ÖREBROPROJEKTET

Delstudier

21.

A STRUCTURAL MODEL FOR TESTING THE AGE-DIFFERENTIATION HYPOTHESIS

by

Ulf Olsson and Lars R. Bergman

Department of Statistics
University of Uppsala

Department of Psychology
University of Stockholm

November 1973

Psykologiska institutionen
Stockholms universitet

Psykologiska Institutionen
Box 6706
113 85 Stockholm
Tel. 08-22 81 60/353
Professor David Magnusson

A STRUCTURAL MODEL FOR TESTING THE AGE-DIFFERENTIATION HYPOTHESIS

by

Ulf Olsson and Lars R. Bergman

Department of Statistics
University of Uppsala

Department of Psychology
University of Stockholm

Research reported in this paper has been supported in part by the Swedish Council for Social Science Research under project "Statistical methods for analysis of longitudinal data", project director K G Jöreskog, and in part by the Swedish National Board of Education through the Örebro Project, project directors D Magnusson and A Dunér. This paper is published simultaneously in the report series of the Department of Statistics, University of Uppsala, and in the Örebro project report series, Department of Psychology, University of Stockholm.

Stockholm and Uppsala
25.8.73

ABSTRACT

This study aims at testing the age-differentiation hypothesis of intellectual abilities. The data consist of various tests of intelligence and achievement, collected for a group of 728 Swedish school children, who were tested at the ages 10 and 13. The hypothesis is tested with linear structural models, and using estimation methods developed by Jöreskog. It was concluded that age-differentiation must be viewed as a multidimensional concept, and that the results indicate integration in certain respects and differentiation in other respects. The present results also suggest that the primary mental abilities V, I, and S can be described as developing fairly independently between the ages 10 and 13.

PREFACE

Since 1965 a large longitudinal project has been carried out at the Department of Psychology, University of Stockholm. This project is named the Örebro project, and it aims at studying the adjustment, behavior, and performance of two successive age groups (cohorts) of Örebro children. Each age group consists of about 1000 children, and the same children have been followed since they were about ten years old in 1965. This follow-up will continue until the children are about 25 years old.

Data have been collected on several occasions, and the collected information involves measures of intelligence, creativity, achievement, peer ratings, teachers´ ratings, and questionnaire data covering school adjustment, norms, delinquency, and choice of career line. Most of the results from the Örebro project have been published in Swedish, but a monograph about the project is now in press (Magnusson, D., Dunér, A., & Zetterblom, G.: *Adjustment - a longitudinal study*. Stockholm: Almqvist & Wiksell, in press).

The present study, "A structural model for testing the age-differentiation hypothesis", has been written by Ulf Olsson and Lars R. Bergman, and it aims at testing the age-differentiation hypothesis of intellectual abilities, using intelligence test and achievement test data collected within the project when the children were 10 and 13 years old.

Stockholm in November 1973

David Magnusson Anders Dunér

CONTENTS

1. INTRODUCTION

1.1 Aims of the study

This paper has two aims. The first aim is to clarify and test certain aspects of the age-differentiation hypothesis. In doing this some longitudinal models will be presented for the structural development of a group of abilities for 728 children during the years 9:6 - 12:6, and these results will be interpreted in terms of the age-differentiation hypothesis.

The second aim regards statistical methodology. Linear structural models which emerge in longitudinal research have frequently been difficult to estimate and test statistically. The estimation method used here, which was originated by Jöreskog (1973), is general enough to be used in a wide variety of applications, and the present paper may be regarded as a means of spreading knowledge of this method The problem of statistical model building will also be discussed.

1.2 Psychological background

When considering the ontogenetic development of intelligence from a multivariate perspective, questions regarding the development of the structure of abilities, and the relationship between abilities, naturally arise. A frequently held hypothesis concerning intellectual development is the age-differentiation hypothesis stated by Garrett (1946, p 373) as: "Abstract or symbol intelligence changes in its organization as age increases from a fairly unified and general ability to a loosely organized group of abilities or factors". It should be noted that age-differentiation appears to be a rather vague concept, and that age-differentiation, or the lack of it, has been inferred from different properties of the analyzed data. For example, the intercorrelations between a set of abilities or the factor patterns of a set of abilities, has been compared between different occasions for the same group, or between different age-groups etc. The assumption that the so called g-factor (often regarded as a second order factor "saturating" all first order ability factors) should diminish and eventually vanish as one moves from early childhood to adolescence is closely linked to the age-differentiation hypothesis. Reinert (1970) has suggested that a hypothesis of "performance-differentiation" may be more parisimonious, i.e. "the degree of differentiation of the factor structure of intelligence is dependent upon the absolute level of intellectual performance" (p 474). This hypothesis points to the fact that different samples of the same age group may exhibit different degrees of differentiation.

Considering the complexity of the problem, and the different data collection strategies and tests used in different investigations, it is hardly surprising that different investigators have come up with different results, and that the age-differentiation hypothesis has caused a prolonged debate, being supported by, e.g., Burt (1954), and being opposed by, e.g., Guilford (1967). Recent reviews by Anastasi (1970) and by Reinert (1970) seem on the whole to support the age-differentiation hypothesis, but this matter is far from settled. For example, in a recent study Fitzgerald, Nesselroade, and Baltes (1973) found no evidence of age-differentiation during adolescence, using factors rotated towards Thurstone´s target pattern of adult intelligence by means of a Procrustes routine. This was done for several age groups,ranging from 12:6 to 17:6.

The conflicting results have been ascribed to various deficiences in the investigations, and it appears useful if a study of age-differentiation has the following properties:
(1) The sample is large and reasonable representative; (2) a repeated measurements design is used; (3) if possible, the same tests are used at all times of measurement; (4) factor analytic techniques, which permit the simultaneous analysis of data from different occasions are employed; (5) the analyses describe how the factor pattern changes with time. In this context a repeated measurements design has two useful properties, namely the property of providing information about individual change, and the property of having all results referring to the same group of individuals. Using cross-sectional data only group change can be studied, and if the several groups involved in such a study are not all comparable it is not clear to what population the results refer. In section 1.3 a more detailed discussion of points (4) and (5) is presented. Only the case where the same sample has been subject to testing on two occasions will be treated.

Some of the conflicting results appearing in earlier research in this area may depend partly on the age range of the study. For example, Reichard (1944) suggested an integration-differentiation hypothesis, and demonstrated a continued integration within the age range 9-12, and a differentiation within the age range 12-15. On the whole, there appears to be rather scanty evidence relevant to the age-differentiation hypothesis for the age-span 9-10 to 12-13. Reinert´s (1970) review only includes three longitudinal studies covering approximately this age-span, namely Asch (1936), Burt (1954), and Cropley (1964). Of these Cropley´s study supports integration, and the other two differentiation. Asch´s and Cropley´s studies appears to be based on selective samples, which makes it difficult to generalize from their results. Studying 326 boys, Burt (1954) found evidence for age-differentiation between the ages 9-10 and 13-14 for a battery of ability tests. It is possible that differentiation parallels the mental growth spurt at pubescence for which there is some evidence (Ljung, 1965). It is then possible that Burt would have obtained different results if he had retested his sample at the age 12-13 instead of 13-14.

In the present investigation, the children are tested at the age of 9:6 and retested at the age of 12:6. A comparison between the sexes may then prove interesting since the boys, but not the girls, can be viewed as mainly prepubescent. Since there is also evidence for sex differences in factor patterns (Anastasi, 1970, p 904), it was decided to analyze boys and girls separately.

I.3 Statistical background

As mentioned above, the question of how to infer age-differentiation from data is not settled. The survey given by Reinert (1970) indicates that the following outcomes have been regarded as supporting the age-differentiation hypothesis:

(1) An increase in the number of common factors;
(2) a decrease in the variance proportion explained by the g-factor;
(3) an increase in the variance proportion explained by other factors;

(4) changes in the factor pattern;
(5) the correlations between factors decrease;
(6) a decrease in magnitude of the communalities;
(7) a decrease in the average size of the intercorrelations between tests.

Most of these indicators (except for the last) are in one way or another influenced by subjective decisions. The first and the sixth indicator depend on the criterion used for deciding the number of common factors, whereas the others are affected by the methods used for rotating the obtained solution (e.g. extracting a g-factor vs extracting several correlated group factors). A certain indicator may also be contradictive, and two indicators may give different results. E.g., it is possible to have increased correlations between some of the factors, whereas the correlations between other factors decrease. It seems clear, that age-differentiation must be viewed as a multidimensional concept, to be treated with multidimensional methods.

Consider now the design of a study of age-differentiation. One approach is to compare data collected from several different cohorts on the same occasion, and another approach is to follow one cohort longitudinally. With the first approach, the problem of how to make the factor analysis solutions for the different cohorts comparable may be solved by Procrustes rotation (cf. Fitzgerald, Nesselroade & Baltes, 1973), or by making a simultaneous factor analysis including all cohorts (Jöreskog, 1971b, van Thillo and Jöreskog, 1970). The longitudinal design, on the other hand, has the potency of offering a more direct description of how change occurs. Consider the complete covariance matrix of responses in the case of measurements on two occasions. The information regarding the between-occasions covariances is not at hand in cross-sectional designs, and this information has not been used in most investigations involving longitudinal data. It appears more powerful to perform some kind of integrated longitudinal analysis which makes effective use of all the information available. By "effective use" is meant that the method for analysis should have the following properties.

(1) It should analyze the complete covariance matrix simultaneously; (2) it should be of a confirmative kind, i.e., when the factor properties of the tests are known, it should be possible to approximate their negligible loadings by zeroes to make sure that no further rotation will be needed; (3) if possible, the structure of the dependence between factors over time should be included in the model. Jöreskog and van Thillo´s (1972) method LISREL for estimating simultaneous linear relations has these three properities.

The LISREL model may be described shortly as follows. Denote the vector of test scores for the first occasion by $\underline{x}$, and the corresponding vector for the second occasion by $\underline{y}$. These vectors may be of different order (corresponding to the number of tests used), say q and p respectively. For both occasions, a factor model is assumed to hold, i.e.

$$\begin{aligned} \underline{x} &= \Lambda_x \underline{\xi} + \underline{\delta} \quad \text{and} \\ \underline{y} &= \Lambda_y \underline{\eta} + \underline{\varepsilon} \end{aligned} \tag{1}$$

Here Λ_x (qxn) and Λ_y (pxm) are the factor matrices for the two occasions, and $\underline{\xi}$ and $\underline{\eta}$ are the corresponding vectors of factor scores, assumed to be of order n and m respectively. $\underline{\delta}$ and $\underline{\varepsilon}$ are vectors of disturbances, i.e., "unique parts", of orders q and p. A general linear relation between $\underline{\xi}$ and $\underline{\eta}$ may be written as

$$B\underline{\eta} = \Gamma\underline{\xi} + \underline{\zeta} \tag{2}$$

where B and Γ are coefficient matrices of orders mxm and mxn, respectively, characterizing the structural relation, and $\underline{\zeta}$ is a vector of residuals. Here, we have neglected the constant term in the relation by imposing the constraint $E(\underline{\eta}) = E(\underline{\xi}) = \underline{0}$. In a longitudinal setting, no influence from $\underline{\eta}$ to $\underline{\xi}$ is assumed, since $\underline{\eta}$ is measured at a later occasion than $\underline{\xi}$. Thus, the model (2) is in this case simplified to

$$\underline{\eta} = \Gamma\underline{\xi} + \underline{\zeta}, \tag{3}$$

i.e. every factor score on the second occasion is regarded as a linear combination of the factor scores on the first occasion, plus a residual. If we assume that the residuals $\underline{\delta}$ are uncorrelated, and similarly that the residuals $\underline{\varepsilon}$ are uncorrelated, their respective covariance matrices Θ_δ and Θ_ε are diagonal, with diagonal elements corresponding to the unique variances. The covariance matrix for $\underline{\xi}$ will be denoted by Φ, and the covariance matrix for $\underline{\zeta}$ by Ψ. When all parameter matrices are known, the covariance matrix for $\underline{\eta}$ may be computed as $\Gamma\Phi\Gamma' + \Psi$, in this case where $B = I$. Otherwise a more complicated formula is used (see Jöreskog & van Thillo, 1973). The parameters in the matrices

$$\Lambda_y, \Lambda_x, \Gamma, \Phi, \Psi, \Theta_\delta \text{ and } \Theta_\varepsilon$$

may be of three kinds:

(1) fixed parameters, that have been assigned given values

(2) constrained parameters that are unknown but equal to one or more other parameters

and

(3) free parameters that are unknown and not constrained to be equal to any other parameter.

The concept of fixed parameter has already been applied above, in that B has been set equal to I. In applications of the model it is necessary to make sure that all parameters are identifiable, so that unique solutions may be found. A computer program which performs all the computations necessary for estimating the parameters by the maximum-likelihood method has been written (Jöreskog and van Thillo, 1973).

The LISREL-program provides a χ^2-test for goodness of fit of the model, and this is the case also for the other computer programs that will be used in the analyses. The hypothesis when using e.g. the LISREL-model is, that formulas (1) and (3) hold exactly in the population, with the proper restrictions inserted. As Jöreskog (1971 a, p 131) points out in another context, any such hypothesis is not very realistic, and a sufficiently large sample would no doubt create a value of χ^2 large enough for rejecting the hypothesis. It seems more reasonable to regard formulas (1) and (3) as an approximation to the true state of the affairs. With this philosophy, the χ^2-value is regarded not as a test in the strict statistical sense, but as a means for comparing different specifications of the model through comparison of their χ^2-values. This problem has recently been discussed by Tucker & Lewis (1973), who designed a reliability coefficient as an indicator of the"goodness" of a model.

2 DATA

2.1 Subjects

This study is based on data originally collected for the Örebro project by Magnusson, Dunér and Zetterblom (1973). The population is defined as all school children in the Swedish town of Örebro who in 1965 were in grade 3 and received normal schooling (about 85 % of all the children in the appropriate age group in Örebro). The educational and occupational status is somewhat higher in Örebro than in normal Swedish urban populations, and migration is comparatively small. Results obtained by Bergman (1973) indicate that the Örebro population is presumably 0.1-0.3 sd units above other "normal" Swedish urban populations with regard to general intelligence.

Since only individuals with complete data from both occasions were included in the analyses, a certain drop out was unavoidable. The population, sample, and drop out sizes are given in Table 1. In Table 2, the sample means and the estimated population means in 1965 are given for composite tests of general intelligence, and of general achievement. From Table 2 it can be seen that the sample means are somewhat higher than the corresponding estimated population means, but it seems reasonable to believe that the sample is representative of the population with regard to general tendences.

Table 1. Population, sample and drop out group sizes

	pop.	sample	drop out	% drop out
boys	476	353	123	25.8
girls	490	375	115	23.5

Table 2. Sample means, and estimated population means in general intelligence (G) and general achievement (A). (All data from 1965.)

	G		A	
	pop.	sample	pop.	sample
boys	106.9	108.4	74.7	75.4
girls	105.4	105.7	77.5	77.7

Note - Estimated sd´s of G and A for the population are about 25 and 19, respectively.

2.2 Variables

Six ability tests from the DBA test battery (Härnqvist, 1962) were given to the subjects, both in grade 3 when they were 9-10 years old, and in grade 6 when they were 12-13 years old. Achievement tests in Swedish and in Mathematics were given both in grade 3 and in grade 6, but different tests were used on the different occasions. In the factor analyses raw scores were used, both with regard to the ability and the achievement tests. A brief description of the tests is given below.

Synonyms (S). The S test is often regarded as measuring Verbal Comprehension (40 items).

Opposites(O). The O test is often regarded as measuring Verbal Comprehension (40 items).

Letter Groups (L). Among four groups of letters, one group is to be found which differs from the other groups. The L test is often regarded as measuring an abstract-logic ability (Inductive ability) (30 items).

Figure sequences (F). The subject indicates which figure follows a given sequence of figures. The F test is often regarded as measuring an abstract-logic ability (30 items).

Cube counting (C). A picture of an arrangement of cubes is given. The task is to count the number of cubes. The C test if often regarded as measuring Spatial ability (40 items).

Metal folding (M). The task is to find the three-dimensional object, among four choices, that can be made from a pictured flat piece of metal with bending lines marked on the drawing. The M test is often regarded as measuring Spatial ability (30 items).

Achievement in Swedish (A_s). The A_s tests contain several subscales measuring various aspects of facility in the Swedish language, e.g., a vocabulary test, reading comprehension tests etc. The test used in 1965 and the test used in 1968 are very similar, although not identical.

Achievement in Mathematics (A_m). The A_m tests contain several subscales measuring facility in performing simple arithmetic operations, in solving simple applied problems etc. The A_m test 1965 and the A_m test 1968 are fairly similar, except that in the 1968 test geometry items are introduced.

Reported split-half reliabilities for S, O, L, F, C, and M range between 0.85 and 0.90 (Härnqvist, 1962). No reliability estimates are available for the A_s and A_m tests, but the reliabilities can be assumed to be quite high (well above 0.90). In grade 6, moderate ceiling effects are present, but probably not to such an extent as to seriously affect the results. (At most 4 % obtain maximum score, which occurs for girls in grade 6 for L.) Measurements taken in 1965 will be indicated by the subscript "1", and measurements taken in 1968 by the subscript "2", e.g., S_1, S_2.

Henrysson (1965) factor analyzed DBA test data collected by Härnqvist (1960), using both graphical, varimax and oblique rotation, and his results seem to support the above grouping of the tests into S and O measuring Verbal Comprehension, L and F measuring Inductive ability, and C and M measuring Spatial ability.

3 MODELS, ANALYSES, AND RESULTS

3.1 The problem of finding a good model

The problem is to find a structural model that clearly and fairly accurately describes change in mental ability structure between grades 3 and 6. In doing this the following steps are involved. (1) The simplest unrestricted solutions giving satisfactory fit to data are obtained; one solution for each occasion, and one solution for the combined data from the two occasions; (2) restricted solutions are obtained for some reasonable models for each occasion, and these solutions are compared to the corresponding unrestricted solution, and to other criteria, with regard to fit; (3) the best restricted model for the first occasion, and the best restricted model for the second occasion, are used in building some reasonable structural models for the relationships between the factors between occasions, and these solutions are compared to the corresponding unrestricted solution, and to other criteria, with regard to fit. However, when testing a structural model against the corresponding, unrestricted solution it is obvious that no reasonably simple model could summarize so much information without giving a significantly worse fit than the corresponding unrestricted model. Therefore, when judging the fit of a structural model it appears better to use some measure of overall fit, which is less affected by moderate deviations from the model. For this purpose we used the Tucker-Lewis (1973) reliability coefficient, and the size of the elements of the residual matrix.

In steps (1) and (2) above, it is necessary to choose between different restricted models. The primary criteria for accepting a model were: (1) The fit of the model, as measured by χ^2, should not be significantly worse ($p<.01$) than the fit for the corresponding unrestricted model; (2) the model should have a reasonable psychogical interpretation. In those cases criteria (1) and (2) were not sufficient for deciding between the models the additional criteria of simplicity and of overall fit, as measured by the Tucker-Lewis (1973) reliability coefficient and the residual matrix, were employed. When deciding between different structural models, the above rationale was followed, except that the main emphasis was put on the Tucker-Lewis reliability coefficient rather than on χ^2.

3.2 Unrestricted analyses

Unrestricted factor analyses were performed on the data from each occasion for k = 2, 3, and 4 factors, and on the combined data from the two occasions for k = 3, 4, 5, 6, 7, and 8 factors. The analyses were performed using a computer program UFABY3 by Jöreskog and van Thillo (1971). The correlation matrices that the analyzes are based on are given in Appendices 1 and 2. The results for the unrestricted analyses are given in Table 3.

Table 3. Values of χ^2 for different numbers of factors (k) for the unrestricted solutions. (Degrees of freedom for χ^2 within parantheses.)

		grade 3		grade 6		grades 3 and 6		
	k	χ^2	prob. level	χ^2	prob. level	k	χ^2	prob. level
	2	63.69(13)	.000	47.85(13)	.000	6	69.14(39)	.002
Boys	3	16.40(7)	.022	22.21(7)	.002	7	47.83(29)	.015
	4	7.79(2)	.020	0.86(2)	.650	8	22.73(20)	.301
	2	25.17(13)	.022	37.79(13)	.000	6	84.54(39)	.000
Girls	3	1.62(7)	.978	16.13(7)	.024	7	33.67(29)	.252
	4	0.32(2)	.853	1.70(2)	.427	8	14.51(20)	.805

It is seen from Table 3 that for both boys and girls in grade 3, the most useful solution is for k = 3. This solution has a very good fit, and it is easily interpreted psychologically. In grade 6, the additional factor for k = 4 has a very clear interpretation as a school achievement factor or knowledge factor. Since the fit is markedly improved for both boys and girls by adding a fourth factor, the solution for k = 4 was chosen for both sexes in grade 6. For the combined grade 3 and 6 data, it appears clear that the solution for k = 7 is preferable for both boys and girls.

For both sexes the solutions for k = 3 in grade 3, for k = = 4 in grade 6, and the solutions for k = 7 for the combined data will be the unrestricted analyses to be compared with the corresponding restricted analyses.

3.3 Restricted analyses

For grade 3 data and grade 6 data separately, a number of restricted factor analyses were performed. The loadings for tests, which were believed to be small in a certain factor, were set equal to zero. The analyses were performed using the computer program SIFASP by van Thillo and Jöreskog (1970), and the factors were permitted to correlate, i.e., oblique solutions were obtained. Using information from the unrestricted analyses, from the expected factor structure of the tests, and from the initial restricted analyses, a number of restricted solutions were obtained. Lack of space permits a presentation of the various restricted models, and in Table 4 is only given the solutions that were finally chosen by applying the criteria stated in 3.1. Since the factors had clear interpretations, they were named for the reader´s convenience with V standing for Verbal Comprehension, I standing for Inductive ability, S standing for Spatial ability, and K standing for a knowledge factor or a school achievement factor.

Table 4. Restricted factor solutions for grades 3 and 6. Decimal points omitted for the factor loadings. (Degrees of freedom and probability level for χ^2 within parantheses.)

		grade 3						grade 6							
		boys			girls			boys				girls			
	factors	V_1	I_1	S_1	V_1	I_1	S_1	V_2	I_2	S_2	K_2	V_2	I_2	S_2	K_2
	S	87	0	0	88	0	0	93	0	0	0	88	0	0	0
	0	85	0	0	87	0	0	90	0	0	0	87	0	0	0
	A_s	54	45	0	57	40	0	55	0	0	43	54	0	0	45
Tests	L	0	71	0	0	68	0	0	89	0	0	0	83	0	0
	F	0	38	32	0	41	28	0	38	42	0	0	53	25	0
	A_m	0	82	0	0	82	0	0	0	38	59	0	0	45	52
	C	0	0	58	0	0	62	0	0	67	0	0	0	68	0
	M	0	0	84	0	0	75	0	0	72	0	0	0	78	0

$\chi^2(15)=26.93$ $\chi^2(15)=12.35$ $\chi^2(11)=16.17$ $\chi^2(11)=12.80$

$(p = .03)$ $(p = .65)$ $(p = .14)$ $(p = .31)$

$\Delta\chi^2(8)=10.53$ $\Delta\chi^2(8)=10.73$ $\Delta\chi^2(9)=15.31$ $\Delta\chi^2(9) = 11.10$

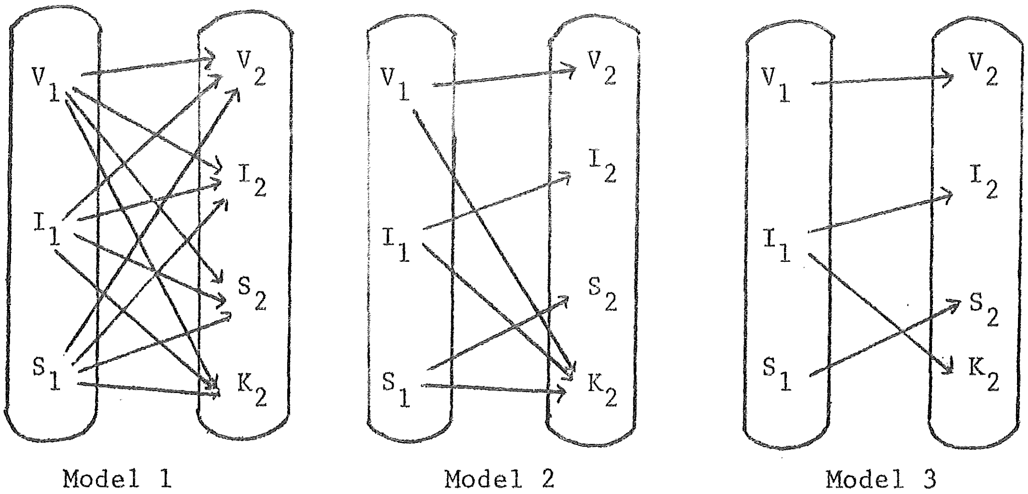
$(p = .24)$ $(p = .22)$ $(p = .08)$ $(p = .27)$

Note - χ^2 indicate the fit of the model, and $\Delta\chi^2$ indicate the χ^2 value of the difference in fit between a restricted model and the corresponding unrestricted model. The zeroes were specified a priori.

3.4 Structural models

The task is now to build structural models for the relationships of the abilities between grades 3 and 6. For this purpose the earlier described LISREL model is used. In building the structural models the factor patterns of the restricted models from grades 3 and 6, respectively, were used.

Figure 1. Simplified graphical representation of three structural models for the relationship of the abilities between grades 3 and 6.



Note - A closed area indicate that all factors within that area are correlated.

As a starting point Model 1 in Figure 1 was fitted to data. This model was considered to be a very general model in the sense that all factors in grade 3 are related to all factors in grade 6. However χ^2, or $\Delta\chi^2$ in relation to the unrestricted 7 factors solution, is quite large with a probability level of .000. It was argued in section 3.1 that this should be expected when a model was fitted to so much data, and that in this case the Tucker-Lewis reliability coefficient and the size of the elements of the residual matrix are better measures of the fit of the model to data. It is seen from Table 5 that juding from these criteria, the fit of Model 1 to data is quite good.

Table 5. Fit of models 1, 2, and 3 to data as measured by χ^2, the Tucker-Lewis reliability coefficient, and average size of the residuals of the residual matrix (df for a certain χ^2 within parantheses)

Model		χ^2	prob. level	$\Delta\chi^2$	prob. level	Tucker´s reliability coefficient	average size of residuals
	1	227.59(74)	.000			.938	.020
boys	2	238.47(80)	.000	10.88(6)	.092	.941	.022
	3	269.66(90)	.000	42.07(16)	.000	.947	.025
	1	184.03(74)	.000			.953	.020
girls	2	188.93(80)	.000	4.90(6)	.557	.955	.022
	3	242.76(90)	.000	58.73(16)	.000	.953	.026

Note - $\Delta\chi^2$ indicates the fit of a model in relation to the fit of model 1.

Since Model 1 is fairly complicated, it was decided to fit two simpler models to data. These models are also illustrated in Figure 1. In Model 2 it is assumed, that all influence on the factors in grade 6 comes from the corresponding grade 3 factors, except for the knowledge factor which is assumed to depend on all grade 3 factors. Model 3 is still more simplified, and it is assumed that the knowledge factor in grade 6 is influenced only by the inductive factor in grade 3.

From Table 5 it is seen that, as judged by χ^2, Model 3 but not Model 2 has a significantly worse fit than Model 1, but as judged by the Tucker-Lewis reliability coefficient and the average size of the residuals both Models 2 and 3 fitted data quite well. Since Model 3 is simpler than Model 2, Model 3 was accepted as describing the relationships of the abilities between grades 3 and 6. It should be noted that the relationships that are assumed to be zero in Model 3 are quite close to zero in Model 1, where they are not assumed to be zero. A fourth model was fitted to data in which it was assumed that all growth operated through the relationship between a g-factor in grade 3 and a g-factor in grade 6. However, this model did not have a satisfactory fit to data. Model 3 is illustrated in Figure 2a for boys, and in Figure 2b for girls. From the model follows that the factor scores in grade 6 can be described as depending only on the corresponding factor scores in grade 3, and that these relationships are quite strong. The factor correlations are somewhat lower in grade 6, but on the other hand an additional factor is extracted in grade 6. Generally, the unigue variances are smaller in grade 6.

4 DISCUSSION

Do these results indicate age-differentiation or integration? In terms of the seven indicators given in section 1.3, the following interpretations are obtained. The increase in the number of common factors supports differentiation, and the decrease of the unique variances support integration. No clear results are obtained with regard to factor correlations, and intercorrelations between tests. Taking into consideration the multidimensionality of the concept of age-differentiation this result is not surprising, and it appears more fruitful to interpret the results in terms of the total picture of the ability development that is given by the structural model.

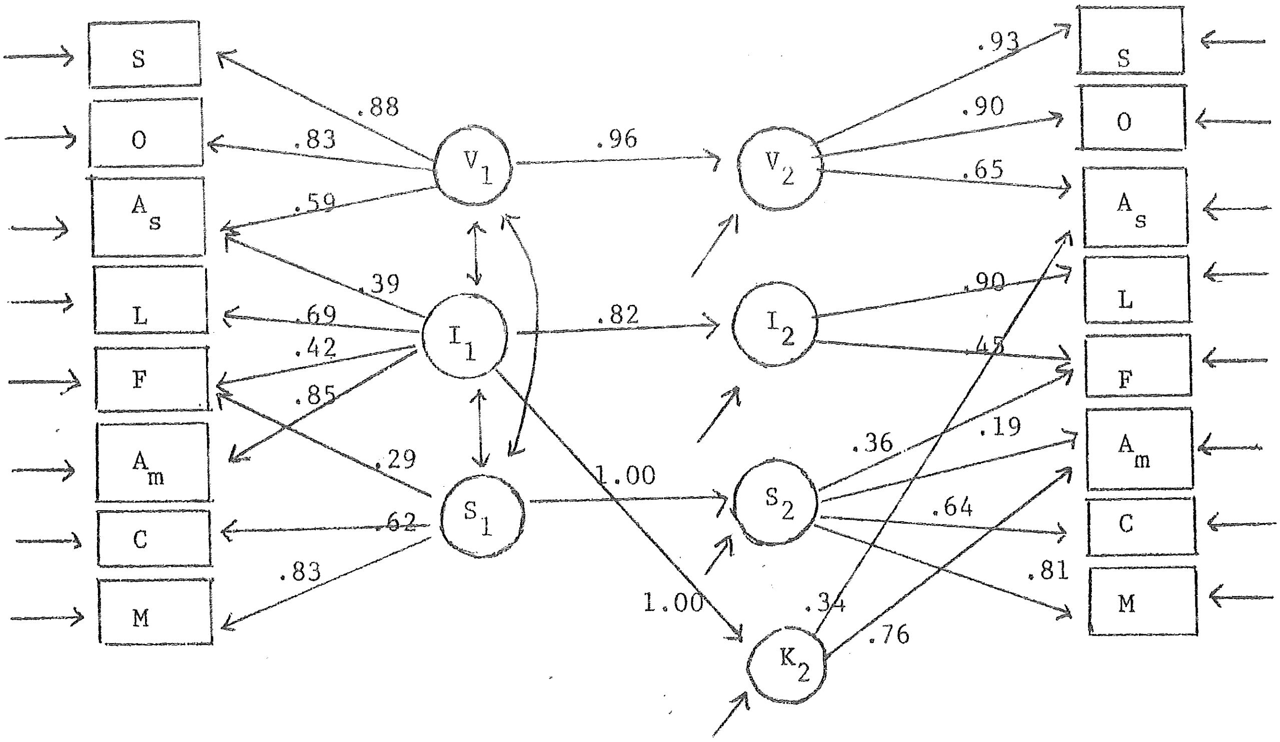
In Model 3, the strong relationship between a grade 3 factor and the corresponding grade 6 factor is notable (this was also the case for Models 1 and 2). This may indicate that the traits measured by the V and S factors are nearly the same on both occasions, although differences in level exist. With regard to the inductive factor, this factor is represented by one factor in grade 3, but by two factors in grade 6. It appears reasonable to interpret these two factors as, respectively, a "fluid" and a "crystallized" inductive ability factor. One aspect of development during the present age-span may be a differentiation between performance on tests measuring "fluid" ability and performance on tests measuring "crystallized" ability.

According to Model 3, development or change in a factor can be described as depending solely on the initial factor score of the corresponding factor, and not on the initial scores in the other factors. This suggests that the primary abilities V, I, and S develop fairly independently during the present age-span. It is interesting to note the marked decrease in the unique variances. More of the variance of the response variates is accounted for by the model in grade 6 than in grade 3. This suggests that although the structural model indicates differentiation in some respects, as indicated above, the relationships are stronger between the factors in grade 6 in the sense that the factors, and the relationships between factors, explain more of the variance of the response variates. In one sense this indicates integration in that the response variates have become more accountable by a number of primary abilities.

No clear sex differences were obtained, except that the girls appeared to fit the model better; in most variables the unique variances were also smaller for girls. These results do not give support to a hypothesis of sex differences in intellectual development during the present age-span.

In comparison to many other studies within this area, the present sample must be considered as fairly good. It appears reasonable to believe that with regard to the aspects studied here, the sample is fairly representative of a "normal" Swedish population. The generalizability of the results are restricted mainly by the age-span, and by the tests used. With regard to the age-span, it appears impossible to generalize to other age-spans since apparently very different results are obtained, depending on the age of the subjects. With regard to the tests used, it is probable that different results would have been obtained if a different set of tests, measuring other abilities, were used. Nevertheless, the tests used here cover some important aspects of intellectual performance,and the results obtained here suggest that, among other things, it may be of interest to investigate to what extent other abilities can be described as developing independently.

Figure 2a. A graphical representation of a structural model for describing the relationships of the abilities between grades 3 and 6 for boys.



Factor correlations:

grade 3

	V_1	I_1	S_1
V_1	1		
I_1	.69	1	
S_1	.43	.61	1

grade 6

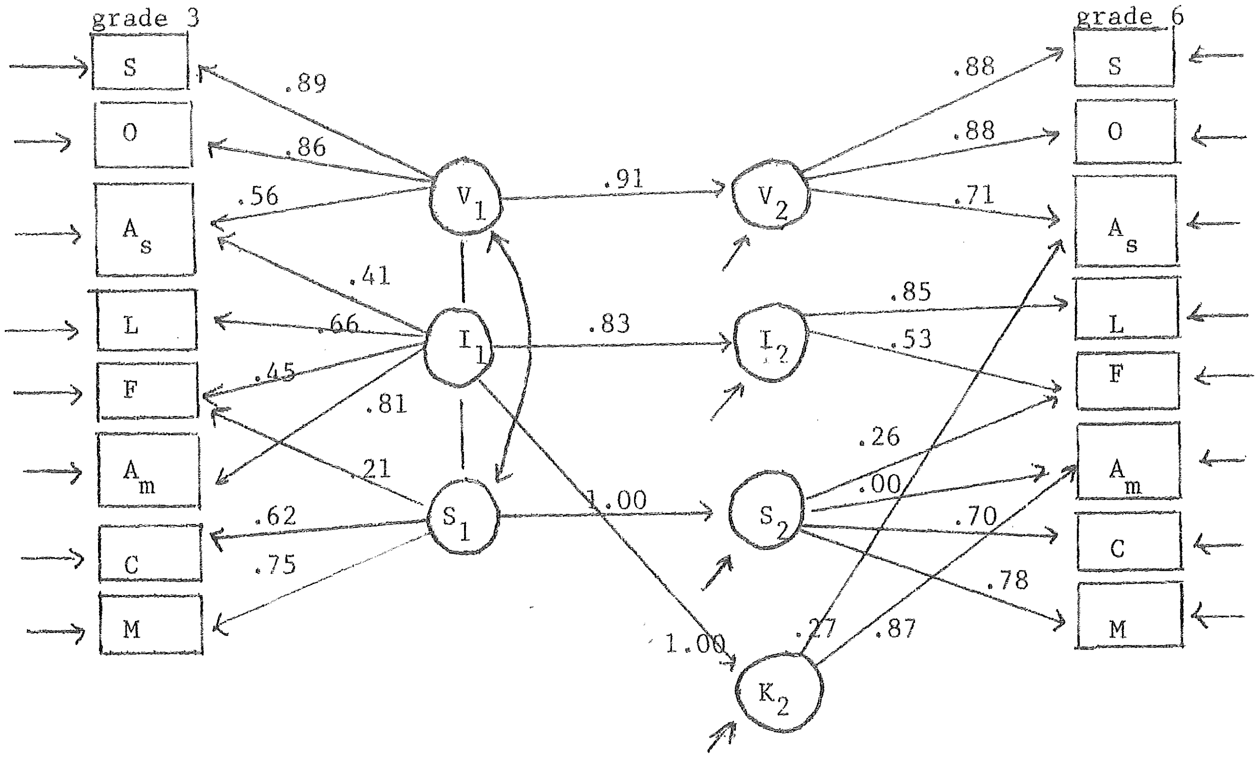
	V_2	I_2	S_2	K_2
V_2	1			
I_2	.64	1		
S_2	.41	.50	1	
K_2	.66	.76	.61	1

Unique variances

	S	O	A_s	L	F	A_m	C	M
grade 3	.48	.56	.43	.73	.77	.53	.78	.56
grade 6	.38	.44	.42	.45	.71	.47	.77	.59

Note - The subscripts 1 and 2 indicate that the factors relate to grade 3 and 6 data, respectively. Arrows "coming from nowhere" indicate residuals.

Figure 2b. A graphical representation of a structural model for describing the relationships of the abilities between grades 3 and 6 for girls.



Factor correlations:

grade 3

	V_1	I_1	S_1
V_1	1		
I_1	.65	1	
S_1	.43	.78	1

grade 6

	V_2	I_2	S_2	K_2
V_2	1			
I_2	.60	1		
S_2	.39	.65	1	
K_2	.59	.81	.78	1

Unique variances:

	S	O	A_s	L	F	A_m	C	M
grade 3	.45	.52	.48	.75	.78	.59	.79	.66
grade 6	.48	.49	.45	.53	.68	.50	.71	.63

Note - The subscripts 1 and 2 indicate that the factors relate to grade 3 and 6 data, respectively. Arrows "coming from nowhere" indicate residuals.

REFERENCES

Anastasi, A. On the formation of psychological traits. American Psychologist, 1970, 25, 899-900.

Asch, S.E. A study of change in mental organization. Archieves of Psychology, 1936, No 195, 1-30.

Bergman, L.R. Parents´ education and mean change in intelligence. Scandinavian Journal of Psychology, 15, 1973, in press.

Burt, C. The differentiation of intellectual ability. British Journal of Educational Psychology, 1954, 24, 76-90.

Cropley, A.J. Differentiation of abilities. Socioeconomic status and the Wisc. Journal of consulting psychology, 1964, 28, 512-517.

Fitzgerald, J.M., Nesselroade, J.R., & Baltes, P.B. Emergence of adult intellectual structure: Prior to, or during adolescence? Developmental Psychology, 1973, in press.

Garrett, H.E. A developmental theory of intelligence. American Psychologist, 1946, 1, 372-378.

Guilford, J.P. The nature of human intelligence. New York: McGraw-Hill, 1967.

Henrysson, S. Faktoranalys av Härnqvists anlagsprov. (Factor analysis of Härnqvist´s ability tests.) Nordisk Psykologi, 1965, 17, 11-20.

Härnqvist, K. Individuella differenser och skoldifferentiering. (Individual differences and school differentiation.) Statens Offentliga Utredningar 13. Stockholm: Ecklesiastikdepartementet, 1960.

Härnqvist, K. Manual till DBA. (Manual for DBA.) Skandinaviska testförlaget, 1962.

Jöreskog, K.G. Statistical analysis of sets of congeneric tests. Psychometrika, 1971, 36, 109-133. (a)

Jöreskog, K.G. Simultaneous factor analysis in several populations. Psychometrika, 1971, 36, 409-426. (b)

Jöreskog, K.G. and van Thillo, M. New rapid algorithms for factor analysis by unweighted least squares, generalized least squares and maximum likelihood. Research memorandum 71-5 Princeton, N.J.: Educational Testing Service, 1971.

Jöreskog, K.G. and van Thillo, M. LISREL - A general computer program for estimating a linear structural equation system involving multiple indicators of unmeasured variables. Research Report 73-5, Department of Statistics, University of Uppsala, 1973.

Ljung, B.O. The adolescent spurt in mental growth. Stockholm: Almqvist & Wiksell, 1965.

Magnusson, D., Dunér, A. and Zetterblom, G. Adjustment - a longitudinal study. Will be published by Almqvist & Wiksell, Stockholm in 1973.

Reichard, S. Mental organization and age level. Archieves of Psychology, 1944, 41, (whole No 295).

Reinert, G. Comparative factor analysis studies of intelligence throughout the human life-span. In L.R. Goulet and P.B. Baltes (Eds.), Life-span developmental psychology. New York: Academic Press, 1970, 467-484.

van Thillo, M. and Jöreskog, K.G. SIFASP- A general computer program for simultaneous factor analysis in several populations. Research Bulletin, 70-62. Princeton, N.J.: Educational Testing Service, 1970.

Tucker, L.R. and Lewis, C. A reliability coefficient for maximum likelihood factor analysis. Psychometrika, 1973, 38, 1-10.

Appendix 1. The correlation matrix for boys. Decimal points omitted.

	1	2	3	4	5	6	7	8	9	10
1	1									
2	737	1								
3	743	724	1							
4	447	460	585	1						
5	346	347	483	449	1					
6	473	465	680	568	453	1				
7	193	210	241	256	338	337	1			
8	323	229	403	369	476	468	486	1		
9	799	726	751	417	370	492	197	356	1	
10	740	702	733	385	342	464	198	315	835	1
11	742	716	824	517	450	619	275	390	804	780
12	416	429	562	577	497	575	281	375	491	524
13	404	322	466	393	515	461	377	471	423	461
14	535	541	680	545	541	777	377	534	592	575
15	232	209	315	290	374	402	597	460	267	302
16	298	288	346	328	390	397	456	724	358	346

	11	12	13	14	15	16
11	1					
12	592	1				
13	466	574	1			
14	694	616	555	1		
15	329	415	462	452	1	
16	387	385	457	524	478	1

Note - The matrix includes the variables in the following order:

Tests 1 - 8: Similarities, Opposites, Swedish, Letter groups, Figure sequences, Cube counting, Metal folding, all tests in 1965.

Tests 9 - 16: The same tests applied in 1968.

Appendix 2. The correlation matrix for girls. Decimal points omitted.

	1	2	3	4	5	6	7	8	9	10
1	1									
2	761	1								
3	723	718	1							
4	323	370	501	1						
5	310	326	464	444	1					
6	461	491	639	552	505	1				
7	244	291	355	354	385	347	1			
8	279	304	422	408	429	476	461	1		
9	737	650	620	261	260	394	202	281	1	
10	716	680	635	309	278	415	251	265	769	1
11	702	710	771	358	345	534	306	358	763	755
12	361	414	519	567	465	514	313	396	396	450
13	319	341	476	414	505	459	328	440	349	414
14	489	532	660	528	502	726	392	483	495	540
15	258	273	371	403	367	484	546	486	288	331
16	265	283	403	422	406	457	432	631	283	301

	11	12	13	14	1	16
11	1					
12	532	1				
13	461	603	1			
14	657	605	570	1		
15	346	444	422	504	1	
16	359	408	532	539	520	1

Note - The matrix includes the variables in the following order:

Tests 1 - 8: Similarities, Opposites, Swedish, Letter groups, Figure sequences, Cube counting, Metal folding, all tests in 1965.

Tests 9 - 16: The same tests applied in 1968.

Örebroprojektets skriftserie

1. Olofsson, B. Vad var det vi sa! Om kriminellt och konformt beteende bland skolpojkar. Stockholm: Utbildningsförlaget, 1971.

2. Dunér, A. Vad skall det bliva? Undersökningar om Studie- och yrkesvalsprocessen. Stockholm: Allmänna förlaget, 1972.

3. Magnusson, D., Dunér, A. & Zetterblom, G. Adjustment - a longitudinal study. The Örebro project. Stockholm: Almqvist & Wiksell, under tryckning.

Huvudrapporter

I Planläggning
David Magnusson - Anders Dunér - Rolf Beckne
(Mars 1965)

II Sammanställning av tidigare forskning
David Magnusson - Rolf Beckne
(Juni 1967)

III Metoder och modeller
David Magnusson - Anders Dunér
(Juni 1967)

IV Datainsamling och bakgrundsvariabler
David Magnusson - Anders Dunér - Rolf Beckne
(Juni 1967)

V Undersökningsvariabler
David Magnusson - Anders Dunér - Rolf Beckne
(Juni 1967)

VI Några undersökningar med utgångspunkt från bakgrundsvariabler.
David Magnusson - Anders Dunér - Rolf Beckne
(Juni 1967)

VII Planläggning av projektets andra etapp
David Magnusson - Anders Dunér - Göran Zetterblom
(Juni 1967)

VIII Studie- och yrkesvalet: Modeller och undersökningsplanering.
David Magnusson - Anders Dunér - Göran Zetterblom
(Maj 1968)

IX Kriminellt beteende: Modeller och undersökningsplanering.
David Magnusson - Anders Dunér - Birgitta Olofsson.
(Juli 1968)

X Undersökning i projektets andra etapp
David Magnusson - Anders Dunér - Göran Zetterblom
(Augusti 1968)

Delstudier

1. En analys av data från elevenkäter
Rolf Beckne
(December 1969)

2. Yrkesenkät, intresseschema och yrkesvalslärares bedömning i årskurs 8
Dorothy Been - Göran Zetterblom
(Maj 1967)

3. Sociala relationer och skolprestation
Inger Josephson
(Oktober 1967)

4. Attityder och värderingar hos skolbarn. En undersökning med semantisk differentialteknik
Karin Nordenstam
(Mars 1969)

5. Självdeklarerad brottslighet bland pojkar i grundskolans årskurs 9
Birgitta Olofsson
(Mars 1969)

6. Sociala relationer i skolan. Del I. Undersökningens uppläggning
Berit Adebäck
(September 1969)

7. Sociala relationer i skolan. Del II. Hemmiljö och familjerelationer
Berit Adebäck
(September 1969)

8. Självvärdering och beteende
Stig Gagnerud
(Oktober 1969)

9. Katekolaminutsöndring och beteende i skolmiljö
Gunn Johansson
(Juni 1970)

10. Social differentiering i skolan. En empirisk studie vid Örebro grundskolor
Anna Nygren
(December 1969)

11 Effekter av klassens sociala sammansättning på elevgruppers prestationsnivå och motivation.
Anna Nygren
(December 1969)

12. Some univariate methods for making inferences about change.
Lars R. Bergman
(Januari 1971)

13. Brottslighet - konformitet. En utvecklingsstudie på grundval av självdeklarerad brottslighet av pojkar i grundskolan
Birgitta Olofsson
(Januari 1971)

14. Tonåringars normer och normklimat
Marta Henricson
(Maj 1971)

15. Symtom eller ålderstypiskt beteende? En studie av 15-åriga flickor
Karin Crafoord
(April 1972)

16. Kreativitet och högre studier I: Bakgrund och undersökningsvariabler
Lennart Elg
(Juni 1972)

17. Kreativitet och högre studier II: Val av utbildning efter gymnasiet -- deskriptiva data
Lennart Elg
(Juni 1972)

18. Studieavbrott i gymnasieskolan
Rolf Beckne
(Februari 1973)

19. Ungdomars postgymnasiala studie- och yrkesval
Lars R Bergman - Margareta Berggren -
Anders Dunér - David Magnusson
(Oktober 1973)

20. Studieavbrott i gymnasieskolan. Rapport nr 2.
Gymnasieelevers syn på sin skola
Rolf Beckne
(Oktober 1973)

21. A structural model for testing the age-differentiation hypothesis
Ulf Olsson - Lars R Bergman
(November 1973)

Övriga rapporter

Bergman, L. R. Some univariate models in studying change. Rep. Psychol. Lab. Univer. Stockholm, 1971. Suppl. 10.

Bergman, L. R. Parents' education and mean change in intelligence. Rep. Psychol. Lab. Univer. Stockholm, 1972. No. 350.

Bergman, L. R. Linear transformations and the study of change. Rep. Psychol. Lab. Univer. Stockholm, 1972. No. 352.