International Education Journal Vol 2, No 4, 2001
Educational Research Conference 2001 Special Issue
http://www.flinders.edu.au/education/iej

Schooling and intelligence: Effects of track of study on level and profile of cognitive abilities

Jan-Eric Gustafsson
Goteborg University *jan-eric.gustafsson@ped.gu.se*

Paper presented at the Spearman Conference, Sydney, 2001

The main purpose of the paper is to investigate effects of schooling on intelligence, with a particular emphasis on two questions: (1) is there a causal effect of schooling on intelligence, and if so how strong is the effect, and (2) what aspects of intelligence are being influenced? Reviews of the empirical research on the effect of schooling on intelligence are presented, and it is concluded that the research indicates that there is an effect of around 2 IQ points per year of additional schooling. Some interpret the effect of schooling on intelligence as an improvement in test performance only, there being no change in intelligence, while others view the change as a real change in intelligence. In order to obtain a better understanding of the nature of the change in intelligence as a function of schooling it is proposed that research should leave the notion of an undifferentiated concept of intelligence in favor of a multidimensional conception of intelligence. It is also proposed that rather than studying effects of an undifferentiated amount of schooling effects of different types of curricula on different aspects of intelligence should be investigated. An empirical study is presented in which changes in intelligence during two years of study on different tracks in upper secondary education are investigated for a group of some 14 000 males. At enlistment to military service at age 18 a cognitive test battery measuring Fluid ability (Gf), Crystallized intelligence (Gc) and General visualization (Gv) ability was administered. A latent variable model with five factors was fitted to the grades in the leaving certificate from grade 9 of compulsory school, and the five factors were used to control for entry differences to the tracks of upper secondary school. Effects corresponding to an improvement of about 3 IQ points are estimated for academic tracks. The academic tracks with technical and science orientation cause at least as strong an improvement in Gv, while for Gc weaker positive effects were obtained for the academic tracks and some of the vocational tracks. It is concluded that certain schooling experiences may cause improvements both in general cognitive ability, and in specific abilities.

INTRODUCTION

The field of research on intelligence is rich in paradoxes. There is, thus, ample evidence of stability of individual differences in intelligence over long periods of time, extending even from childhood to old age(e. g., Deary et al., 2000).Assumptions of stability are also made in mnay practical applications where it has proven useful to treat results obtained on tests of intelligence as representing fairly fixed characteristics of individuals that may be used for purposes of prediction, selection and diagnosis. But there also is evidence that intelligence can change, and that these changes may be quite dramatic. The best example of this is the so called secular change of intelligence, according to which there is a strong increase in the intelligence level of the population over time. This has been observed to take place since at least the 1930s, and since then the level of intelligence has risen by about one standard deviation (Flynn, 1984, 1987, 1999; Neisser, 1998). While no one challenges the empirical finding of the secular trend there is little agreement as to the explanation and meaning of the finding (Neisser, 1998). One of the many explanations that have been put forward is that the secular trend is due to the increased general level of education, and there is evidence that attending education may affect individual level of intelligence in a positive way (e. g., Husén & Tuijnman, 1991).

The possibility that education may affect intelligence is, however, an idea that has generated much controversy. The discussion following the publication of Jensen’s (1969) paper on how much intelligence may be boosted through preschool participation is an example of this. Reviews of research by, among others, Ceci (1991), Herrnstein and Murray (1994), and Winship and Korenman (1997) also arrive at quite different conclusions about to which degree schooling affects intelligence. One of the reasons for the divergence in opinions among reseachers is that the problem of determining the degree of effect of quality and quantity of schooling offers great methodological challenges.

Let me illustrate this by looking at the design and results of one of the classical studies in this field, namely the study by Härnqvist (1968a, 1968b). In this longitudinal study Harnqvist tested a representative 10% sample of the Swedish population of 13-year olds with a test battery comprising a verbal, a spatial and an inductive test. At the age of 18, the male subset of the sample took another test battery of a similar composition at the enlistment to military services. Harnqvist also had access to information about the educational careers of the subjects, which were differentiated into different tracks both in secondary school (grades 7-9) and in upper secondary school (grades 10-12; subjects still in school were tested either during the second semester of grade 11 or during the first semester of grade 12). Some subjects left school for work after grade 7, and quite a few did so after grade 9. In the analysis of the data Härnqvist used the test results at age 13 to partial out differences in entry characteristics to the different tracks of education, in order to determine treatment effects on intelligence of the differing kinds and amounts of schooling. In order to obtain more correct estimates of treatment effects in the covariance analytic design, Härnqvist devised a method for correcting the intelligence measures at age 13 for unreliability. A main finding was that students who had the most academically oriented education gained approximately .6 standard deviation units in intelligence as compared to those with the least amount of academic education. This is a fairly strong treatment effect, which exceeds the level of 0.5 stipulated for a medium effect size. Estimated on a yearly basis the effect is approximately 0.12 sd units per year, or 1.9 IQ points.

This empirical result raises, however, a number of questions concerning validity and interpretation. The most fundamental question is, of course, the one about internal validity: is the conclusion about a causal effect correct, or may this conclusion be challenged? Since this is a non-experimental study in which subjects were not randomly assigned to treatments there is the possibility that selection effects are an alternative explanation of the findings. This alternative explanation was stressed by Brody (1992), who said:

> Individuals who chose or were assigned to an academic track might have gained in IQ even if they had been randomly assigned to a less rigorous academic education. Consider two individuals with the same IQ who elect to enter different educational tracks at the secondary school level. The student who chooses the academic track may like to read books more than the individual who chooses a less academic track. Differences in intellectual interests may be related to changes in IQ.

This objection certainly is valid in principle, as it raises all the fundamental problems which are related to making causal inferences from non-experimental data. As was observed by Lord (1963) there is in such situations no dependable method for determining what is the appropriate adjustment of initial differences between groups. The only way out of this situation seems to be the recommendation made by Campbell (1963) to: "...vigorously attend to the specific plausible rival hypotheses appropriate to each situation" (p. 214). I will return to the evidence available concerning the validity of the alternative hypothesis put forward by Brody (1992).

If the Härnqvist finding is, temporarily at least, accepted to be internally valid, there is still the question of interpretation of what the finding means. Ceci (1999) argued, in essence, that there is a change of 1.8 IQ points for each additional year of academic schooling (which estimate, incidentally, agrees quite well with the estimate of 1.9 IQ points obtained by Härnqvist), but he also emphasized that the change is a change in intelligence test scores and not in intelligence: "... while schooling seems to prop up IQ test performance, it does not seem as obvious that it increases intellectual development, particularly if we conceptualise the latter in terms of novel problem solving" (Ceci, 1999, p. 171). This interpretation may be a reasonable one, but Ceci does not supply any empirical support for it. The question what the change in observed test performance signifies thus seems to be an open one.

The main purpose of the present chapter is to investigate effects of schooling on intelligence, with a particular focus on the two questions identified in relation to the Harnqvist example: (1) if it is possible to claim that there is a causal effect of schooling on intelligence, and in that case how strong the effect is, and (2) what aspects of intelligence are being influenced?

REVIEWS OF STUDIES OF EFFECTS OF SCHOOLING ON INTELLIGENCE

As has already been mentioned, several reviews of the literature on effects of schooling on intelligence have been published, and since the literature is very large, I will first review the reviews.

One of the fundamental methodological problems in research which aims to determine effects of schooling on intelligence is that strong designs which involve randomization of subjects are out of the question for practical and ethical reasons. This makes it difficult to establish casual relations, but it has spurred researchers into being creative to find other ways to establish the effects of schooling on intelligence. Ceci (1991) identified some 200 studies with eight different types of designs, and even though it is not possible to discuss all these here, it may be worthwhile to bring up some of the main designs and findings.

One class of evidence concerning the effect of schooling on intelligence is that during the summer vacation there is a small decrement in IQ, which effect seems to be stronger for children of low socio-economic status (e. g., Jencks et al., 1972). Another, similar, type of studies has investigated effects of intermittent school attendance on intelligence. In this group, Ceci (1991) refers to several early studies of children who were prevented from attending school regularly because of lack of travelling, handicap, or living in areas without access to education. These studies demonstrate a successive drop in IQ as a function of age for these groups of children. Ceci also refers to other studies showing cumulative deficits in IQ as a function of how much of school is missed.

Another class of studies investigates effects on intelligence of delayed onset of school. An example is a Dutch study (DeGroot, 1951) of the effects of schools being closed during World War II. This study estimated that when school onset was delayed several years, IQ dropped approximately 7 points. Ceci concluded that studies of effects of missed or delayed schooling show that these factors cause consistently large negative effects on IQ.

Ceci also reviewed studies investigating length of academic schooling. One example of such a study is the Härnqvist study which has already been described. There also are other such studies, several of which are Scandinavian. One is the Husen (1951; se also Husén & Tuijnman, 1991) study, in which 613 Swedish men were tested at the age of 10 in 1938, and then again at the induction to military service 10 years later. This study showed that there is an effect of completing secondary school (junior high school) and an even stronger effect of completing upper secondary school (high school). Lund and Thrane (1983) gave a large sample (N = 7 703) of Norwegian school children in grade 7 a set of military enlistment tests. Five years later the boys got the same tests at the military enlistment, leaving a sample of 2 485 individuals who had test results at both testing occasions. Lund and Thrane used several different methods to analyze the data, and among them were similar analys of covariance techniques as were used by Härnqvist. With correction for attenuation in the pretest scores they arrived at estimates of increases of from 2.5 to 2.8 IQ points per year of education. Lorge (1945) in an early study also used a longitudinal design to study effects of amount of schooling on intelligence, and concluded that there were substantial school-related increases in intelligence.

Yet another category of studies reviewed by Ceci (1991) argue that there are effects of schooling on intelligence, because the pattern of relations between schooling variables and achievement test scores are similar to the pattern of relations between schooling variables and intelligence scores. This suggests that IQ scores are just as influenced by schooling as is academic achievement.

Ceci (1991) argued that one of the strongest designs for investigating effects of schooling on intelligence is a cohort-sequential design in which children of the same age enter school at different times (i. e., they go in different grades), and stay there for similar lengths of time. One such a study was reported by Cahan and Cohen (1989) who investigated the slope of within-grade regressions of intelligence test scores on chronological age, comparing the regressions across grades to determine the impact of years in school. They found that the effect of schooling was substantial, and, indeed, about twice as strong as the effect of chronological age.

Ceci (1991) observed that for each of the different categories of evidence reported, there is evidence that schooling affects intelligence. He also noted that it would be possible to challenge this as a causal conclusion in virtually every study, because they are all basically correlational. However:

> ... despite the many interpretive snarls one confronts with correlational data, when one considers the entire corpus of correlations that have been reported this century; the high correlations between IQ and schooling are difficult to account for on the basis of genetic selection or any other explanation (e. g., motivational differences or parental SES), because these mechanisms appear farfetched in many of the studies that were reviewed. The most parsimonious account of the correlations that have been reviwed is that of a direct causal link, namely, that the processes associated with schooling influence performance on IQ tests through a combination of direct instruction and indirect inculcation of modes of cognizing ... (Ceci, 1991, p. 711).

On the basis of a review of the literature and an empirical study, Herrnstein and Murray (1994, Chapter 17) arrived at quite a different conclusion, namely that there is little evidence that differences in the amount of schooling accounts for much of the intellectual variation. They did, though, observe that a large scale systematic attempt to raise the intelligence in Venezuela and long-term coaching on tests like the Scholastoc Aptitude Test can have an effect of the order of 3 IQ points. Herrnstein and Murray (1994, Appendix 3) also reported an empirical study of their own, in which they analyzed changes in the Armed Forces Qualification Test (AFQT; a subset of the Armed Services Vocational Aptitude Battery) for a subset of 1408 cases in a longitudinal study (NLSY), for which results were available on one or more previously administered IQ tests. The results showed an increase of about 1 IQ point per year of additional schooling. This is a positive effect, but the estimate is lower than what has been obtained in other similar studies.

Winship and Korenman (1997) reanalyzed the Herrnstein and Murray (1994) data, correcting technical problems with the data (e. g., missing data problems, and the treatment of individuals entered multiple times) and using alternative model specifications (e g., comparing models with different assumptions about the reliability of the independent variables). The reanalysis resulted in higher estimates of the effects of schooling on intelligence: according to the model preferrred by Winship and Korenman (1997, p. 231) there was an estimated effect of 2.7 points of IQ per year of education.

Winship and Korenman (1997) also reviewed the research on effects of schooling on intelligence, with a special emphasis on studies using an analysis of covariance design. They tried to obtain a quantitative estimate of the magnitude of effects of schooling from each study. These estimates covered a quite large interval, the lowest being around 1 IQ point per year of schooling and the highest being around 4 IQ points per year of schooling.

If more emphasis is placed on the Winship and Korenman (1997) reanalysis of the Herrnstein and Murray (1994) study than on the original analysis, these reviews do seem to agree that there are positive effects of at least a couple of IQ points for each year of additional schooling.

Ceci (1991) brought up possible mechansisms that may account for the effects of schooling on intelligence, but he also observed that the empirical results are often too coarse to allow a more detailed analysis. Different types of schooling practices may thus be expected to influence some types of IQ tests more than others, but studies have typically not investigated effects on tests measuring different aspects of intelligence.

Ceci (1991) also pitted two possible explanations against one another. According to one of these schooling influences intelligence because the experience of being in school alters the individuals' cognitive processes in a fundamental manner, such as through fostering more abstract and disembedded ways of thinking about the world. According to the other, schooling only supplies the individuals with knowledge relevant for responding to IQ tests, and efficient ways of responding to items in such tests. Evaluating the support for these hypotheses, Ceci seems to favour the latter one. He emphasized that schooling involves teaching answers to questions on IQ tests, and that it is in school that one is most likely to come across information relevant to IQ tests. Ceci also observed that schools teach modes of thinking that are rewarded on intelligence tests, such as using paradigmatic classifications. As a third explanation, Ceci emphasized that schools prepare students with values that support efficient test-taking.

In the absence of more specific information about how different types of schooling affect different aspects of intelligence, it does seem, however, that conclusions about the mechanisms involved are premature.

There is also another reason why it does seem necessary to focus more on effects of different types of education than has been done in previous research. This is because education nowadays cover a much greater part of life for most individuals, at the same time as the concept of education, particularly at upper secondary and tertiary levels has been broadened to include much more of vocational education.

Commenting upon a finding by Flynn (1987) that only 5 % of the Dutch population increase in intelligence between 1952 and 1972 was due to increases in schooling, Ceci (1991) observed that " ... perhaps dropping out of high school and going into work-study apprenticeships during the early half of this century has been replaced by completing high school on a nonacademic track (i. e., taking relatively little academic course work). There has been an enormous growth of occupational and alternative education during the 1960s and 1970s, and this may have resulted in many students attending school even though little academic work was accomplished, whereas in former times these same students would not have remained in school.

From the literature reviwed so far it thus seems that we may draw the general conclusion that schooling has an effect on intelligence. However, several concerns which need further attention have also been identified. There is, thus, a need for better control of factors which are correlated both with the independent variable and the residuals of the dependent variable, and which therefore are threats to the validity of the inference that it is schooling that causes improvements in intelligence. One example of such an omitted variable may be differences in intellectual interests, which was proposed by Brody (1992) as a possible alternative explanatory factor in the Härnqvist (1968) study. Another factor which needs closer attention is type of education. Ceci (1991) suggested that only academically oriented types of programs cause effects on intelligence, and it certainly is a worthwhile question to ask to what extent there are differences in the effects of different types of programs. A third aspect which needs refinement is the conceptualization and measurement of the construct of intelligence. The large majority of studies have relied on global and undifferentiated IQ measures, but considerably more information may be gained if instead a model of intelligence is adopted which allows for multidimensionality, such as a hierarchical model (Carroll, 1993; Gustafsson, 1984, 1988, in press-a). Particularly when effects of different types of educational programs are studied it would seem necessary to adopt a more multifacetted conception of intelligence.

There is, indeed, a largely unnoticed study of effects of schooling on intelligence which improves the design in all these three respects. This is a Swedish study by Balke-Aurell (1982), which is described in somewhat greater detail below.

The Balke-Aurell Study

The study is a follow-up and extension of the Härnqvist (1968) study, and it partly relies on data from the same large-scale longitudinal study. While the main purpose of the Härnqvist study was to investigate changes in the general factor of intelligence, the main purpose of the Balke-Aurell study was to investigate changes in specific ability factors (primarily verbal/ reasoning and spatial/technical abiliies). In order to study the influence of different types of educations and occupations on the pattern of abilities, Balke-Aurell classified the educations and occupations into verbal and spatial/technical domains, and used these classifications as independent variables.

The study comprised two cohorts of data, one representative 10 per cent sample of males from the Swedish population born in 1948, and one born in 1953. The reason why the study is restricted to males is that it uses information about test results from the military enrollment at age 18. The number of cases from the 1948 cohort with complete data was 4 443 (82.6% of the originally drawn sample). For the 1953 cohort there were 3 847 cases with complete data (80.8% of the originally drawn sample).

The first wave of data collection took place when subjects were 13 years old, and typically attended grade 6. At this time a test battery with three tests (a number series test measuring inductive ability, a vocabulary test measuring verbal ability, and a spatial visualization test measuring spatial ability) was administered along with a questionnaire, asking, among other things, about preferred and actual leisure-time activities, attitudes toward school, and future plans for education and occupation.

Another wave of data collection took place when the subjects were tested at age 18 at the enlistment to military service. At this time a test battery with four tests was administered (an instructions test measuring verbal ability, a concepts test measuring verbal reasoning, a form-board test measuring spatial ability, and a mechanical comprehension test measuring spatial-technical ability). In order to form maximally comparable composites of scores from the two batteries Härnqvist (1968) used canonical correlation analysis to form one g-factor composite (g) from each battery, along with a bipolar factor contrasting verbal/reasoning ability with spatial/technical ability (v-s). Balke-Aurell used the same procedure, but she also treated the test batteries as measuring two correlated factors, one verbal (V) and one technical-spatial (T/ S).

Information about educational and occupational experience was also collected at age 18. Balke-Aurell categorized type of educational experience into three categories (educational structure): lines of study dominated by verbal subjects, those dominated by spatial/technical subjects, and a residual group not dominated by any of these. Educational experience also was classified into four educational levels: compulsory school only, practical-vocational education, lower secondary education, and upper secondary education.

Occupational experience was classified into five categories. One category comprised occupations with mostly verbal tasks (V). There were three categories of occupations making different kinds of demands of a spatial-technical nature: occupations involving spatially oriented tasks (T1); occupations involving tasks with demands on comprehension of technical-mechanical matters (T2), and occupations with tasks with demands on independent decisions concerning technical-mechanical problems (T3). The fifth occupational category did not pose any of these demands.

Balke-Aurell analysed the data by regression analysis and by fitting path models with latent variables. One analysis replicated the Härnqvist analysis on data from the 1953 cohort, with highly similar results. For the 1948 cohort the effect was .61 sd unit for the difference between those with the highest level of education as compared to those with the lowest level of education, while for the 1953 cohort the corresponding effect was .54 sd units. Balke-Aurell (1982, p. 105) suggested that the slight decrease of the effect may be due to the fact that most of the boys of the 1953 cohort attended the newly introduced comprehensive school with a less streamed curriculum than the school system attended by those born in 1948.

Changes in the v-s factor at age 18 for the different educational and occupational factors were also estimated, using the v-s factor at age 13 as control variable. These analyses revealed that for occupations at all educational levels there were quite consistent changes in the expected directions for both cohorts. The V group changed towards the verbal end of the factor, while T1, T2, and T3, generally in that order, changed towards the spatial end of the bipolar factor. The overall effect, expressed in sd units was .34 for the 1948 cohort, and .42 for the 1953 cohort. There was, however, a tendency for the effect to be stronger at higher educational levels.

For those in education until age 18 there also were changes in the expected directions for the three groups. The overall effect was .38 sd units for the 1948 cohort and .26 for the 1953 cohort. Balke-Aurell cautioned that it is difficult to compare the two cohorts because of changes of the Swedish educational system during these 5 years which makes it difficult to fit the educations into the same coding system.

Balke-Aurell also conducted multiple regression analyses in which a wide range of variables from the first wave of measurement was included as additional control of self-selection effects. Among these were school marks, social class, municipality, interests and ambitions, and social relations. Inclusion of these variables caused considerable increases in the amount of explained variance, but the estimates of treatment effects generally were unaffected. For the students the effect for the 1948 cohort increased slightly, from .38 to .39, while for workers the effect dropped from .34 to .28.

It was noted, however, that several of the variables introduced into the multiple regression analysis had low reliability, which causes bias in the estimates. The correction introduced and used by Härnqvist (1968a) could not be applied in this situation. This was one reason why Balke-Aurell also performed a series of analyses using latent variable models. Another reason for fitting such models was that they allow specification of two correlated latent ability variables at age 18, which would make it possible to study not only whether there has been a relative change in the balance of a bipolar ability, but also which ability has been affected.

These models were path models in which Verbal-scholastic ability, Verbal interest, Spatial ability, Technical interest and Social background were independent variables, Educational level and Educational structure were intervening variables, and Verbal ability and Technical/ spatial ability at age 18 dependent variables. According to this model there was no effect of Educational level or of Educational structure on the Verbal ability factor. However, for the Technical/spatial factor there was a significant standardized partial regression of .07 for Educational level, and a coefficient of .08 for Educational structure.

A models also was fitted in which the g and v-s factor were used as dependent variables. This model showed the g-factor to be affected by Educational level (standardized regression coefficient .07), and the v-s factor to be affected by Educational structure (.11).

Balke-Aurell concluded that the higher the educational level, the stronger is the increase in general intelligence. She also concluded that specific ability factors develop in accordance with verbal and technical types of education, and, to a lesser extent, with type of occupation. These effects were mainly found, however, for spatial/technical ability, and to a lesser extent for verbal ability.

The Present Study

The Balke-Aurell study demonstrates that it is necessary to take into account the fact that education is not a homogeneous activity, but that the characteristics of different educational tracks are important for effects on intelligence. The study also shows that it is necessary to investigate differential effects of schooling on different aspects of intelligence. Below an empirical study will be presented which aims to extend the study of effects of different educational programs, on different aspects of intelligence.

Sweden has 9 years of compulsory education, and 3 years of voluntary upper secondary education (or high school). However, even though upper secondary education is not compulsory, about 90 % of the cohort participates in 12 years of education (OECD, 2001). This increase in participation rate compared to the cohorts studied by Härnqvist (1968a, 1968b) and Balke-Aurell (1982) is mainly due to an expansion of vocational education at the upper secondary level. This had started for the 1954 cohort, who went through upper secondary school in the early 1970s, but the great expansion occurred after that. Until the mid 1990s most vocational programs comprised two years of study, but after this time all upper secondary education programs in Sweden comprise three years of study, and have a higher level of ambition as to academic content.

The current study investigates the cohort born in 1976. Normally students born this year left compulsory school at the age of 16 (school starting age is 7 in Sweden) in 1992 and started upper secondary school in the autumn the same year. At this time the old system of 2-year vocational programs was still in operation, but in some parts of the country the new 3-year programs had already been implemented. This makes it possible to make some comparisons between the vocational programs of different types.

This study also takes advantage of the fact that a new military enlistment battery was introduced in 1994. During the latter part of this year the old test battery consisting of four paper and pencil tests, which were designed to measure the general factor only, was replaced by a computer administered test, consisting of 10 subtests designed to measure the factors Fluid intelligence (Gf), Crystallized intelligence (Gc) and General visualization (Gv) (Carlstedt 2000; Mårdberg & Carlstedt, 1998). The three factors are estimated within a hierarchical modeling framework, according to which the Gf-factor is identical with the g-factor (Gustafsson, 1984, 1988, in press-a) and Gc and Gv correspond to second-stratum factors of Carrolls (1993) model.

This new test battery provides much more information about different aspects of intelligence than the battery studied by Balke-Aurell. What is of special interest to investigate in the current study is, of course, if schooling affects the g-factor, or if it affects the factors at lower levels of the hierarchical model, as was hypothesized by Ceci (1999). It is also of great interest to investigate if programs of education with special curricular emphasis affect special abilities, in a continuation of the investigation conducted by Balke-Aurell (1992).

The empirical study relies on information from official registers only, while the Härnqvist and Balke-Aurell studies used data collected through tests and questionnaires as well. Register information can easily be obtained for the entire population, so this method of data collection brings the advantage that a large set of data is available for analysis. The disadvantage is that certain types of information are not available, such as test scores and information about interests. However, there is information about school achievement and socio-ecenomic status, so a rather high degree of control of intial differences may be achieved.

METHOD

The design of the empirical study is fairly simple. Differences in results on the military enlistment test battery at age 18 for students who have followed different tracks in upper secondary school are analyzed, controlling for initial differences in grades and socio-economic background. To the extent that the control for differences due to self-selection is successful this design makes it possible to make causal inferences about effects of track of study on level and profile of intelligence.

Subjects

The data from the present study are taken from a large data base consisting of everone born between 1972 and 1979 living in Sweden in 1996. The data base includes a large number of variables taken from different registers containing information about educational choice and educational achievement (see Gustafsson, Andersson & Hansen, 2000, for a description in Swedish). Because every person in Sweden has a unique identification number which is used in almost all registers it is possible to create longitudinal data bases by combining the register data at individual level. Most of the registers are kept by Statistics Sweden, where the combination of the different registers has been conducted, after which the data has been delivered to Göteborg University in anonymous form.

The data to be analyzed here are from the 1976 cohort. This is the only cohort for which information is available both about track of study at upper secondary school and results on the new version of the military enlistment test battery. This cohort also is interesting to study because it represents a large variety of upper secondary programs. As has already been mentioned a new version of the upper secondary school curriculum was successively implemented until the mid 1990s. In the 1996 cohort there thus are persons who have followed the older version of the upper secondary school curriculum, in which vocational lines comprise two years of study, and there are persons who have followed the new curriculum, in which vocational programs comprise three years of study, and have a higher level of amibition when it comes to academic subject matters.

Because enlistment to military service in Sweden is compulsory for males only, the analyses reported here will be restricted to this gender. This not only limits the generalizability of the findings to males only, but it also entails a restriction on which programs of study can be investigated, because in many instances choice of program is highly correlated with gender. This is an unfortunate but inescapable consequence of relying on data collected for other purposes than research.

The data base comprises 52 113 males born in 1976, while the analyses to be reported are based upon a subset of 13 903 cases. The main reason why so many cases are lost is that only 17 588 cases have a score on the 10 subtests of the test battery analyzed here, which is because the new enlistment test was operational only from August 1994 (Mårdberg & Carlstedt, 1998). Only few cases were not tested at all. Those who are not Swedish citizens are excluded from the military enlistment procedure, as are those who have been diagnosed to be mentally retarded. With these restrictions, there is little reason to believe that those taking the new enlistment test do not form a representative sample from the population.

Another reason why not all cases are analyzed is that some attend programs of study which are too small to be analyzed. Because of the successive implementation of the new curriculum for the upper secondary school there were during the years 1992 to 1994 an unusually large variety of upper secondary school tracks. Not only did the old and new curricula co-exist in the country, as a function of whether the local authorities had decided to adopt the new curriculum or not, but in some school districts an experimental pre-cursor to the new curriculum, with 3-year vocational programs, was in operation. These experimental programs are not so well defined as the others, so they have been left out of the analysis.

Information about the socio-economic background of the students is available from the 1990 Census. Parental occupation has been used to classify the background of the students according to the so called Socio-Economic Index, which here involves a categorization into one of eight groups (e. g., unskilled labor, positions requiring academic education). In the analyses group belongingness has been represented as a vector of dummy variables.

Treatments

The old version of the upper secondary school curriculum consisted of five academic lines (Liberal Arts, Economics, Social Science, Natural Science and Technology), which prepared for entry to university level education. These programs all comprised three years of study. There also was a rather large number of vocational programs, which all comprised two years of study, and which did not allow entry to university level education without further education at the upper secondary level. The new system consists of 17 national programs of study and a large number of local programs. These programs are all three years long, and they all provide formal access to university level education, even though the range of available educational programs varies highly. When the students born in 1976 entered upper secondary school the new system had been decided upon, and in some parts of the country implementation started before the official starting date (i. e., the academic year 1994/95). A smaller part of the cohort thus have their upper secondary education from one of the new programs. For most of the programs the number of students is, however, to small to allow analysis.

Table 1 presents the number of students born in 1976 who have graduated from different lines of the old upper secondary school curriculum, and for students who have graduated from some of the programs of the new upper secondary school curriculum. For the analyses only lines and programs have been included for which at least 100 students is available for analysis. This implies, regrettably, that the Liberal arts program, for which only 90 cases are available is excluded. However, when attempts were made to include this program in the models, they failed to converge, so it is necessary to exclude the Liberal arts program from the analyses.

For most programs the proportion of cases who have data on the variables to be analyzed is around 30 per cent. For students from the new programs the proportions tend to be somewhat higher. One explanation for this may be that the new programs were implemented earlier in urban areas, where also the proportion of students who have complete data for grades as well as the SEB is larger than in other areas of the country. About 25 per cent of the students who miss information about line of study have grades from compulsory school and results on the SEB. The reason why information is missing is that they have either not entered upper secondary school, or have entered but not completed their education at the age of 20. The reason why there is a larger attrition from this group than the others is that it includes a larger proportion of cases with a non-Swedish background, who are less likely to take the SEB, and who are more likely to miss one or more of the grades from compulsory school.

In the analyses to be reported here all the different lines and programs will be analyzed as different treatments, and comparisons will be made with the group of cases for which no information is available about their upper secondary education. To understand the nature of possible treatment effects it would be desirable to have a fairly detailed description of the amount of time spent on different subject matters. Given the large number of tracks this is not possible, but it may be useful to have a more detailed presentation of one example of each of the main categories of educations.

Table 1. Number of male students born 1976 from different lines and programs of study

	Graduated	Included	Percent
Two year lines			
Building and Construction Line	2346	695	29.6
Consumer Line	148		
Distribution and Clerical Line	1016	297	29.2
Electro-Telecommunications Line	2874	924	32.2
Food and Manufacturing Line	721	239	33.1
Horticultural Line	194		
Motor Engineering Line	1910	601	31.5
Music Line	181		
Nursing Line	135		
Operation and Maintenance Line	690	232	33.6
Social Line	1609	503	31.3
Processing Line	36		
Woodwork Line	343	120	35.0
Workshop Line	972	372	38.3
Three year lines			
Economics Line	3861	1136	29.4
Liberal Arts Line	311		
Natural Sciences Line	3458	1007	29.1
Social Sciences Line	3717	1101	29.6
Technology Line	5356	1551	29.0
Three year programs			
The Construction Programme	495	215	43.4
The Electrical Engineering Programme	1191	445	37.4
The Vehicle Engineering Programme	834	271	32.5
The Business and Administration Programme	961	309	32.2
The Industrial Programme	822	324	39.4
The Use of Natural Resources Programme	843	280	33.2
The Natural Science Programme	1425	488	34.2
The Social Science Programme	1913	634	33.1
No information	8404	2087	24.8

The Electro-Telecommunications line is the largest among the two-year lines. This line had during the two years 160 lessons (40 minutes) of Swedish and 240 lessons of an optional theoretical subject. For electro- telecommunications engineering 2360 lessons were alloted, divided between work technique and vocational theory, with a stronger emphasis on the former than on the latter. This curriculum thus has a weak emphasis on general academic subjects, and a strong emphasis on practical skills.

This curriculum may be compared with the new Electro- Telecommunications program. In the new upper secondary school the curriculum is based on courses, which are of three different kinds: core courses, which are taken by everyone, program specific courses, and individual choice courses. Each course has a number of credit points assigned to it, and currently a total of some 2 500 points is required for the three years. The core courses, almost all of which are in different academic subjects, comprise 750 points, and the program specific courses about the same number of points. Some 900 points may be taken as individually elective courses, which may be either general academic courses (e. g., mathematics, English) or program related courses (e. g., electronics, computer technology). In addition a project work must be completed. The proportion of academic courses thus is higher in the new program, even though the large number of elective courses makes it impossible to identify a single common curriculum. However, according to the recommended time table specified in 1997/98, 1370 hours (60 minutes) should be devoted to vocational subjects, while 680 hours should be devoted to core subjects. Thus, even though there is a stronger emphasis on academic subjects in this program than in the two-year program, the main emphasis still is on vocational skills.

For the three-year Natural Science line, to take yet another example, approximately half the time was allotted to mathematics, science subjects and technology, approximately 25 per cent of the time to Swedish and foreign languages, and the remaining time roughly evenly split between social subjects and other subjects. The main emphasis of this program thus is academic.

In the analysis each different track will be represented by a dummy variable of its own, using the No information group as the reference group.

Grades from compulsory school

As has already been pointed out there is a strong need to identify and control for possible self-selection of students into the different tracks. This will be done through relying on the grades in 17 different subject matters given on the leaving certificate from compulsory school. The grades were assigned on a scale from 1 to 5 according to a norm-referenced grading system, in which national tests were used to achieve comparability of grades from different schools.

Andersson (1998) has fitted a five-factor model to the grades data, which holds promise not only to capture differences in general level of school achievement for the different lines and programs of study, but also differences in the profile of performance over different areas. Andersson fitted her model to the population data for students born in 1972, and the standardized factor loadings for males are shown in Table 2, along with the results for the current sample.

Table 2. Standardized factor loadings in the five-factor model for grades

	SchAch		**Non-Verb**		**MathSci**		**Lang**		**Ad**	
	Males (72)	Current	Males (72)	Current	Males (72)	Current	Males (72)	Current	Males (72)	Current
Child studies	0.78	0.79							0.23	0.30
Art Education	0.49	0.50	0.25	0.25			0.12	0.12	0.14	0.12
English	0.73	0.73			0.10	0.12	0.42	0.41		
Domestic science	0.72	0.72	0.19	0.21					0.25	
Physical education	0.39	0.44	0.23	0.21						
Mathematics	0.75	0.75	0.14	0.16	0.29	0.28	0.12	0.14		
Music	0.63	0.62	0.10	0.12			0.19	0.19	0.13	0.12
Biology	0.85	0.84	0.08	0.11	0.24	0.27			0.11	0.11
Physics	0.83	0.82	0.15	0.17	0.34	0.36				
Chemistry	0.84	0.84	0.09	0.11	0.31	0.34			0.08	0.07
Technology	0.59	0.60	0.36	0.35	0.22	0.23				
Geography	0.88	0.89	0.03	0.02						
History	0.90	0.90							-0.09	-0.08
Religious studies	0.90	0.90							-0.02	-0.04
Civics	0.90	0.90								
Crafts	0.42	0.44	0.59	0.58					0.11	0.12
Swedish	0.83	0.84					0.25	0.24	0.04	0.07

The estimated factor loadings generally are very close. As may be seen from the structure of the estimated factor loadings, the model fitted is an orthogonal model with so called nested-factors (Gustafsson & Balke, 1993). There is a general factor (*SchAch*), which has positive and fairly strong relations to all the grades. The highest loadings are observed for social science subjects which require heavy reading and much home-work. Andersson (1998) interpreted this factor as involving a strong component of verbal ability, and also a strong motivational component.

There is also a rather broad factor which Andersson labeled NonVerb, and which is most highly related to technology, crafts, and physical education, but also to several other subjects which involve figures or numbers. Andersson (1998) interpreted this to be a spatial-practical factor. The third factor relates to grades in mathematics and in science subjects, and was labeled MathSci by Andersson (1998). The fourth factor has fairly strong relations to the grade in English and in Swedish, and weaker relations to grades in Art education, mathematics, and Music. This was interpreted as a language (Lang) factor by Andersson (1998). The fifth factor, finally, has low relations to a fairly broad range of grades, and it was labelled the aesthetic-domestic factor (Ad) by Andersson (1998).

These five factors will be used to control for possible selection and self-selection effects into the different lines and programs.

The Swedish Enlistment Battery

Until 1994 the enlistment battery consisted of four tests (one inductive, one verbal, one spatial and one technical; see Carlstedt, 2000), which where combined to measure a general factor. In August 1994 this batttery was replaced by a computer administered test, consisting of 10 (or rather 12, if two experimental tests are counted as well) subtests designed to measure the factors g, Gc and Gv within a hierarchical modeling framework (Carlstedt 2000; Mårdberg & Carlstedt, 1998). This test is referred to as the Swedish Enlistment Battery (SEB). The following tests are included:

- *Synonyms 1* and *Synonyms 2*, with 25 and 20 items, respectively. These are multiple-choice tests with four or five options, from which the synonym of a given word is to be selected. There is also a test called *Opposites* (25 items) in which the task is to select the antonym of a given word. These three tests have been shown to measure Crystallized Intelligence (Gc).
- *Figure Series* (20 items) presents sequences of four figures, and the task is to complete the series by selecting two figures out five given ones. In the *Groups* (20 items) test items five figures are presented, and the task is to identify the figure that does not fit thematically. These two tests have been shown to measure Fluid Intelligence (Gf).
- In the items in *Dice 1* and *Dice 2* (20 items each) two cubes, on which three surfaces are visible, are presented. According to the instructions there is a unique symbol on each side of the cube, and on identical cubes, the symbols are placed in the same relation to each other. The task is to find out if the two cubes, if turned, could be identical, or if they are different. In the *Metal folding* test (16 items) a drawing of an unfolded piece of metal is presented and the task is to find the three-dimensional object out of four which corresponds to the two-dimensional drawing. The items in the *Block rotation* test (20 items) present a three-dimensional target object, and the task is to select the identical three-dimensional rotated object out of five. In *Technical comprehension* (16 items) the items all constitute illustrated technical and physical problems, and one out of three suggested solutions should be selected as the correct one. These five tests have been shown to measure General visualization (Gv). The Technical comprehension test has a loading on Gc as well, and there also is overlap between the specific components of Dice 1 and Dice 2 (Mårdberg & Carlstedt, 1998).

Factor models have been fitted to this test battery in several previous studies (e. g., Carlstedt, in press; Carlstedt, Gustafsson, Ullstadius, 2000; Mårdberg & Carlstedt, 1998). These models have genereally been taken to be hierarchical models of the nested-factor kind (Gustafsson, in press-a; Gustafsson & Balke, 1993) in which a general factor factor (g) is related to every test and Gc’ and Gv’ are introduced as residual factors to account for the remaining covariance between these two groups of tests. The model does not include a residual Gf-factor, because of the empirical equivalence between g and Gf (Carroll, in press; Gustafsson, 1994, in press-a).

Such a model has been fitted to these data as well, and the standardized factor loadings are shown in Table 3.

Table 3. Standardized loadings of the tests in the SEB on the three factors of the NF-model

	g	Gv	Gc
Figure series	0.89		
Groups	0.80		
Dice 1	0.70	0.20	
Dice 2	0.68	0.29	
Metal folding	0.71	0.38	
Block rotation	0.69	0.34	
Technical comprehension	0.68	0.24	0.17
Synonyms 1	0.61		0.64
Synonyms 2	0.66		0.59
Antonyms	0.68		0.59

As may be seen there are substantial loadings of all the tests on the g-factor, and particularly so for the two Gf-tests. The tests hypothesized to load on Gv' do so, but the loadings tend to be relatively low, which indicates that it is difficult to separate the g- and Gv'-factors in this battery of tests. For Gc' the three vocabulary tests have quite substantial loadings. Being restricted to vocabulary tests only it may be noted that the Gc factor is a fairly narrow factor.

The model includes a covariance between the residuals of Dice 1 and Dice 2 as well, and when this covariance is introduced the model fit is quite acceptable, the RMSEA being 0.044, with a narrow confidence interval.

This hierarchical model will be used to estimate the g-factor. However, this model is less useful for investigating efffects of schooling on Gc and Gv, because g is partialled out from these dimensions, and g may be affected by treatment effects. For purposes of analysis of effects of schooling on the abilities at lower levels of the hierarchy an oblique measurement model will be used as well. This model specifes three correlated factor Gf, Gv, and Gc, with the pattern of standardized loadings shown in Table 4.

Table 4. Standardized loadings of the tests in the SEB on the three factors of the oblique model

	Gf	Gv	Gc
Figure series	0.90		
Groups	0.80		
Dice 1		0.74	
Dice 2		0.74	
Metal folding		0.79	
Block rotation		0.77	
Technical comprehension		0.56	0.23
Synonyms 1			0.87
Synonyms 2			0.89
Antonyms			0.91

In this model there is no g-factor, and the Gf-factor is measured by Figure series and Groups alone. Otherwise the pattern of loadings matches that of the hierarchical model. In the oblique model there are, however, high correlations among the factors. Gf and Gv thus correlate 0.90, Gc and Gv correlate 0.69, and Gf and Gc correlate 0.72. This model too fits well, with an RMSEA of 0.046.

RESULTS

The analysis and reporting of results proceeds in three steps. First relations between the factors of the model for the grades and the SEB-factors will be investigated. Next relations between tracks of study and the grades will be studied, in order to investigate the extent of selection and self-selection to the tracks. In the third step relations between lines of study and the factors of the SEB-models are investigated with control for grades and socio-economic status.

Relations among grades from compulsory school and test performance at 18

Unless there is a strong relationship between the grades from compulsory school and performance on the SEB the grades cannot be used to control for entry differences in level of performance among the different tracks. It is thus necessary to investigate the amount of relationship among these two sets of factors, which has been done by regressing, in two separate analyses, the g-factor and and the factor of the oblique SEB-models onto the five factors of the model for the grades. The standardized regression coefficients are presented in Table 5.

Table 5. Standardized relations among the latent variables for the grades and SEB models

	SchAch	Non-Verb	MathSci	Lang	Ad	Expl var
g	0.64	0.19	0.31	0.30	-0.09	64%
Gf	0.63	0.19	0.33	0.28	-0.09	60%
Gv	0.52	0.36	0.36	0.24	-0.12	60%
Gc	0.71	-0.09	0.22	0.36	-0.05	70%

The pattern of results are highly similar for the g- and the Gf-factors, and no less than 64% of the variance in g is accounted for. The general SchAch factor contributes most, but MathSci and Lang also have substantial relations to g. For Gv there is a strong positive correlation with SchAch and Lang, which is because there is a very substantial amount of g-variance in this factor, but this factor also has rather strong relations with NonVerb and MathSci. For Gc there is a very strong relation of .71 to SchAch, and there also are relations to Lang and MathSci.

Between 60% and 70% of the variance in abilities is accounted for by the grade factors. In spite of the fact that intelligence scores are not available the amount of variance accounted for is at least as high as in previous studies. This is because a latent variable modeling approach is used which takes full advantage of the multidimensionality of grades and SEB tests, and which allows estimation of error-free latent variables. The problem of disattenuation of relations by taking unreliability into account is thus not present here.

Tracks and grades

In the Swedish school system grades have traditionally served an important function as a tool for selection into upper secondary school from compulsory school. During the 1990s, however, little selection has been necessary because by and large study places have been available to match the demand. However, even though little explicit selection is made on the basis of grades, processes of self-selection may cause considerable differences in means of grades over the different tracks in upper secondary school.

In order to investigate such differences the five factors of the model for the grades have been regressed into the full set of dummy variables for the lines of study, and the partial regression coefficients have been estimated. This analysis expresses on a convenient scale, which is comparable across the different dimensions, track differences in the level of the five factors. As has been observed by Lubinski and Humphreys (1996) correlations are often misinterpreted to indicate weaker relations and smaller differences than other measures, such as effect size measures. However, they also demonstrate that correlations and measures which express mean differences can be translated into one another, so such misinterpretations can be avoided. Some simple rules of thumb are obtained from knowing that a small effect size (> .20) translates into a correlation of .10, that a medium effect size (> .50) corresponds to a correlation of .24, and that a large effect size (> .80) translates into a correlation of .37.

As may be seen in Table 6 choice of, and successful completion of, track accounts for no less than 55 % of the variance in SchAch. This is primarily due to the fact that the academically oriented three year lines have a much higher mean on the SchAch factor than have the vocationally oriented two-year lines. Among the new programs the Natural Science Program and the Social Science program also have a higher mean on the SchAch factor, but not as high as the old lines. The two-year Electro-Telecommunications line and the three-year Electrical Engineering program have a higher level on SchAch than the other vocational programs.

Table 6. Standardized coefficients for the differences in grades between lines

	SchAch	**Non-Verb**	**MathSci**	**Lang**	**Ad**
Two year lines					
Building and Construction Line	0.05	0.15	-0.01		0.04
Distribution and Clerical Line	0.04	0.01	-0.01		0.04
Electro-Telecommunications Line	0.15	0.19	0.12		0.08
Food and Manufacturing Line	0.06	0.03	0.00		0.09
Motor Engineering Line	-0.01	0.12	0.01		0.03
Operation and Maintenance Line	0.03	0.09	0.02		0.00
Social Line	0.10	0.02	-0.02		0.03
Woodwork Line	0.00	0.12	-0.01		0.02
Workshop Line	-0.07	0.04	-0.01		-0.02
Three year lines					
Economics Line	0.36	0.07	0.07		0.11
Natural Sciences Line	0.53	0.06	0.32		0.07
Social Sciences Line	0.43	0.02	0.02		0.05
Technology Line	0.50	0.21	0.32		0.16
Three year programs					
The Construction Programme	0.01	0.06	-0.03		0.03
The Electrical Engineering Programme	0.10	0.10	0.10		0.07
The Vehicle Engineering Programme	-0.02	0.08	0.03		0.01
The Business and Administration Programme	0.05	0.02	-0.03		0.03
The Industrial Programme	-0.03	0.06	0.00		-0.02
The Use of Natural Resources Programme	0.04	0.07	0.00		0.01
The Natural Science Programme	0.31	0.09	0.24		0.07
The Social Science Programme	0.25	0.02	0.05		0.05

In the MathSci factor 24 % of the variance is accounted for, and this is almost entirely due to the fact that the Natural Sciences line, the Technology line, and the Natural Sciences program have a high mean on this factor. Here too the Electro-Telecommunications line and the Electrical Engineering program have intermediate means.

The lines of study account for 8% of the variance in the NonVerb factor. The Technology line, the Electro-Telecommunications line, the Building and Construction line, the Motor Engineering line, and the Woodwork line all have correlations higher than .12.

For the Lang factor 11.6% of the variance is accounted for. The academically oriented lines and programs have a higher level on this factor. Had the Liberal Arts line been included in the analysis this line would have been seen to have a considerably higher mean on this factor.

For the Ad factor only 4% of the variance is accounted for, and there does not seem to be any clear pattern of differences among the lines with respect to this factor.

The results presented in Table 6 show that the different tracks have widely varying levels of entry performance, not only on the general school achievement factor, but also on the narrow achievement factors. It is quite interesting to see that these differences go along with the different contents and requirements of the lines. Thus, for the programs with an emphasis on mathematics and science the MathSci factor is high, and for programs in which spatial-practical skills are important the NonVerb factor is high. Unless these differences, which primarily are due to self-selection, are controlled for, they will be confounded with treatment effects.

Effects of track of study on test performance at age 18

In the final step of the analysis a model has been constructed in which the intelligence factors at age 18 have been regressed onto the dummy variables representing track and the latent variables representing individual differences in the compulsory school grades and the SES variables have been included as control variables. The model thus estimates the direct effects of line of study on test performance at age 18, controlling for entry differences.

The parameter estimates are presented in Table 7. Only significant ($p < .01$) parameter estimates are shown.

Table 7. Standardized coefficients for the differences the SEB factors between tracks, controlling for grades and ses

	g	Gf	Gv	Gc
Two year lines				
Building and Construction Line	-0.03	-0.03	-0.03	-0.03
Distribution and Clerical Line			-0.03	-0.03
Electro-Telecommunications Line	0.04	0.04	0.04	0.03
Food and Manufacturing Line				
Motor Engineering Line				
Operation and Maintenance Line	0.02	0.03		0.04
Social Line				0.02
Woodwork Line				
Workshop Line				
Three year lines				
Economics Line	0.09	0.10		
Natural Sciences Line	0.09	0.12	0.11	0.07
Social Sciences Line	0.07	0.07		0.08
Technology Line	0.14	0.17	0.17	0.10
Three year programs				
The Construction Programme				
The Electrical Engineering Programme	0.02	0.02		
The Vehicle Engineering Programme				
The Business and Administration Programme				-0.03
The Industrial Programme				
The Use of Natural Resources Programme				0.05
The Natural Science Programme	0.08	0.10	0.10	0.05
The Social Science Programme	0.06	0.07	0.03	0.03

For the g and Gf factors positive effects of attending an academic program may be observed. The strongest effect is obtained for the Technology program, but a standardized regression coefficient around .10 or slightly lower is observed for all the academic programs, including the new programs. There are also weak positive effects on g/Gf of attending the Electro-Telecommunications line and the Electrical Engineering program.

For Gv there are effects of the Natural Sciences line (0.11), the Technology line (0.17), and of the Natural Sciences program (0.10). The latter program may be regarded a combination of the Natural Sciences line and the Technology line.

For Gc there are effects of the academic lines and programs, with correlations varying between .03 and .11. There also are very weak effects, both positive and negative ones, of some of the vocational lines, but these do not seem to form an easily described pattern.

DISCUSSION AND CONCLUSIONS

The results indicate that there are effects on the factors of intelligence of schooling at the upper secondary level. However, the effects seem to be restricted to the academic programs.

The estimates for the general factor are marginally stronger when the Gf-factor in the oblique model is taken to represent this factor, than when it is estimated as a truly general factor in the NF-model. There also is some variation between the programs, the lowest effect being observed for the Social Sciences program and the strongest for the Technology program. However, the mean effect amounts to about .10, which translates into an effect size of 0.20 (Lubinski and Humphreys, 1996). Assuming an IQ scale with a standard deviation of 16 this implies a change of 3.2 IQ units during the approximately two years that the treatment lasted, or 1.6 IQ units per year. This is quite close to the estimated increase of 1.8 IQ units for each additional year of academic schooling reported by Ceci (1999), but it is at the lower end of the interval of estimates of 1 - 4 IQ points arrived at by Winship and Korenman (1997).

There is no sign that there would be any effect of the higher academic level of amibition of new 3-year vocational programs. However, it must be emphasized that this increased level of ambition has implied that the length of the programs have been extended from two to three year, while during each of the years the proportion of academic subject matter is still fairly small compared to vocational subjects. The present data was collected during the autumn semester of the third and last year of upper secondary school, which implies that for most subjects there remained almost a full year of education in the track. These data thus were collected to early to tell about the full effect of of this change.

For Gv there are fairly strong effects of the three academic programs which have an orientation towards technicology and science. For the Electro-Telecommunications line there is also a weak effect of .04, but otherwise there is no effect on this ability of the vocational tracks. Balke-Aurell (1982) found that both educations and occupations with a spatial/technical content caused a shift in the ability profile in the direction of the spatial/technical end of the bipolar v-s factor, but she also found that the effects were stronger at higher levels of education. The present results suggest that there is an effect on Gv only for the most academically oriented tracks.

For Gc there also are effects of the academic tracks, except for the Economics line, even though these effects tend to be somewhat weaker than for the other to factors of intelligence. There also are still weaker effects of some of the vocational tracks (Electro-Telecommunications line, Operation and Maintenance Line, Social line, Use of Natural Resources Program). While for Gv there seems to be a fairly strong effect for certain specific programs, the effects of Gc are weaker and associated with a wider range of programs. It must be observed, however, that the fact that the tracks under study here are those which are popular among males implies a restriction of which tracks are actually included in the analyses. The Liberal Arts program is an example of a program which too few males had selected to allow inclusion in the study. Had it been possible to include this program, a strong effect on Gc might have been found of this track. Had females been included in the study as well, a wider range of vocational tracks would also have been included in the study, some of which may have had a strong influence on Gc.

When interpreting the Gc findings it must also be borne in mind that the Gc-dimension in the present study is quite narrowly defined as a vocabulary factor. It seems, however, that it would have been difficult in the present study to adopt such a broad definition of Gc as to also include achievement in subject matter areas (Carroll, 1993), becuase then specific content taught in the different programs might have been misinterpreted as general effects on intelligence.

These results thus indicate that participation in an academic track at upper secondary school has a positive effect on Gf and, to a lesser extent, on Gc. Furthermore, academic tracks with an orientation towards technology and science affect Gv positively. This shows that even after the length of education has been extended so that the natural variation in the number of years of education up to age 18 has been eliminated, there are so large differences between the content and methods of different tracks that there are effects on intelligence which are closely similar to those obtained when variation in years of schooling were investigated.

One of the main questions in the research on effects of schooling on intelligence is how the change should be interpreted. One hypothesis is that the change is a real and profound one, which fully reflects the meaning of the construct of inytelligence. Another hypothesis is that the observed change only is a change in test performance without any concomitat change in the underlying ability. Ceci (1999) expressed the latter hypothesis as follows: "... while schooling seems to prop up IQ test performance, it does not seem as obvious that it increases intellectual development, particularly if we conceptualise the latter in terms of novel problem solving" (Ceci, 1999, p. 171). The results show quite clearly, however, that the strongest effect is observed for the Gf-factor, there being a marginally weaker effect on the g-factor which is estimated from the full set of tests in the battery. The Gf-factor is measured by tests of inductive reasoning and it clearly involves novel problem-solving, which provides support for the idea that the effects of schooling are not surface-level effects due, for example, to acquisition of specific pieces of information. One possible explanation for the positive effect on problem-solving abilities may be that the academic curricula to place an emphasis on problem solving rather than on knowledge acquisition.

The weaker effects that have been obtained for Gc should, however, probably be accounted for in terms of the increased opportunities in education to acquire knowledge about the measning of new words.

It seems, however, that the estimates of the size of effects of the current study are somewhat smaller than what has been obtained in previous Scandinavian research. One hypothesis to account for this is that the effects of schooling are diminishing with age. It seems that the highest estimates of the effects of one year of education have been obtained in studies of young students, such as the Cahan and Cohen (1989) study in which grades x to x were investigated. The present study only covers the last two years of the developmental period up to 18, while most previous study have covered a wider time span, so if the effects of education on intelligence are diminishing with age lower estimates would be expected in the present study.

It could, of course, be asked if the design of the current study allows causal inferences to be made. Given that this is a correlational study, with a less than perfect statistical control of initial differences, the results could be rejected as being expressions of mechanisms of self-selection, rather than treatment effects. With Ceci (1991) it could, however, be argued that "... in an inductive enterprise ... the usual inferential processes are relevant, namely that if studies using different methodologies, different mental measures, and different samples all converge on the same conclusion and if these studies possess uncorrelated weaknesses, then one can infer that their collective power is greater than that of their individual conclusions (Ceci, 1991, p. 718).

Compared to previous studies using the analysis of covariance design the present study employs a weaker set of measures to control for entry differences in that only school achievement and no IQ measure is being used. However, this is partially compensated for by employing a multidimensional latent variable model which fully utilizes the information available, and which uses latent independent variables in which there is no attentuation of relationships due to errors of measurement. The amount of variance accounted for in the dependent latent variables therefore rather seems to be higher (60 % to 70%) than in most previous studies.

REFERENCES

Andersson, A. (1998).The dimensionality of the leaving certificate in the Swedish compulsory school. *Scandinavian Journal of Educational Research*, *42*(1), 25-40.

Balke-Aurell, G. (1982). *Changes in ability as related to educational and occupational experience*. Goteborg, Sweden: Goteborg Studies in Educational Sciences.

Brody, N. (1992). *Intelligence (2nd ed)*. New York: Academic Press.

Carlstedt, B. *Cognitive abilities – aspects of structure, process and measurement*. Göteborg, Sweden: Acta Universitatis Gothoburgensis.

Carlstedt, B. (in press). Differentiation of cognitive abilities as a function of level of general intelligence. A latent variable approach. *Multivariate Behavioral Research*.

Carlstedt, B., Gustafsson, J.-E, & Ullstadius, E. (2000). Item sequencing effects in the measurement of fluid intelligence. *Intelligence*, *28*(2), 145-160.

Carroll, J. B. (1993). *Human cognitive abilities*. Cambridge: Cambridge University Press

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components?: A reassessment of the evidence. *Developmental Psychology*, *27*, 703-722.

Cahan, S., & Cohen, N. (1989). Age versus schooling effects on intelligence development. *Child Development*, *60*, 1239-1249.

Campbell, D. T. (1963). Factors relevant to the validity of experiments in social settings. In C. W. Harris (Ed.) *Problems in measuring change.* Madison: Wisconsin University Press.

Deary, I. J., Whalley, L. J., Lemmon, H., Crawford, J. R., & Starr, J. M. (2000). The stability of individual differences in mental ability from childhood to old age: follow-up of the 1932 Scottish Mental Survey. *Intelligence*, *28*. 49-55.

DeGroot, A. D. (1951). War and the intelligence of youth. *Journal of Abnormal and Social Psychology*, *46*, 596-597.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, *95*, 29-51.

Flynn, J. R. (1987). Massive IQ gains in 14 nations. What IQ tests really measure. *Psychological Bulletin*, *101*, 171-191.

Flynn, J. R. (1999). Searching for justice: the discovery of IQ gains over time. *American Psychologist*, *54*, 5-20.

Gustafsson, J. E. (1988). Hierarchical models of individual differences in cognitive abilities. In R. J. Sternberg, *Advances in the psychology of human intelligence* Vol. 4, (pp. 35-71). Hillsdale, NJ: Erlbaum.

Gustafsson, J. E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, *8*, 179-203.

Gustafsson, J.-E. (in press-a). On the hierarchical structure of ability and personality. To be included in a volume based on papers presented at the Second Spearman Seminar, Plymouth, 1997

Gustafsson, J.-E., & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research*, *28*(4), 407-434.

Harnqvist, K. (1968a). Relative changes in intelligence from 13-18. I. Background and methodology. *Scandinavian Journal of Psychology, 9*, 50-64.

Harnqvist, K. (1968b). Relative changes in intelligence from 13-18. II. Results. *Scandinavian Journal of Psychology, 9*, 50-64.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve. Intelligence and class structure in American life*. New York: The Free Press.

Husén, T. (1951). The influence of schooling upon IQ. *Theoria, 17*, 61-88.

Husén, T., & Tuijnman, A. (1991). The contribution of formal schooling to the increase in intellectual capital. *Educational Researcher, 20*(1), 17-25.

Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen D., Gintis, H., Heyns, B., & Mitchelson, S. (1972). *Inequality: A reassessment of the effects of family and schooling in America.* New York: Basic Books.

Lorge, L. L. (1945). Schooling makes a difference. *Teacher's College Record, 46*, 483-492.

Lund, T., & Thrane, V. C. (1983). Schooling and intelligence: A methodological and longitudinal study. *Scandinavian Journal of Psychology, 24*, 161-173.

Lubinski, D., & Humphreys, L. G. (1996). Seeing the forest from the trees: When predicting the behavior or status of groups, correlate means. *Psychology, Public Policy, and Law, 2*(2), 363-376.

Mårdberg, B., & Carlstedt, B. (1998). Swedish Enlistment Battery (SEB). Construct validity and latent variable estimation and profile prediction of cognitive abilities by the CAT-SEB. *International Journal of Selection and Assessment, 6*(2), 107-114.

Neisser, U., et al. (1996). Intelligence: knowns and unknowns. *American Psychologist, 51*, 77-101.

OECD (2001) *Education at a glance. OECD indicators 2001 edition.* Paris: OECD.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262-274.

Winship, C., & Korenman, S. (1997). Does staying in school make you smarter? The effect of education on IQ in *The Bell Curve*. In B. Devlin, S. E. Fienberg, D. P. Resnick, & K. Roeder (Eds.). *Intelligence, genes, & success. Scientists respond to the Bell Curve*. New York: Springer-Verlag, Inc.