to be set up and reinforced by the immediate environment. This would explain the boys' side of the pattern of correlations.

Still there is not information enough to weigh influences from the boys' own families against possible contextual influences from the opportunities, values and beliefs characterizing the communities where they grew up. Ecological change measures based on statistics from the whole communities might have helped somewhat, but considering the high correlations between almost all kinds of ecological indices in the initial state, and the relatively small number of observational units, such additional measures would not have been likely to bring this analysis much forward. As already has been stated a complete multilevel analysis is outside the possibilities of the present approach.

Another deficiency from an analytical point of view is the lack of time series information on ecological characteristics which makes the study of time lag impossible. This lack of information is not inherent in the design but it would have taken an immense amount of work to supply such additional data.

Still it is possible to speculate about time lags. The ecological changes that seem to have influenced gains among boys might very well be just the latest observable part of a development that has been going on for a long time in the direction of increased educational level and transfer to non-agricultural occupations in the less densely populated areas. The full model of Figure 1 would have made it possible to estimate an optimal interval between changes in independent and dependent variables. It seems to us quite likely that this is longer than the interval observed in the study.

The impact on girls of comprehensive school and TV looks less gradual because both innovations were quite new and, on the individual level, one either had them or not at the two points in time. The "message", on the other hand, that especially the comprehensive school was thought to convey -- the changing notion of female roles in society -- was not that new, but its dissemination certainly gained momentum during the '60s. If there had been explanatory variables that recorded the spread of the message instead of the media, it might have been possible to estimate time lags also in this connection. But this is far out from what could be foreseen when this study was planned.

In conlusion we think that we have been able to establish several important relations between ecological characteristics and test score changes. The most interesting results have to do with decreasing differences between boys and girls and their differential response to ecological conditions. These results cannot be verified through replication of the study because it is of necessity bound to a rather special context of societal changes and innovations. Still it may give circumstantial evidence to hypotheses that also can be studied in other contexts. Furthermore it points to methodological approaches and improvements that can be applied in other studies of changes in psychological variables as related to societal trends.

Finally we have asked ourselves whether this study of increasing test scores has any implications for the study of decreasing scores also. In spite of the fact that the observations relate to different domains (ability vs. achievement) and different age groups (elementary vs. secondary) some suggestions seem to be possible.

The present study indicates that it may be fruitful to analyze the causation separately for male and female students even where overall trends are similar for both sexes. The generally positive development over time of the socioeconomic indices that were related to growth among boys may seem incompatible with decreasing scores. However, these indices can be looked upon as proxies for parental and environmental encouragement, and the level of encouragement might very well have decreased over time in relation to the educational and occupational level of the family and its surroundings. A more direct measure of encouragement then would be needed in order to establish similar relations in a phase of declining scores. As to the girls' test scores Harnischfeger and Wiley (1975, p. 117) already have the hypothesis that score losses among girls have been more affected by changes in curriculum, work habits and television viewing than those among the boys. This is a somewhat different interpretation than ours of the impact of these factors, but still compatible with the notion that boys and girls react differently to environmental conditions.

REFERENCES

Alwin, D. F. Structural equation models of contextual effects. Conference paper, October 1976.

Cronbach, L. J. Research on classrooms and schools: Formulation of questions, design, and analysis. Occasional Papers of the Stanford Evaluation Consortium, July 1976.

Flanagan, J. C. Changes in school levels of achievement: Project TALENT ten and fifteen year retests. Educational Researcher, 1976, 5, (September), 9-12.

Hanushek, E. A. & Jackson, J. E. Statistical methods for social scientists. New York: Academic Press, 1977.

Harnischfeger, A. & Wiley, D. E. Achievement test scores decline: Do we need to worry? St. Louis, Missiouri: CEMREL, 1975.

Harnischfeger, A. & Wiley, D. E. Achievement test scores drop. So what? Educational Researcher, 1976, 5,(March), 5-12.

Härnqvist, K. Social factors and educational choice. International Journal of Educational Sciences, 1966, 1, 87-102.

Härnqvist, K. Relative changes in intelligence from 13 to 18. Scandinavian Journal of Pscychology, 1968, 9, 50-82.

Härnqvist, K. Individual demand for education. Paris: OECD (in press).

Härnqvist, K. & Svensson, A. A Swedish data bank for studies of educational development. Sociological Microjournal, 1973, 7, 35-42.

Hauser, R. M., Sewell, W. H. & Alwin, D. F. High school effects on achievement. In W. H. Sewell, R. M. Hauser & D. L. Featherman (Eds.) Schooling and achievement in American society. New York: Academic Press, 1976.

Maccoby, E. E. & Jacklin, C. N. The psychology of sex differences. Stanford, Ca.: Stanford univ. Press, 1974.

Nesselroade, J. R. & Baltes, P. B. Adolescent personality development and historical change: 1970-1972. Monographs of the Society for Research in Child Development, 1974, 39, no. 1.

Öberg, S. Methods of describing physical access to supply points. Lund: Gleerup, 1976.

Sewell, W. H. & Hauser, R. M. Causes and consequences of higher education: Models of the status attainment process. In W. H. Sewell, R. M. Hauser & D. L. Featherman (Eds.) Schooling and achievement in American Society. New York: Academic Press, 1976.

Stahle, G. Regionala förändringar i testresultat under en femårsperiod. Licentiatavhandlingar från Pedagogiska institutionen, Göteborgs universitet, Maj 1973.

Svensson, A. Relative achievement: School performance in relation to intelligence, sex and home environment. Stockholm: Almqvist & Wiksell, 1971.

Thorndike, R. L. Mr. Binet's test 70 years later. Educational Researcher, 1977, 4, (May), 3-7.

Wernersson, I. Könsdifferentiering i grundskolan. Göteborg: Acta Universitatis Gothoburgensis, 1977.

Table 1. Pupils with complete test data.

	1961	1966
Males	5382	4258
Females	5180	4673
Total	10562	9431

Table 2. Average correlations within and between categories of ecological indices. (Density x Location group as unit of analysis).

Indices from	Indices from Project data	Official statistics
Project data	.58	.56
Official statistics	.56	.62
Density	.71	.73

Table 3. Means and standard deviations of test scores in 1961 and 1966 (Individual pupil as unit of analysis).

	Males			Females		
	1961	1966		1961	1966	
(a) Means						
Opposites	22.27	23.49	1.22	22.46	23.99	[illegible]
Number Series	19.64	20.24	0.60	19.05	20.09	[illegible]
Metal Folding	21.84	22.38	0.54	20.26	21.49	[illegible]
Total	63.75	66.11	2.36	61.77	65.58	[illegible]
(b) Standard deviations						
Opposites	6.92	6.73		7.12	6.74	
Number Series	8.08	8.27		7.81	7.91	
Metal Folding	7.52	7.64		6.88	7.03	
Total	18.26	18.41		17.86	17.55	

	61	66
	♂ - ♀	♂ - ♀
O	-0.19	-0.50
N	0.59	0.15
M	1.58	0.89

Table 4. Group-differences in Table 2 expressed as square roots of variance components (σ) (Significant values underlined).

	Sex difference	Cohort difference	Interaction
Opposites	.24	.97	.12
Number Series	.25	.58	.18
Metal Folding	.87	.62	.33
Total	.87	2.17	.67

Table 5. Average ecological indices 1961 and 1966 (Density x Location group as unit of analysis).

In percent of samples:	1961	1966
Comprehensive school coverage	32.0	81.3
Socioeconomic group A (professional and managerial)	4.6	6.4
Socioeconomic group B (white-collar, higher level)	8.3	10.8
Socioeconomic group D (agricultural)	14.7	11.2
Secondary education among fathers	13.5	17.7
Secondary education among mothers	12.4	16.0
Salaried employees, lower level, among mothers	4.0	5.4
Occupational activity among mothers	21.5	25.5
Per thousand inhabitants:		
TV licenses	141	267

Table 6. Regression slopes for test scores on density (Individual pupil as unit of analysis).

	Males		Females	
	1961	1966	1961	1966
Opposites	.265	.172	.291	.127
Number Series	.253	.138	.103	.020
Metal Folding	.268	.138	.125	.054
Total	.787	.449	.511	.200

Table 7. Correlations between test score averages in 1961 and density, and part correlations between test score averages and ecological variables with density controlled (Density x Location group as unit of analysis).

		Oppo-sites	Number Series	Metal Folding	Total
Correlations with density index	Males	.66	.59	.49	.68
	Females	.59	.34	.22	.47
Males Part correlations with:					
Secondary education among mothers (P)		.42	.	.33	.
Socioeconomic group B (white-collar, higher level) (P)		.	.	.33	.
Taxable income (S)		.	.	.38	.
Women in agricultural occupations (S)		-.53	.	.	-.42
Females Part correlations with:					
Men in white-collar occupations (S)		.37	.30	.	.32
Taxable income (S)		.42	.	.33	.32
Socioeconomic group A (professional and managerial) (P)		.40	.	.	.31
Salaried employees, lower level, among mothers (P)		.35	.	.31	.
Housing units with bath or shower (S)		.40	.30	.29	.39

Table 8. Zero-order, multiple and canonical correlations for selected variables (Density x Location group as unit of analysis).

	Oppo-sites	Number Series	Metal Folding	Total
Males				
Secondary education among fathers (P)	.68	.48	.55	.67
Men in non-agricultural working-class occupations (S)	.28	.31	.22	.32
Housing units built in 1951-1960 (S)	.21	.13	.15	.20
Commercial occupations among fathers (P)	.31	.13	.09	.21
Multiple correlations	.72	.54	.58	.72
Canonical correlation		.74		
Females				
Secondary education among fathers (P)	.56	.32	.23	.46
Men in non-agricultural working-class occupations (S)	.36	.25	.09	.27
Housing units built in 1951-60 (S)	.41	.25	.07	.31
Commercial occupations among fathers (P)	-.12	.06	-.17	-.09
Multiple correlations	.65	.37	.32	.52
Canonical correlation		.69		

Table 9. Correlations between test score increases, density, and initial ecological variables (Density x Location group as unit of analysis).

	Opposites	Number Series	Metal Folding	Total
Males				
Density	-.27	-.21	-.05	-.21
Women in agricultural occupations (S)	.31	.	.	.
Occupationally active women (S)	-.30	-.31	.	-.31
Socioeconomic group C (white-collar, lower level) (P)	-.32	-.41	.	-.34
Workers among mothers (P)	.	-.30	.	.
Socioeconomic group E (workers) (P)	.32	.	.	.32
Commercial occupations among fathers (P)	-.37	.	.	.
Females				
Density	-.30	-.09	.08	-.13
Men in white-collar occupations (S)	-.33	.	.	.
Taxable income (S)	-.33	.	.	.
Socioeconomic group (A) (professional and managerial) (P)	-.37	.	.	.
Men in agricultural occupations (S)	.33	.	.	.
Salaried employees, lower level, among mothers (P)	-.35	.	.	.
Women in agricultural occupations (S)	.29	.	.	.
Occupationally active women (S)	-.32	.	.	.
Housing units with bath or shower (S)	-.39	.	.	.
TV licenses (S)	-.30	.	.	.

Table 10. Correlations between test score increases in ecological variables (Density x Location group as unit of analysis).

Increase in:	Opposites	Number Series	Metal Folding	Total
Males				
Secondary education among mothers (P)	.46	.	.	.
Salaried employees, lower level, among fathers (P)	.	.	.31	.
Salaried employees, lower level, among mothers (P)	.	.	.50	.31
Socioeconomic group B (white-collar, higher level) (P)	.33	.	.	.31
Females				
Salaried employees, lower level, among mothers (P)	.	.34	.	.
Socioeconomic group E (workers) (P)	.30	.	.	.
Comprehensive school coverage (P)	.	.	.46	.40

Table 11. Correlations between test score increases and increases in TV and comprehensive school coverage (Density x Location group as unit of analysis)

	Increase in: Opposites	Number Series	Metal Folding	Total
TV increase				
Males	.10	.04	-.06	.03
Females	.19	-.01	-.15	.00
Comprehensive school increase				
Males	.05	.21	.27	.23
Females	.24	.23	.46	.40

Table 12. Distributions of correlations between test changes (4 variables) and ecological changes (7 variables) (Density x Location group as unit of analysis).

Correlation		Males	Females
below -.20		-	3
-.20 - -.01		3	10
.00 - +.20		14	13
above +.20		11	2
	All	28	28

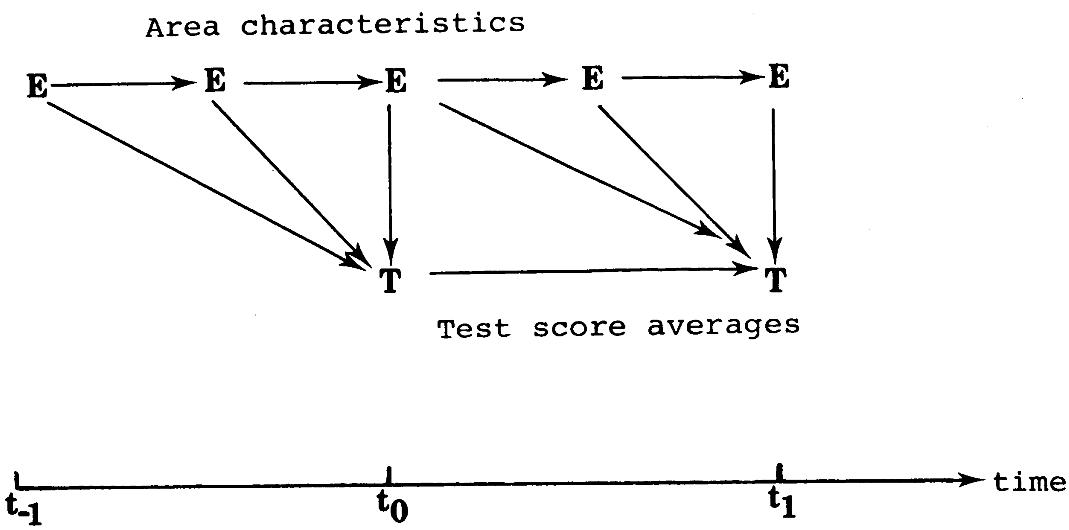

Figure 1. Path model for test score changes in relation to ecological characteristics.

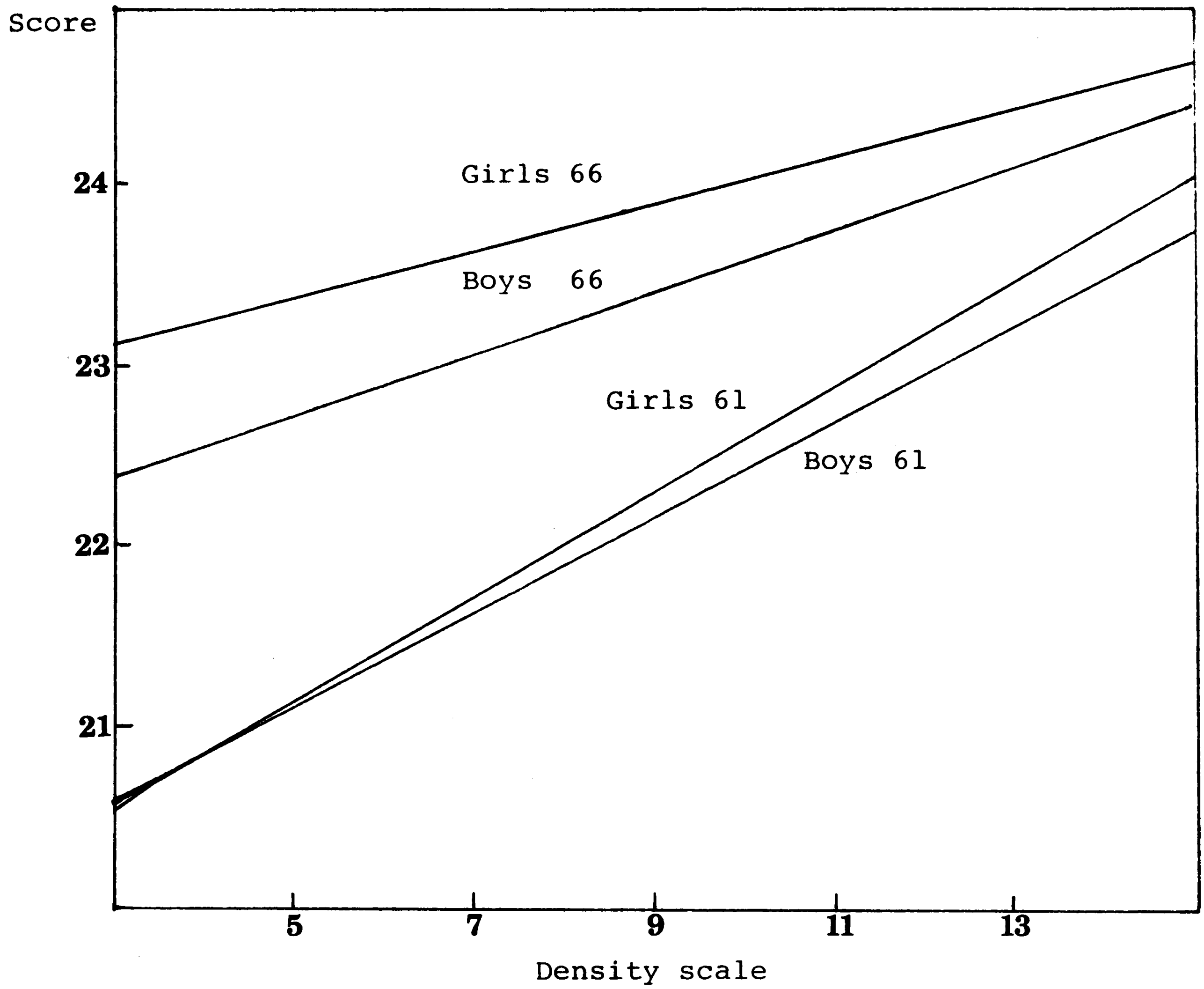

Figure 2. Opposites - Individual test score regressions on density. Significant differences:

Means	Slopes
61<66	61>66 M
M<F	61>66 F

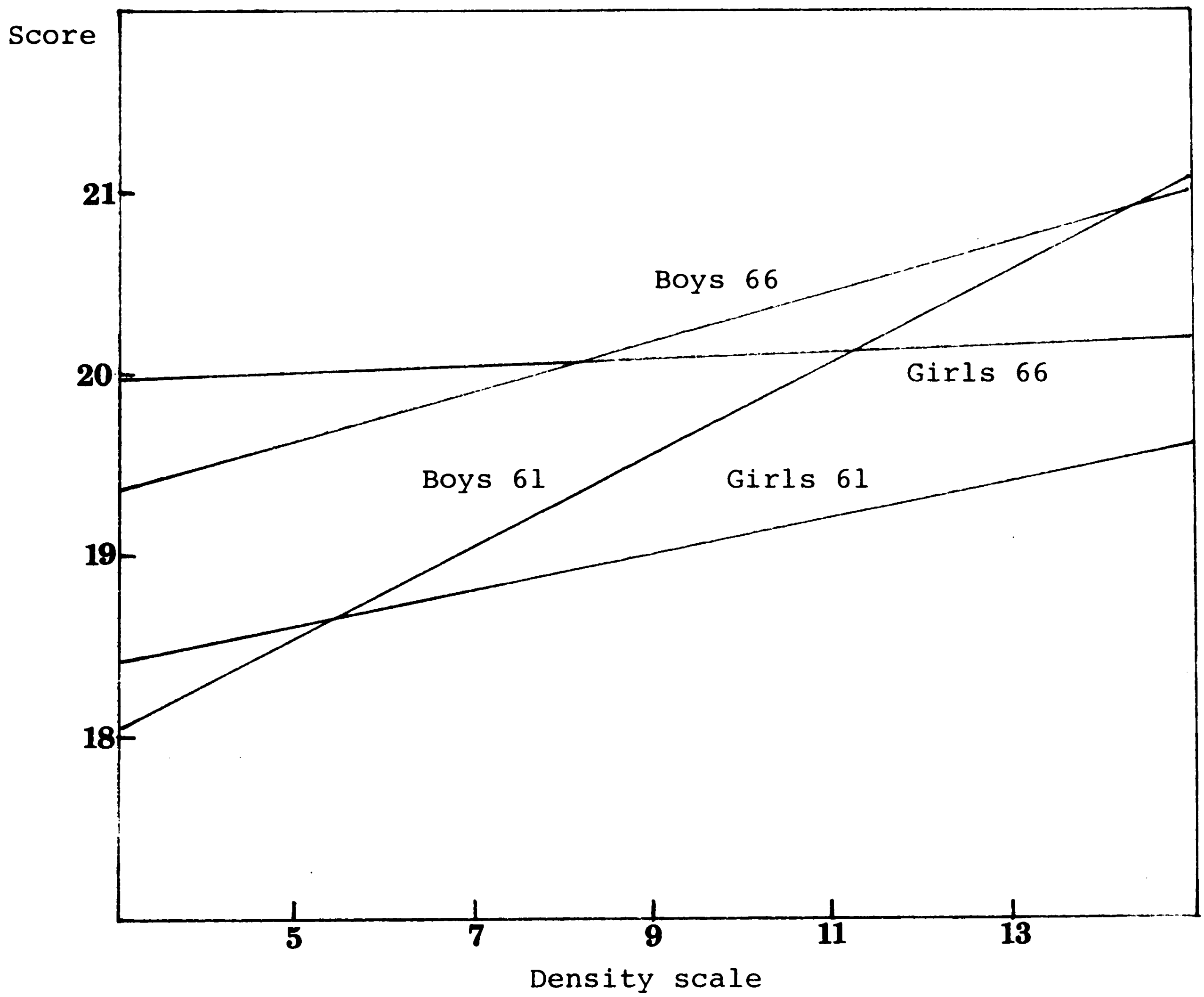

Figure 3. Number Series - Individual test score regressions on density. Significant differences:

Means	Slopes	
61<66	61>66	M
M>F	M>F	61
	M>F	66

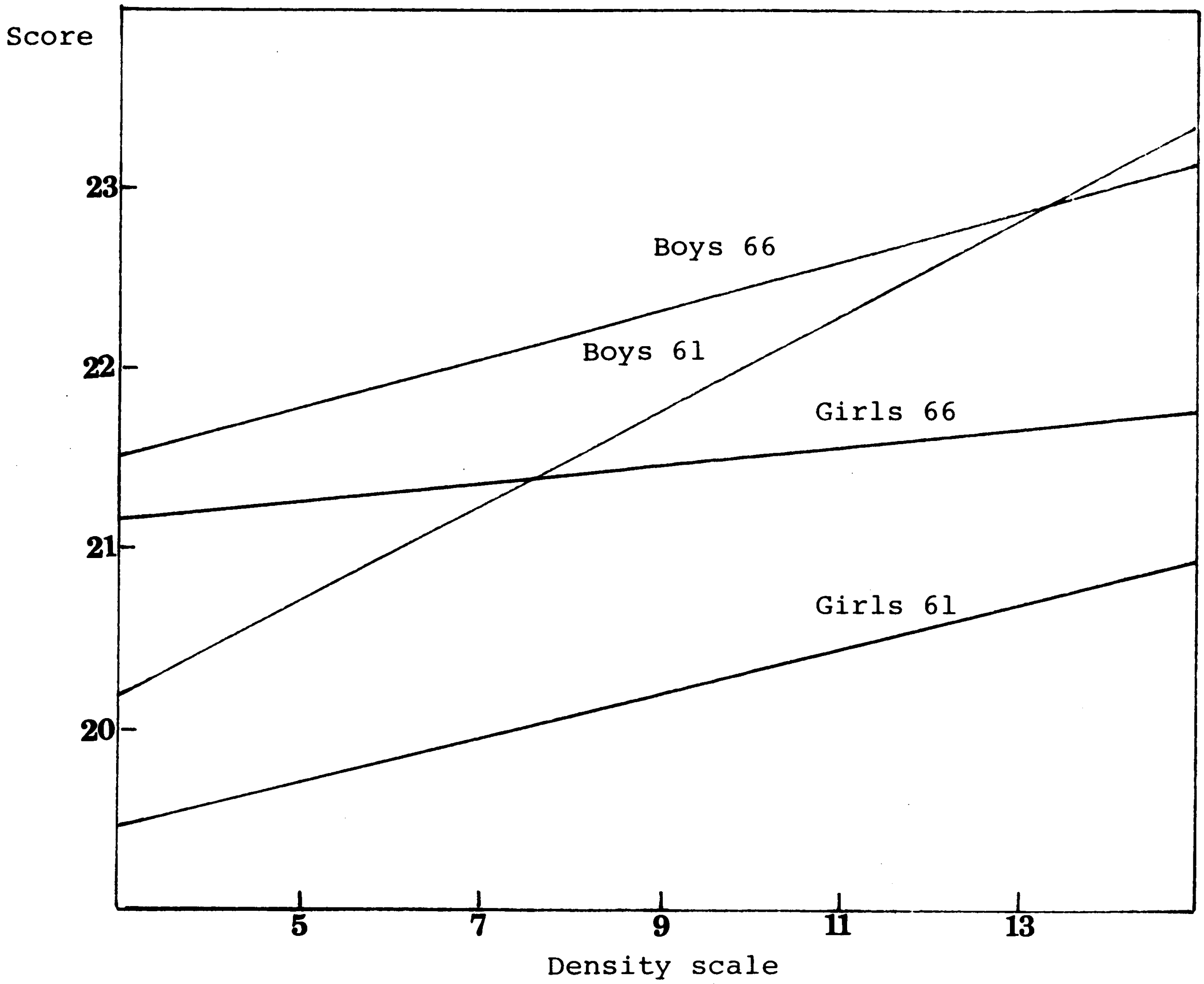

Figure 4. Metal Folding - Individual test score regressions on density. Significant differences:

Means	Slopes
61<66	61>66 M
M>F	M>F 61
Interaction	

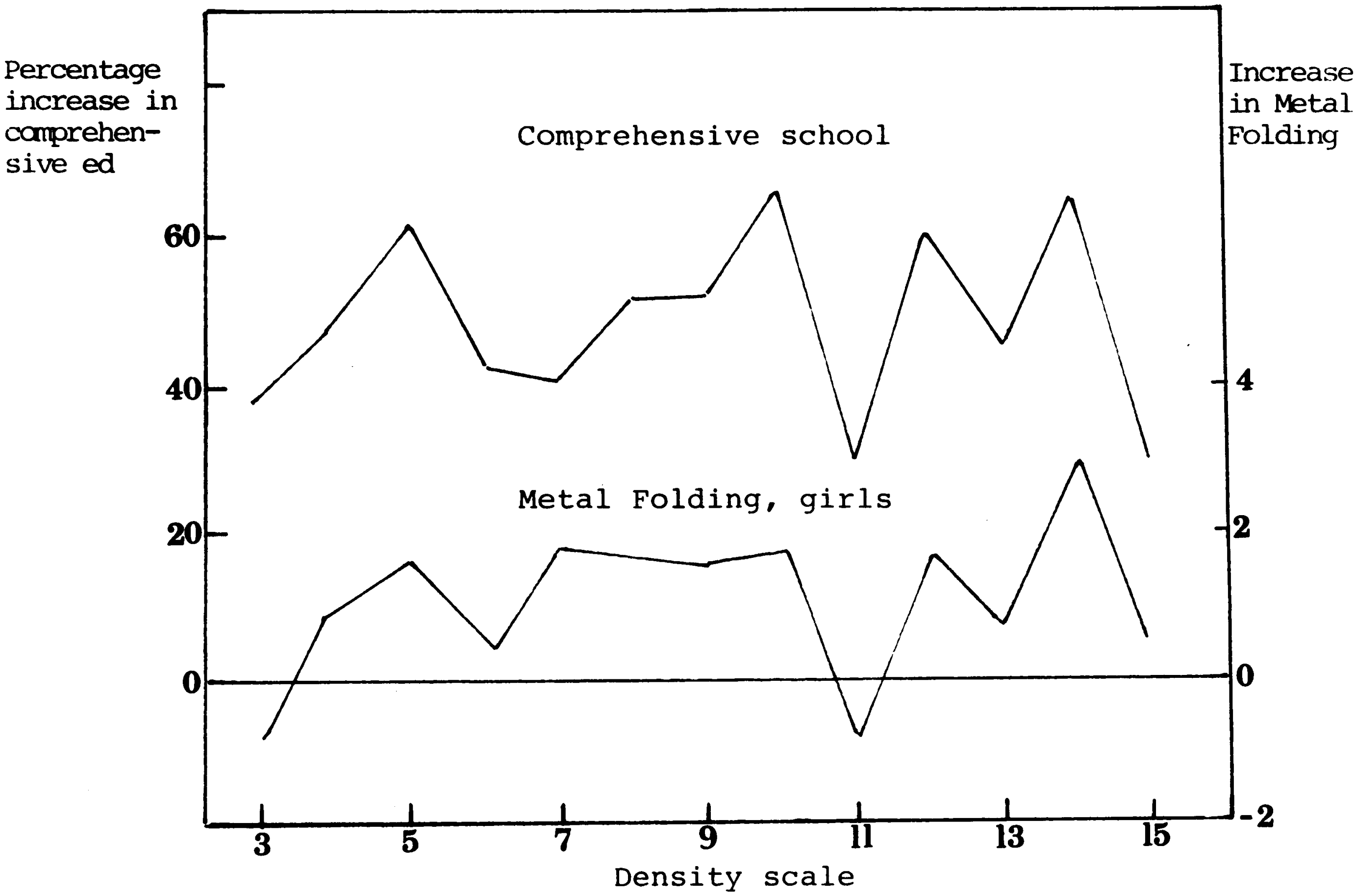

Figure 5. Relations between increases in spatial test scores and comprehensive school coverage, with density controlled. (Individual pupil as unit of analysis)

Appendix 1. Ecological variables derived from project data.

	Initial	Change
Secondary education among fathers	x	x
Secondary education among mothers	x	x
Commercial occupations among fathers	x	x
Salaried employees, higher level, among fathers	x	x
Salaried employees, lower level, among fathers	x	x
Occupational activity, among mothers	x	x
Salaried employees, lower level, among mothers	x	x
Workers among mothers	x	x
Socioeconomic group A (professional and managerial)	x	x
B (white-collar, higher level)	x	x
C (white-collar, lower level)	x	x
D (agricultural)	x	x
E (workers)	x	x
Comprehensive school coverage		x

Appendix 2. Ecological variables derived from official statistics.

	Initial	Change
Population increase 1960 to 1961	x	
Taxable income	x	
TV licenses	x	x
Housing units built in 1951-60	x	
Housing units with bath or shower	x	
Men in white-collar occupations	x	
Men in agricultural occupations	x	
Men in non-agricultural working-class occupations	x	
Women in white-collar occupations	x	
Women in agricultural occupations	x	
Women in non-agricultural working-class occupations	x	
Occupationally active women	x	

46. Gustafsson,Jan-Eric: Inconsistencies in aptitude-treatment interactions as a function of procedures in the studies and methods of analysis. Project MID 16, March 1976.

47. Gustafsson,Jan-Eric: Differential effects of imagery instructions on pupils with different abilities. Project MID 17, April 1976.

48. Ekholm, Mats: Social development in school. Summary and excerpts. Project SOS 23, May 1976.

49. Stangvik, Gunnar: Approaches to the analysis of learner-task interactions and some implications for the study of pedagogical processes. Project YP 7, January 1976.

50. Andrae, Annika: Non-graded instruction in small rural lower secondary schools. A presentation of the PANG-project. Paper read at the INTERSKOLA conference, July 1976, Project PANG 20, July 1976.

51. Patriksson, Göran: Attitudes toward olympic games of Swedish adolescents. Paper presented at the internationa congress of physical activity sciences in Quebec City 11-16 July 1976. September 1976.

52. Gustafsson, Jan-Eric: A note on the importance of studying class effects in aptitude-treatment interactions. Project MID 19, September 1976.

53. Gustafsson, Jan-Eric: Spatial ability and the suppression of visualization by reading. Project MID 20, September 1976.

54. Dahlgren,Lars Owe & Marton, Ference: Investigation into the learning and reaching of basic concepts in economics - a research project on higher education. Paper presented at the Congress of the European Association for Research and Development in Higher Education. August 30 - September 3, 1976, Laouvain la Neuve, Belgium, September 1976.

55. Marton, Ference: Skill as an aspect of knowledge. Some implications from research on students conceptions of central phenomena in their subjects. Paper presented at the Second International Conference on Improving University Teaching, July 13-16, 1976. Heidelberg, F.R. Germany, September 1976.

56. Andrae. Annika (Ed.): Non-graded instruction. Research organization and design. Administration and daily teaching experiences in small rural lower secondary schools. Experiences from the PANG-project. Paper read a the INTERSKOLA conference in Sveg, July 1976. Project PANG 23, August 1976.

57. Sandgren, Björn & Åsberg, Rodney: On cognition and socia change. A report from a pilot study regarding the effect of schooling on cognitive growth and attitudes towards social change in Pakistan, October 1976.

58. Sandgren, Björn: Relation between cognition and social development. January 1977.

59. Entwistle, Noel: Changing approaches to research into personality and learning. February 1977.

60. Härnqvist, Kjell: Primary Mental Abilities at Collective and Individual Level. October 1977.

61. Härnqvist, Kjell: Enduring Effects of Schooling - A Neglected Area in Educational Research, October 1977.

62. Härnqvist, Kjell: A Note on the Correlations between Increments, Cumulated Attainment and a Predictor. October 1977.

63. Gustafsson, Jan-Eric: The Rasch Model for Dichotomous Items: Theory, Applications and a Computer Program. December 1977.

64. Härnqvist, Kjell: An Ecological Analysis of Test Score Changes over Time. December 1977.