RR – 84-05

Correlation between Raven Progressive Matrices test scores of fathers and sons

J. de Leeuw
P.A. Vroon
A.C. Meester
J.L. Leyten

Vakgroep Datatheorie FSW/RUL
Vakgroep Theoretische Psychologie RUU

ABSTRACT

The paper studies parent off-spring resemblance in intelligence scores.

None of the assumptions underlying the measurement of intelligence in general and parent-child resemblance in particular is ever met in actual investigations and most (if not all) of the studies sofar have shortcomings with respect to sampling and representativity. The present study overcomes at least some of the more serious handicaps of older ones. A large sample of father-son pairs tested at the same age and with the same test is analysed. The sample is representative for the male population in the Netherlands.

Regression of fathers' scores on sons' and of sons' on fathers' is non-linear and therefore the correlation coefficient is a misleading measure of association, if computed in a direct way. First a polychoric correlation coefficient is computed on the basis of a grouped binormal model. In a second step this coefficient is corrected for restriction of range.

The grouped bivariate normal model fits the data very well and the association can be described by a correlation coefficient of .34 with standard error .018.

Further investigation of the data is needed, considering possible systematic variations in the correlation if professional and educational level of the parents are also included in the analysis.

Keywords: parent-child resemblance in intelligence, polychoric correlation

Correlation between Raven Progressive Matrices-test scores of fathers and sons

J. de Leeuw[1)], P.A. Vroon[2)], A.C. Meester[1)], J.L. Leyten.

1. Introduction

In this paper we study parent-offspring resemblance in intelligence. As McAskie and Clarke (1976) have pointed out in an important review paper, this resemblance can be studied in several ways. The studies they review show many variations. This makes it difficult to summarize and to compare them. We add another study and it does not seem to have many of the disadvantages of older studies. Because of our material we can eliminate one important source of variation right away. We restrict our attention to the resemblances between fathers and sons.

If we say that we want to study this relationship we first need a definition of intelligence. In view of most of the previous research in this field we first try to use a psychometric or operational definition based on the classical Spearman model. To that effect we must assume that there is a universe of intelligence tests or intelligence test items. All these tests have a single common factor, by which we mean that the residuals after projection on a hypothetical common factor are orthogonal. Scores of the individuals on this common factor can be determined uniquely, and they are identified with the intelligence of that individual. In practice, of course, we have only a sample of tests or items and the persons' intelligence is to some extent undetermined.

1) Dept. of Data Theory, University of Leiden
2) Dept. of Theoretical Psychology, University of Utrecht.

The authors express their thanks to the Dept. of Behavioral Sciences of the Dutch Ministry of Defense (Mr. M. Harsveld) who did everything possible to help in carrying out this study. Mr. H. van Ree assisted in tracing the data in the files of the army. Mr. G. van den Andel assisted in coding and preparing the data for analysis.

We do not assume that an individual's level of intelligence is stable throughout his or her lifetime. Thus our previous discussion only defines intelligence at time t or age t, and intelligence is a real valued function on the time-axis. In the population of individuals we are studying, we have a real valued stochastic process $\chi(t)$, which is called "father's intelligence", and a real valued process $\eta(t)$, "son's intelligence". It is clear that this definition gets us into trouble right away. Intelligence is defined as the hypothetical common factor of an, in principle infinite, battery of intelligence tests. But the structure of tests which can be described by the Spearman model at time t, may very well be more complicated at time $s \neq t$. Moreover, it is clear that some intelligence tests cannot be given to very young children, and that it does not make sense to give tests for young children to adults either. The fact that we assume a universe of test items, which can be used to measure intelligence at each and every age, already presupposes that intelligence is at least conceptually the same trait throughout the life-span. And this supposes that intelligence is a priori given, i.e. genetically defined at conception. More particularly, our definition supposes that the intelligence of fathers and sons are comparable. They are the same trait, measurable by using the same universe of test items. The factor structure of the tests is not supposed to change during the period we study (although of course means and variances of individual tests may change).

After spelling out the assumptions that are needed to make the psychometric approach to father-son resemblance in intelligence operative, it becomes clear that none of these assumptions is ever met in actual investigations. It is possible to engage oneself in deep discussions if the psychometric research program is at all feasible, but because the requirements are so clearly violated in everyday psychometric practice, such a discussion is of academic interest only. We mention some of the 'short-comings' of actual studies. Longitudinal studies on the development of intelligence are carried out on small samples, usually heavily selected and far from representative. Often a great variety of tests is used. Father-son resemblance is investigated by testing fathers once and and sons once. If fathers are tested as adults, and sons as children, they often get different tests. If fathers and sons are tested at the same age, different tests are also used in many cases, because of the 20-30 year time lag between the testings. Because only a single test is used, with a problematic factor structure for its items, measurement of intelligence at a given time will allready be poor according to psychometric standards. Approximation to the latent processes $\chi(t)$ and $\eta(t)$ is even poorer, because we only use one time point and only one indicator for each process.

The program becomes much more realistic if we study the actual performance of fathers and sons on a given battery of tests. This means, of course that we in fact abandon the attempt to measure intelligence. The intelligence concept is used mainly for guiding our interpretations. We have a well-defined concrete task, with observable outcome, given to both fathers and sons. The latent processes $\chi(t)$ and $\eta(t)$ have vanished, the observed processes $x(t)$ and $y(t)$ are no longer interpreted as indicators, but are studied in their own right. But even if interpreted in terms of observed processes $x(t)$ and $y(t)$ the existing parent-offspring resemblance studies are not very satisfactory. Studies with different tests for fathers and sons cannot be used any more. Using different tests can only be justified by supposing that they are all measures of the same latent process, and we have decided not to use this latent process any more. Of course it is still possible to compare, for instance, father's Raven at time t with son's Binet at time s. But although this correlation may be of some interest, there is no theory which tells us that it is more interesting or more fundamental to our problem than the correlation between father's reaction time and son's skull circumference at times s and t.

Because we do not have a satisfactory testable theory of intelligence measurement, it is not advisable to compare measurements on different tests. Neither is it advisable to compare the performance of fathers and sons at different ages. As long as there is no satisfactory theory of test performance development, the comparison of sons of age 20 with fathers of age 45 can be misleading. It is even more misleading to compare a sample of sons, with average age 20, with their fathers tested at the average age 45. We agree with McAskie and Clarke in this respect. 'Even in this narrow area, there are a number of ways of looking at the data. Firstly, the notion of resemblances can be interpreted as meaning either resemblance in actual performance or resemblance in relative performance with respect to an age group. For intelligence, the former requires a direct comparison between the performance of parents and offspring at the same age on the same test. Most studies use the latter method, the comparison being based on measures derived from performances on different sets of items at different ages. Unfortunately, psychological theories underlying test construction are inadequate to treat resemblance in type as meaningful'. (l.c., p. 243). These considerations provide the major motivation for our paper. We analyze a large sample of father-son pairs, tested at the same age and with the same test. The sample is representative for the male population of the Netherlands, except for some fairly unimportant selection factors which are mentioned below. We compare the actual performance of fathers and sons, and investigate if the resemblance between the two can be quantified appropriately by using product moment correlation methods. We investigate this appropriateness in considerably more detail than is usual, by applying newly developed statistical methodology.

2 REVIEW OF LITERATURE

The most famous summary of early studies on parent-offspring resemblance in intelligence is, of course, the short paper by Erlenmeyer-Kimling and Jarvik (1963). It has been used by many authors, and the corresponding plot has been reproduced in many books and papers. Nevertheless the studies summarized there are almost useless for our purposes. Samples are often small and rarely representative. Parents and offspring were tested at different ages, and with different tests. The resulting correlation estimate, the median of twelve very heterogeneous studies, was .50. Jencks (1973) recomputed the estimate .48 on the basis of a similar survey. These summaries, and the studies on which they are based, have been criticized severely by Kamin (1974), Taylor (1980), McAskie and Clarke (1976). We discuss the studies which may seem relevant briefly below.

Higgins et al. (1962) collected data on 1016 mothers, 1016 fathers, and 2039 children. They were selected from a large pool of relatives of 300 patients of the Minnesota State School and Hospital for the mentally retarded. IQ's were collected for as many persons as possible, using the testing programs in the various school systems. If IQ-values of both parents and of at least one child were known, then the family was included. The father child correlation found was .43 ± .02, the mother-child correlation .45 ± .02. Mean age of parents at testing was 14.24 years, mean age of children was 8.65 years. The tests used were heterogeneous, the age at testing was heterogeneous too.

Waller (1971) selected 131 fathers and their 173 sons from the same population. The criterion was that the sons had to be 24 years or older, and that scores of both father and son were available. Mean age at testing was 15.90 ± .52 for fathers and 13.38 ± .20 for sons. Again the tests were the various (Kuhlmann and Otis) group tests used in the school system. IQ-correlation was .360. Of course in some cases the IQ's will be derived from different tests for father and son, or from the same test with a different norming for father and son. The mean ages of fathers and sons match pretty well, but the situation with respect to the tests is rather hopeless.

One of the very few attempts to correlate x(t) and y(t) at various ages is the study by McCall (1970). He used a sample of 35 parent-child pairs from the Fells longitudinal study. Most of the tests were Binet tests, but they were standarized for each age group separately. At ages 42, 48, 54, 60, 66, 74, 84, 96, 108, 120 and 132 months the parent-child correlations were .36, .29, .27, .19, .43, .45, .50, .35, .28, .21, .17. Tests are fairly homogeneous in this case, matching for age is perfect, but the sample is very small indeed.

According to the review of McAskie and Clarke (1976) these three studies seem to be the only ones in which parents and children from a fairly representative sample were tested at the same age. Because of the particular test we use in our study (a version of the Raven Progressive Matrices) three other studies are of some importance. Guttman (1974) compared 89 Israeli father-son pairs using the Progressive Matrices, and found a correlation of .36. De Fries et al. (1976) studied 672 father-son pairs of European ancestry and 241 father-son pairs of Japanese ancestry on Hawaii. The progressive

Matrices correlations were .23 and .09. Park et al. (1978) repeated a similar study in Korea. There were 100 father-son pairs, and the Progressive Matrices correlation was .33.

A general tendency in the values of parent-offspring correlation now becomes clear. Compare the reviews of Plomin and De Fries (1980), and of Scarr and Carter-Saltzman (1982). Erlenmeyer-Kimling and Jarvik still estimated parent-offspring correlation around .50. This figure was also used by Burt, Jensen, Herrnstein, Eysenck and Jencks. More recent data suggests a correlation around .35 in large representative studies. Of course the precise value depends on the test, the sample, the age, the country, and other factors, but the general tendency is that in older studies the correlation has been seriously overestimated.

3 DATA COLLECTION

In the Netherlands military service is compulsory. The Ministry of Defence calls up all males for a medical and psychological examination in the year they reach the age of 18. The summons are sent out by a national centre and the actual examination takes place in various places all over the country. In the fall of 1981, the national centre added 10.000 questionnaires to the standard forms. Since each day about 500 males receive this order, after about 20 days all questionnaires were sent out. The respondents were identifiable by their registrationnumbers, but anonimity was guaranteed. Furthermore both son and father had to agree to be used in the sample. The questionnaires could be returned to the university in a postage-paid envelope. We asked the draftees to write down their registration number which contains the year, the month and the day of birth and three other digits, their most recent educational certificates, their current education and the city and the date on which the examination would take place. Furthermore, we asked if their father was still alive, if he was a part of the family, his profession and registration number or date of birth. During the spring of 1982 another 10.000 questionnaires were sent out. They were identical to the first set, apart from the fact that we also asked if the mother was alive, and if so, her educational level and school certificates.

We received about 5.000 questionnaires back. About 10% proved to be useless since the registration number, the registration number of the father, and/or the date of birth of the father was lacking, or because of the fact that the father did for some reason or other no longer belong to the family. The data were ordered on date of birth and put on a computer file. The Ministry of Defense collects all raw data of the examinations that have taken place. After coding and filing, the forms with the raw data are destroyed within about half a year.

Since 1945, the Ministry of Defence uses several intelligence tests in the field of technical insight, perceptual speed, etc. Most of these tests have not changed much over the years. Specifically, one test has remained unchanged since 1945, the Raven Progressive Matrices. This test contains 60 items. In 1945, an item-test analysis was carried out. On the basis of this, the 40 best discriminating items were chosen. Consequently, everybody who has been tested since 1945 has a maximal Raven score of 40.

On the basis of the computer file and the examination forms the raw Raven score of the son was looked up, and subsequently the raw score of his father. Fathers who had been tested before 1945, were excluded from the data. As far as the fathers are concerned, we also recorded their educational level and school certificates at the moment they were tested. As a final check, all surnames of the fathers and the sons were compared on the basis of the registration numbers. In this way, virtually all possibly wrong father-son pairs could be detected. Due to incorrectly filled in forms, missing data and doubtful cases, 2847 father-son pairs were left for analysis.

A small percentage of the population is not called up for the examinations. This is the case if two brothers completed military service, if the draftee is not resident in the Netherlands, and in case of general disablement. Another small percentage is medically examined, but not psychologically tested because of serious medical defects or unability to speak Dutch.

In this paper we only use part of the data we have collected. The raw Raven-scores of fathers and sons are used to estimate the correlation coefficient in the bivariate distribution. The data having to do with socioeconomic status and educational levels will be analyzed in subsequent studies.

4 RESULTS AND ANALYSIS

The basic results are given in tables 1 and 2. Table 1 gives the complete bivariate distribution in 41 x 41 = 1681 cells, with sons as rows and fathers as columns. Table 2 shows various sets of univariate marginals. In the first column we have the father-marginal, which are the column-sums of table 1, and in the second column the son-marginal, the row sums of table 1. The third and fourth column contain the population distribution (all persons tested) at two time points corresponding with our investigation. These last two columns are used in section 4.2 to investigate selection bias.

INSERT TABLES 1 AND 2 ABOUT HERE

4.1 Raw univariate marginals

To compare the test score distribution of fathers and sons we have plotted the cumulative distributions in figure 1. It is clear that there has been a considerable shift to the high end of the scale in the course of one generation. Although a test seems superfluous we perform one for completeness. Thus we test the hypothesis that the distributions of fathers and sons are the same (i.e., we test marginal homogeneity of the bivariate distribution). The chi-square value is 4036.7853, with 40 degrees of freedom.

INSERT FIGURE 1 ABOUT HERE

We can also compare both distributions by computing moments. This is done in table 3, in which we give mean, variance, skewness, and excess. (The skewness is the square root of the third-order moment about the mean divided by the third power of the standard deviation. The excess is the fourth moment about the mean, which is first divided by the square of the variance, and then diminished by three. Both skewness and excess are invariant under linear transformations. For symmetric distributions skewness is zero, for the normal distribution excess is also equal to zero.) Table 3 indicates that for the fathers we have a distribution which is more or less symmetric around the midpoint of the scale. The excess also indicates a shape like a normal distribution. For the sons, however, the situation is completely different. The distribution is very skew as well as peaked. There is a considerable loss of variance, and the mean has shifted almost eight points. For the sons the matrix test is too easy. This dramatic 'increase in intelligence' has been studied in more detail by Dronkers (1978) and by Meester and the Leeuw (1984), using similar test results, but more time-points.

INSERT TABLE 3 ABOUT HERE

We use skewness and excess merely as descriptive statistics here, and not as a test for normality. Such a test can be derived from figure 2, which applies the inverse cumulative normal to the empirical distributions in figure 1. If the distributions are discreticized normals, with equally spaced discretization points, then the curves in figure 2 should be straight lines. We test the hypothesis that this is the case by a simple chi square test. For fathers chi square is 195.0600, with 38 degrees of freedom. For sons it is 676.7622, with 38 degrees of freedom as well. Although the fit is much better for fathers than for sons, it is definitely unsatisfactory for both. Consequently, the marginals are not equally spaced discrete normal.

INSERT FIGURE 2 ABOUT HERE

4.2 Selection bias

We have already seen that is is important to find out if the sons in our sample can be considered to be a random sample from the population of all sons tested in the same period. We have compared our son-distribution with the population of all individuals tested between September 15 and December 11, 1981. These are 28.769 individuals. We have also compared it with the population of all individuals tested in September, October and November 1982, another 29.128 individuals. The three cumulative distributions are plotted in figure 3. The two population distributions cannot be distinguished in the plot, they are almost identical. The lower curve is the one of our sample.

INSERT FIGURE 3 ABOUT HERE

Although the deviations are not large, they show clearly that lower test scores are under-represented in the sample. A partial explanation might be that many persons with very low scores were not motivated to fill in the questionnaire. Figure 3 shows that an even larger portion of the tail is censored away by our sampling procedure. But all in all the distortion is rather small. We have tested the deviations in figures 3 for significance. For the 1981 population the chi square is 69.4050, with 40 degrees of freedom. For the 1982 population it is 69.6032, again with 40 degrees of freedom. Although both chi squares are significant, they are certainly not large.

4.3 Raw bivariate distribution

We now study the complete bivariate distribution given in table 1. First we compute the regression functions. In figure 4 we have plotted the average score of the son as a function of the score of the father. These averages are labeled with S in the figure. We have also computed the 95% confidence intervals for each of these averages. They are indicated as lines in the figure. We see that the regression of sons' scores on fathers' increases only slightly. The instability is greatest near both endpoints of the scale, because fathers with very high scores and fathers with very low scores are about equally rare.

INSERT FIGURE 4 ABOUT HERE

The situation is quite different for the regression of fathers on sons shown in figure 5. The instability is concentrated on the lower part of the scale, and above the midpoint of the scale the regression function is much steeper than the one in figure 4.

INSERT FIGURE 5 ABOUT HERE

We tested the linearity of these regressions with the appropriate chi square test. For the regression of sons on fathers we find a chi square of 123.3234 with 37 degrees of freedom, and for the regression of fathers on sons a chi square of 52.2824 with 32 degrees of freedom. Linearity of regression must be rejected for both curves, although only marginally for fathers on sons.

Due to the nonlinearity of the regressions the correlation coefficient is a doubtful measure of association. Because it is traditionally the statistic that is considered to be the most interesting one we have computed it any way. It equals 0.3003. Our subsequent computations have the purpose of finding out if and to what extent this estimate makes sense.

4.4 Grouped bivariate distributions

A more refined statistical analysis is only possible if we group the test scores into a relatively small number of classes. We can only fit more specific parametric models on such grouped tables, in which cell frequencies are considerable. How to choose the grouping is to a large extent arbitrary. We have followed the procedures used by the Ministry of Defence, and grouped into six classes. The first class and the last class must contain approximately 10% of the distribution, the four middle classes contain 20% each. The grouped distribution with class intervals and marginals is given in table 4. If we group the two population distributions in the same way as the distribution of the sons in the sample, we can again compute chi squares to measure the extent in which selection has taken place. The first chi square is 36.2759, the second one is 47.0946. Both have 5 degrees of freedom. The residuals indicate, again, that the lowest classes are under-represented in our sample.

INSERT TABLE 4 ABOUT HERE

4.5 Computing the polychoric correlation

We have seen in 4.1 that the marginals of the raw distribution are not univariate normals, and in 4.3 that the regressions are not linear. This definitely implies that the raw bivariate distribution in table 1 is not bivariate normal. But it can be grouped binormal, or, to use a term of Udny Yule, strained binormal. By this we mean that there are bivariate normal latent variables χ and η, which we might call the true test scores of fathers and sons. They vary continuously. They are 'rounded' to the observed test scores x and y, but not rounded by using equal intervals. If $\chi \leqq \chi_0$ then the son has score 0, if $\chi_0 < \chi \leqq \chi_1$ the son gets score 1, and so on. Thus there are 40 parameters corresponding with category boundaries for sons, and in a similar way there are 40 parameters for the fathers. The usual bivariate normal model has 5 parameters: two for the means, two for the variances, and one for the correlation. A grouped binormal model for table 1 has 81 parameters: 80 for boundaries and one for the correlation. Correlations estimated on the basis of the grouped binormal model are called polychoric correlations.

It is not possible to estimate a polychoric correlation coefficient directly from table 1. It has too many cells, more specifically too many empty cells. Therefore we need grouping. Observe that the grouped multinormal model is invariant under grouping. After the grouping has taken place, it is still true that a grouped binormal model is true, with the same parameters as before grouping (only some parameters have disappeared). Thus if the grouped binormal model is true, we need not be concerned about the fact that the grouping is arbitrary.

To compute polychoric correlation we use a technique described by Van der Pol and De Leeuw (1984) on table 4. In this table the grouped binormal model uses 5 + 5 + 1 = 11 parameters to describe 6 x 6 - 1 = 35 independent proportions. Consequently the chi square indicating the fit of the model, which is 31.0036 for these data, has 24 degrees of freedom. This is a very good fit, and on the basis of this analysis we cannot reject the grouped binormal model. The polychoric correlation is 0.3186, and its standard error is .0180. This is clearly a better estimate of the correlation coefficient than our earlier one, because it computes essentially the same parameter on the basis of a more appropriate model. In fact the good fit of the grouped binormal model indicates that summarizing the associations by the use of a correlation coefficient makes sense.

4.6 Correcting the polychoric correlation

Using polychoric correlation can be thought of as applying a correction for grouping, without assuming that the groupings are equally spaced. We can also apply a correction for selection (for restriction of range) in the polychoric framework. Of course we must make assumptions about the mechanism that generated the non-response. Basically, we assume that the non-response mechanism is ignorable, in the technical sense in which this word is used by Rubin (1976) and Little (1982). This does not mean that we suppose that responding or non-responding is independent of the test-score of father and son. In fact we merely assume that responding or non-responding is independent of the score of the father, given the score of the son. To put it differently: non-response does not depend directly on the score of the father, only on that of the son. Or: the conditional distribution of the scores of the fathers, given the scores of the sons, is the same in the responding and in the non-responding groups. With the available data it is not possible to test this additional assumption. It seems plausible, however. If we make the ignorability assumption, we can correct the bivariate distribution by using the two population marginals in table 2. If we apply our polychoric algorithm to these corrected tables, we find estimates of .3391 and .3393, and chi squares of 32.4289 and 32.2505 with 40 degrees of freedom. The polychoric model continues to fit well. Thus our final conclusion is that the bivariate normal model (in its grouped form) fits our data nicely. The association can be described by a correlation coefficient. Our best estimate of this correlation coefficient is .339 with standard error .018.

5 CONCLUSIONS

The present study has several advantages over earlier ones. We have a large sample, and father and son are tested at the same age with the same test. Our computation of the correlation coefficient is also much more careful than usual, with the net result that the rough estimate .30 is replaced by the better estimate .34. In more or less comparable studies which (much) smaller samples Guttman (1974) and Park et al. (1978) also found correlations around .34.

As stated earlier, the Raven test is just one of the arbitrary ways to measure intelligence in the psychometric tradition. The Raven correlates about .75 with the WAIS. Although the matrix test requires no verbal fluency, it relates no more closely to so called performance IQ than verbal IQ, both correlations are about .70. Furthermore, the matrix test is relatively independent of specific educational abilities as compared with many other tests. For these reasons this test is used in The Netherlands and also in Great Britain for military classification. It is important that normally intelligent recruits are not rejected because of poor school education. The rather dramatic increase of the mean scores on this test over the past decades has also been observed elsewhere, but it remains largely unexplained. Possibly, a reason is that the general level of education has also increased considerably in the western countries.

It is difficult to interprete the correlation we have found in the framework of the IQ-debate. One of the main problems with this debate is that the so called hereditarians and the environmentalists attach value to different data. The environmentalists do not predict correlations and are mainly interested in intelligence levels. A further analysis of the data is needed in order to discuss this position, i.e., correlating the Raven score with the profession of the father, his education, the educational level of the son, the school certificates of the mother etc. The hereditarians assume that parent and child have 50% of their genes in common, which is reflected in a predicted average parent-offspring correlation of .50. In case of assortative mating this may easily increase to .60. It is clear that this value is a gross overestimate of the observed correlation. Subsequent analyses of the data will show whether or not this correlation varies systematically when performance of the son is correlated with the profession and the educational level of his parents.

6 REFERENCES

DeFries, J.C., Johnson, R.C., Kuse, A.R., McClearn, G.E., Polvina, J., Van den Berg, S.G. & Wilson, J.R. (1979), Familial resemblance for specific cognitive abilities, Behaviour Genetics, 9, 23-43.

Dronkers, J. (1978), De stijging van intelligentiescores. Hollands Maandblad, 19, no. 363, 15-19.

Erlenmeyer-Kimling, L. & Jarvik, L.F. (1963), Genetics and intelligence: a review. Science, 142, 1477-1479.

Guttman, R. (1974), Genetic analysis of analytical spatial ability: Raven's Progressive Matrices. Behavior Genetics, 4, 273-284.

Higgins, J.V., Reed, E.V. & Reed, S.C. (1962), Intelligence and family size: a paradox resolved. Eugenics Quarterly, 9, 84-90.

Jencks, C. (1972), Inequality: a reassessment of the effect of family and schooling in America. New York, Basic Books.

Kamin, L.J. (1974), The Science and politics of IQ. Potomac, Md, Earlbaum

Little, R.J.A. (1982), Models for nonresponse in sample surveys. Journal of the American Statistical Association, 77, 237-250.

McAskie, M. & Clarke (1976), A.M., Parent-offspring resemblances in intelligence: Theories and evidence. British Journal of Psychology, 67, 243-273.

McCall, R.B. (1970), Intelligence quotient pattern over age: comparisons among siblings and parent-child pairs. Science, 170, 644-648.

Meester, A.C. & De Leeuw, J. (1984), Over het intelligentieonderzoek bij de militaire keuringen vanaf 1925 tot heden. Mens en Maatschappij, 59, 5-26.

Park, J., Johnson, R.C., DeFries, J.C., McClearn, G.E., Mi, M.P., Rashad, M.N., VandenBerg, S.G. & Wilson, J.R. (1978), Parent-offspring resemblance for specific cognitive abilities in Korea. Behavior Genetics, 8, 43-52.

Plomin, R. & DeFries, J.C. (1980), Genetics and intelligence: recent data. Intelligence, 4, 15-24.

Rubin, D.B. (1976), Inference and missing data. Biometrika, 63, 581-592.

Scarr, S. & Carter-Saltzmann, L. (1982), Genetics and intelligence. In: Sternberg, R.J. (Ed.), Handbook of human intelligence. London, Cambridge University Press.

Taylor, H.J. (1980), _The IQ-game: a methodological inquire into the heredity-environment controversy_. New Brunswick, N.J., Rutgers University Press.

Van der Pol, J. & De Leeuw, J. (1984), Aspects of the estimation of the polychoric correlation coefficient. Submitted for publication.

Waller, J.H. (1971), Achievement and social mobility: relationships among IQ-score, education, and occupation in two generations. _Social Biology_, 18, 252-259.

table 1: raw bivariate distribution for sons' (rows) and fathers' scores on Ravens Progressive Matrices (columns)

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
0																																									
1																																									
2																																									
3																		1								1															
4												1					2																								
5													1						1																						
6																		1																							
7																	1																								
8															1								1																		
9																																									
10																2																									
11											1				1																										
12					1						1	1							1					1																	
13									1			1										1	1					1													
14				1		1			1	1					2		1			1																					
15									1	1		1					2								2																
16							1									1			1	1				1						2											
17										1		1		1			1					1	1	1	1			1													
18							1			1	3	2	1			1	1	2		1		1		1		2		1													
19						2	2	2	1	1	1	1	2	1		2	2	1		1	2				1	1	2	2	3												
20				2				2	1	3	1	1	1	1		3	2	5		2	1	2		3	1				1	1				1							
21		1				1			2		2	1	1	1			1	2	3	2	2	2	1	1	1	1		1					1								
22			1	1		1				2	2	2	2		1	4	2	3	1	5	3	6	1	5	5	1	2	2	3	3			1								
23			1			1		2	1	1	2	2	3	5	3	1	3	3	3	6	2	1	7	2	3		3	3	3	4		1									
24			1	1	1	1			2	3		3	2	1	4	1	7	2	8	3	10	7	5	5	8	6	6	4	6	3	2	1	1		1				1		
25			1		1	2	1	1			3	2	3	5	8	4	5	6	3	4	13	8	6	5	8	12	5	6	3	2			1		1	1		1			
26		1			1	2	2	6	1	4		6	5	5	7	8	7	3	12	5	9	5	11	9	13	7	9	2	9	3	1	3	1	4	1						
27		1			1	1	4		2	2	3	2	4	6	4	8	12	12	20	6	7	13	11	12	5	11	10	11	3	12	2	1	2	1	2	2	2		1		
28				1	2	2	2	1		2	2	7	3	11	6	7	4	12	7	7	16	10	8	15	19	14	15	8	9	7	7	4	3	3	1			1			
29	1				1	1		3	2	3	6	5	4	1	1	4	6	15	2	13	14	13	19	11	23	15	16	11	16	10	13	3	2	2	2	1					
30				1	1	1	2	1	2	3	6	4	10	3	4	10	12	11	15	8	16	12	13	13	15	14	13	13	12	15	9	6	2	4	1		1				
31		1			3		1	1	1	2		2	3	4	7	5	7	18	9	12	15	16	14	11	9	16	13	20	13	7	6	8	4	1	2		1	1			
32	1				2		4	1	2	4		3	4	4	5	9	3	11	11	9	16	10	14	13	12	19	24	20	16	16	5	8	4	2		1	1	1			
33				1			2		2	1	1	2	5	3	2	3	4	5	7	9	12	12	24	10	23	13	16	11	8	10	7	9	7	7	3	2					
34								1	1			1	3	3	5	1	2	7	5	9	14	12	9	9	15	14	13	11	13	8	4	3	2	5	3		2	1			
35					1				1	2	2	1	2	2		2	2	3	4	5	8	8	6	5	14	15	12	9	14	10	6	3	6	4	3	1	1				1
36						1	1			1		1	1		2	1	2	1	4	2	2	4	7	3	7	10	14	4	8	9		9	3	4	2	1	2	1		1	
37					1						1	1			1	2	1	1	3	3	1	1	4	4	4	6	11	3	2	4	3		4	2	2	1		1			
38															1				2	1	2					6	3	2	4	1	1	1			1						
39																	1			1	2			1			2		1	4	4				1						
40																				1			2			2												1			

	F	S	P1	P2
0	2	0	27	25
1	4	0	7	2
2	4	0	11	12
3	8	2	30	33
4	16	3	49	30
5	17	2	38	39
6	23	1	33	53
7	21	1	36	25
8	24	2	37	31
9	38	0	29	37
10	37	2	50	58
11	54	2	63	45
12	60	5	62	54
13	57	5	103	69
14	68	8	83	119
15	79	7	149	108
16	93	7	160	134
17	125	9	179	218
18	122	18	215	244
19	117	30	320	334
20	167	34	356	424
21	145	27	512	494
22	165	59	628	673
23	141	66	800	921
24	189	106	1038	1108
25	186	121	1351	1403
26	189	162	1656	1658
27	146	196	1921	1882
28	147	216	2209	2157
29	131	242	2187	2236
30	70	253	2236	2363
31	60	233	2276	2360
32	44	255	2296	2289
33	40	221	2135	2159
34	26	176	1929	1880
35	10	153	1412	1464
36	10	108	1074	965
37	8	67	588	599
38	2	25	316	264
39	1	17	114	112
40	1	6	54	47
	2847	2847	28769	29128

Table 2: Univariate marginals for fathers' and sons' Ravens scores, and for two populations

	fathers	sons
mean	+ 21.6178	+ 29.3737
variance	+ 45.4823	+ 24.0774
skewness	- 0.4632	- 1.0020
excess	- 0.0423	+ 2.4278

Table 3: Moments of distributions of fathers' and sons' Raven scores

	0-11	12-17	18-21	22-25	26-29	30-34	
0 - 23	70	78	53	47	38	4	290
24 - 26	46	83	87	95	58	20	389
27 - 29	57	123	128	163	128	55	654
30 - 32	49	130	149	163	182	68	741
33 - 35	19	54	105	157	135	80	550
36 - 40	7	14	29	56	72	45	223
	248	482	551	681	613	272	2847

Table 4: Grouped bivariate distribution of sons' (rows) and fathers' (columns) Raven scores.

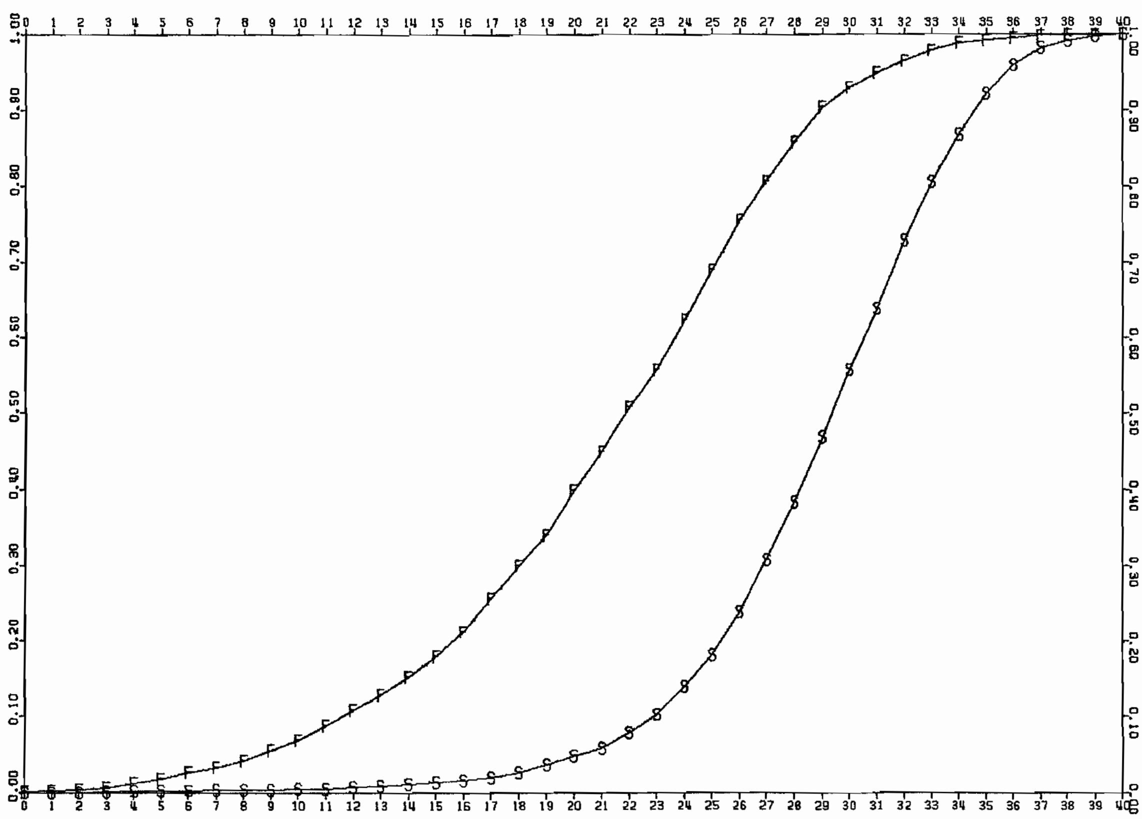

Figure 1. Cumulative distribution of frequencies for fathers' (F) and sons' (S) scores on Raven's Progressive Matrices.

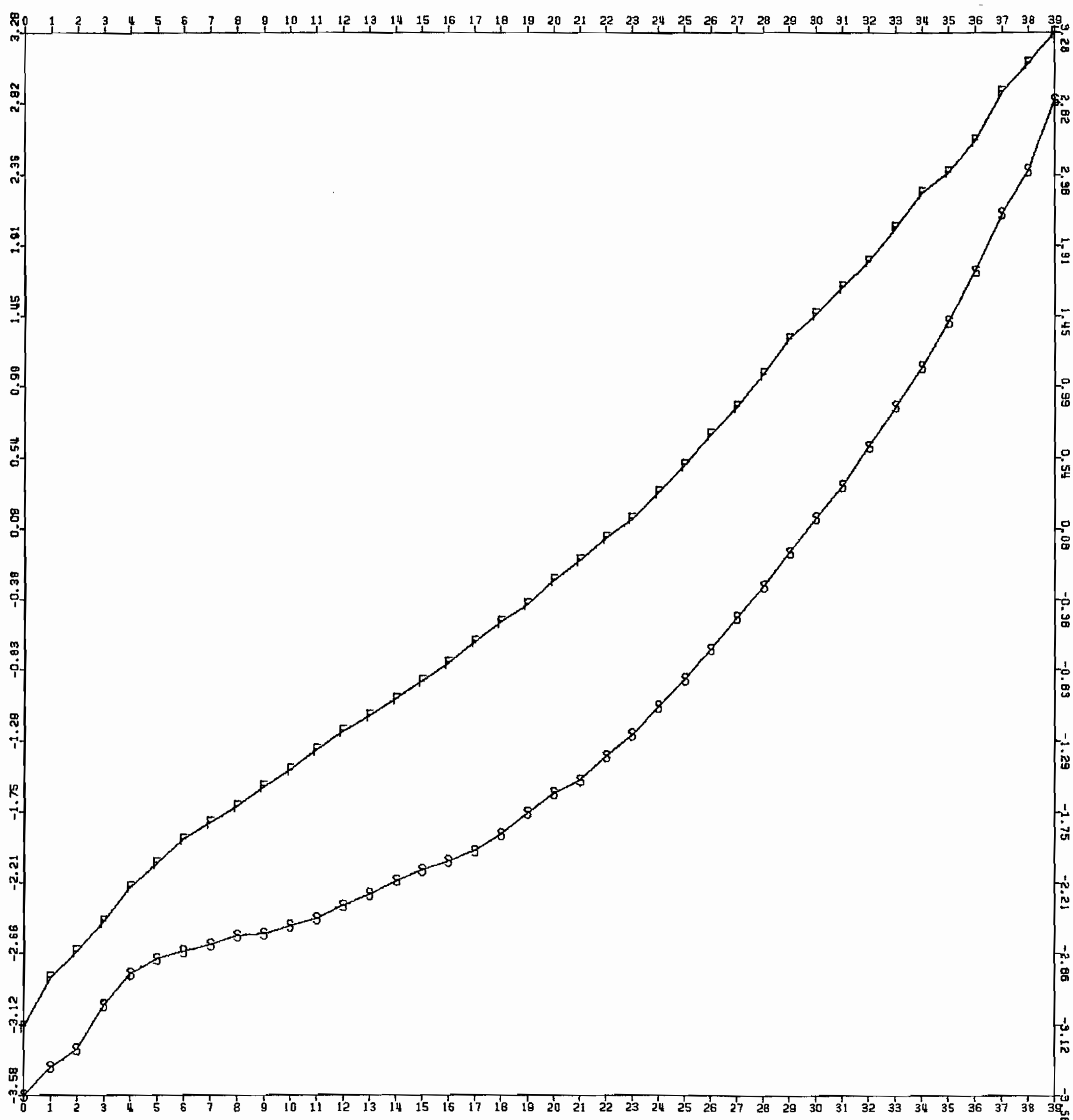

Figure 2. Half-normal plot of distribution of fathers' (F) and sons' (S) scores on Raven's Progressive Matrices.

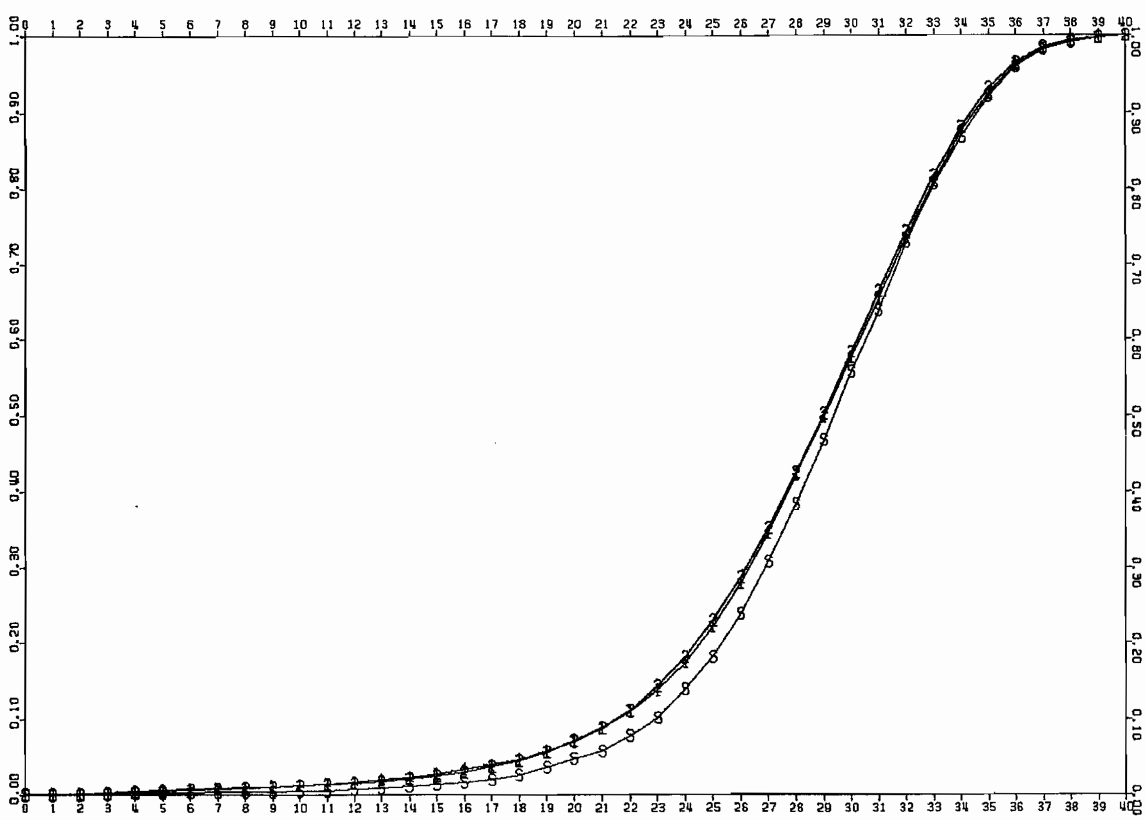

Figure 3. Empirical cdf of sons' scores compared with two population cdf's.

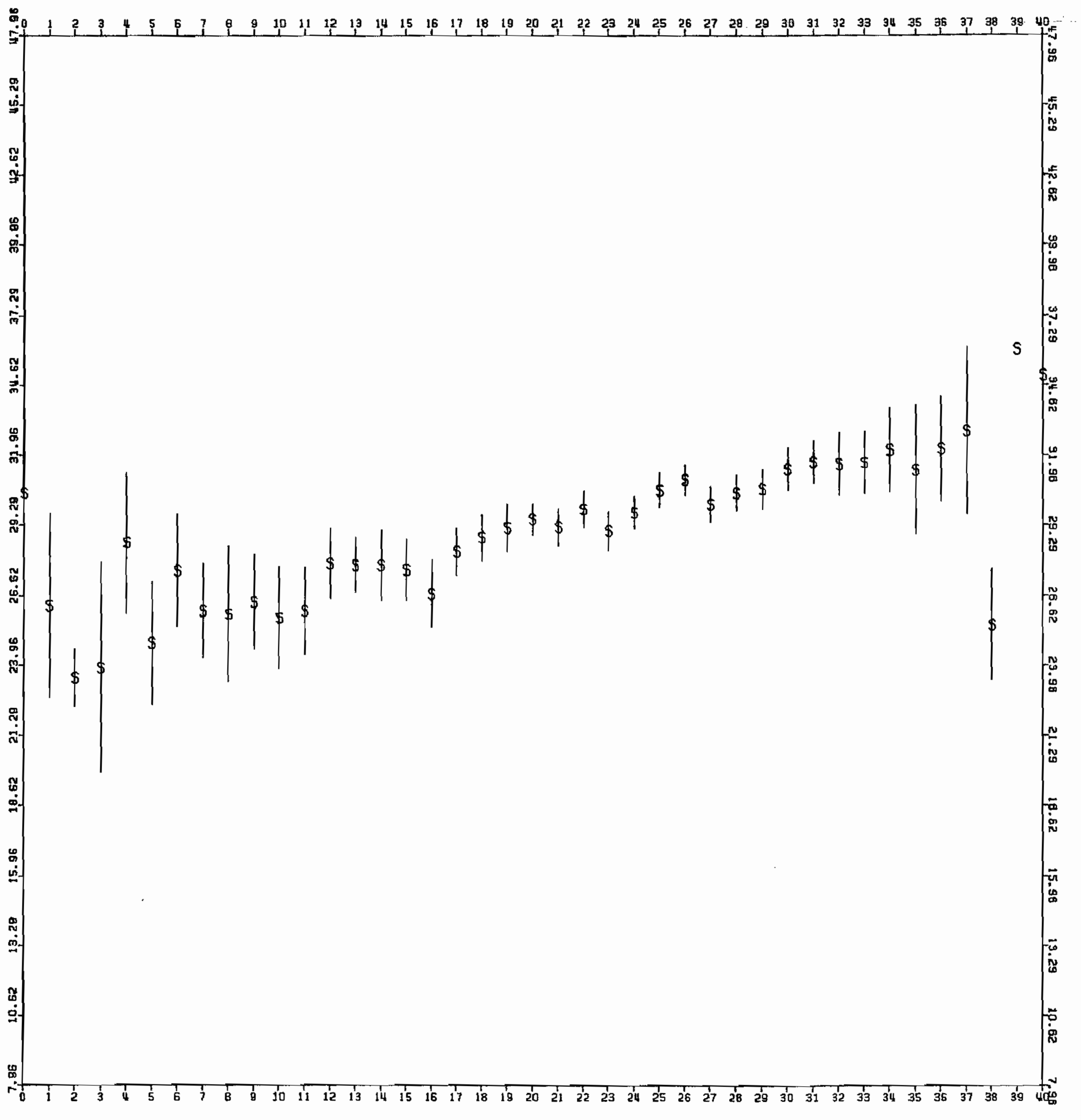

Figure 4. Regression of sons' Raven scores on fathers'.

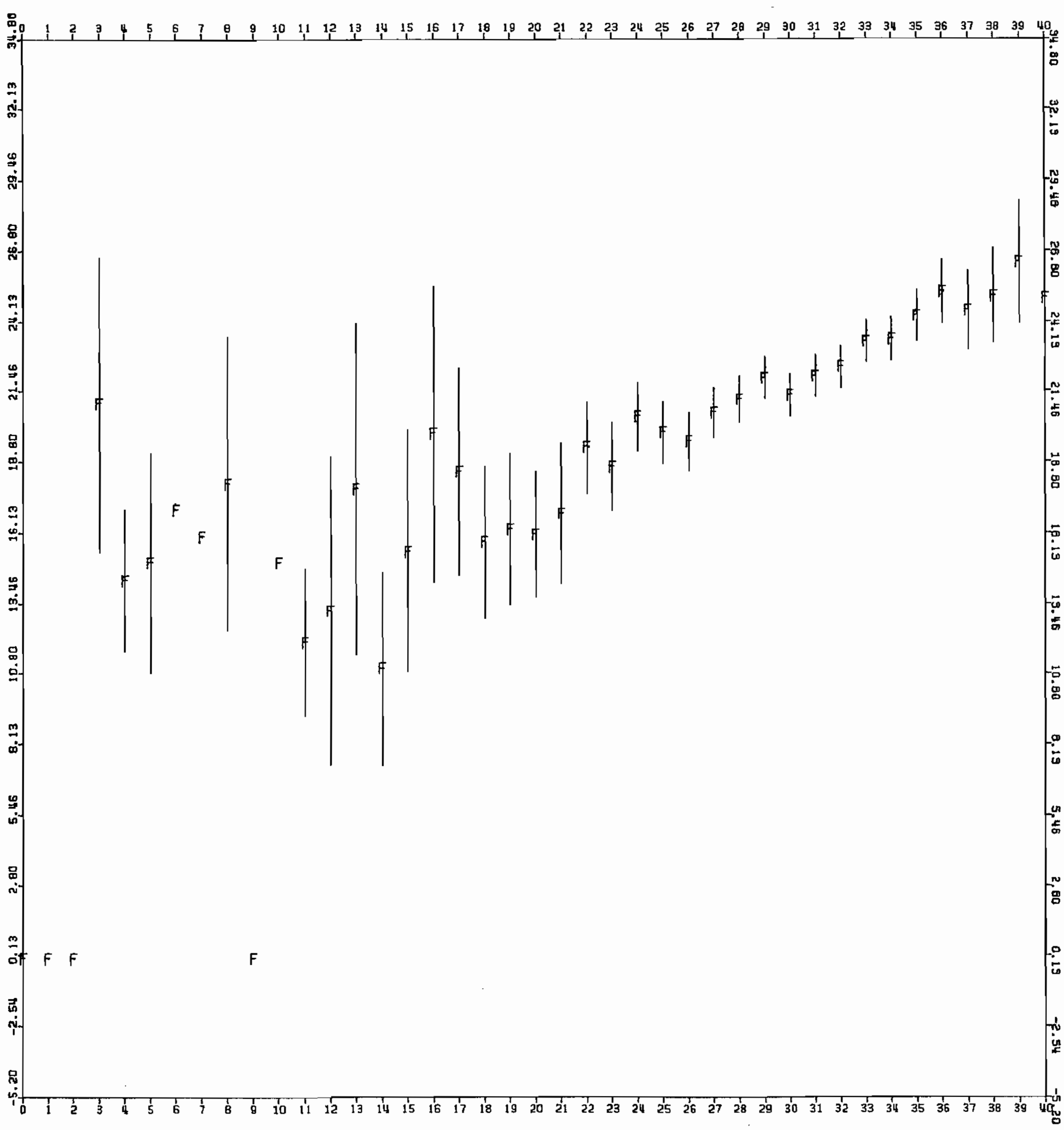

Figure 5. Regression of fathers' Raven scores on sons'.