Reports from the department of Education and Educational Research

EVALUATION THROUGH FOLLOW-UP

SELECTION OF SWEDISH SCHOLASTIC APTITUDE TEST TAKERS AS A FUNCTION OF SOCIOECONOMIC BACKGROUND AND ABILITY

Sven-Eric Reuterberg

S

Report no 1994:05
Department of Education and Educational Research
Göteborg University

SELECTION OF SWEDISH SCHOLASTIC APTITUDE TEST TAKERS AS A FUNCTION OF SOCIOECONOMIC BACKGROUND AND ABILITY

Sven-Eric Reuterberg

Evaluation Through Follow-up is a research program aiming at a continous evaluation of the Swedish school system. The program was initiated by Statistics Sweden and the National Board of Education.

The present report was financially supported by the Swedish Council for Planning and Coordination of Research and the National Agency for Higher Education.

ABSTRACT

Reuterberg, S-E. Selection of Swedish Scholastic Aptitude Test takers as a function of socioeconomic background and ability.

ISSN 0282-2156

Number of pages: 33

Since 1977 the Swedish Scholastic Aptitude Test (SweSAT) has been used as an admission test for applicants to higher education in Sweden. From the beginning only applicants who were 25 years old and who had been working for four years were allowed to apply for higher education on the basis of SweSAT scores, but in 1991 the admission rules were changed so that also younger applicants were allowed to do so. The new rules implied that the number of SweSAT takers increased from about 5,000 per year to more than 100,000.

An admission test like the SweSAT must satisfy strict demands on fairness in relation to applicants from different demographic groups. When the SweSAT was introduced most discussion concerned fairness in relation to socioeconomic background. Later, most attention has been paid to fairness between men and women.

However, since the test takers constitute self-selected groups, the fairness of the test cannot be determined unless relevant information is available about the groups tested. In this study such information is taken from the ETF project - a longitudinal data base for nationally representative samples - and from the data base BACE 72.

The present study focuses on the process of self-selection to the SweSAT among individuals from different socioeconomic groups and the strength of the self-selection process to the SweSAT is measured by the differences between test takers and others within each socioeconomic group on marks from compulsory school, achievement tests and intelligence tests.

According to all variables studied the test takers constitute a positively selected group within all socioeconomic groups. The strongest selection effects are found within group III and the weakest ones in group I. However, due to great social differences in the total sample the test takers from group III still obtain lower means on all variables studied than do the test takers from group I.

1. INTRODUCTION

In 1977 the admission rules to higher education in Sweden were changed so that even persons without an upper secondary education were eligible to enter universities and colleges. Since these persons had no leaving certificates from upper secondary school, there was a need for a supplementary selection instrument. By introducing this new instrument, the marks from upper secondary school were also supposed to be of less decisive importance for the individual's possibilities for entering higher education. The new instrument was to give the applicants a "second chance". The commission for constructing this instrument was given to a group of researchers at the Institute of Education, University of Umeå and the group is still responsible for the Swedish Scholastic Aptitude Test (SweSAT), which became the name of the new instrument.

By Act of the Parliament only applicants who were 25 years old and who had been working for at least four years were allowed to take the SweSAT and this group constituted a selection group of its own. However, in 1991 the admission rules were changed so that younger applicants could also apply for higher education on the basis of their result on the SweSAT, so that now admission is based either on marks from upper secondary school or on the SweSAT results. These new admission rules have resulted in an enormous increase in the number of test takers. During the 80's the number of test takers was constantly about 5,000 a year, but since 1991 between 100,000 and 150,000 persons per year take the SweSAT (Ingerskog & Stage, 1993).

The SweSAT consists of six different subtests, which measure verbal and non-verbal abilities, the capacity to make use of information, and knowledge of a general character:

Vocabulary, WORD, measures understanding of words and concepts. The task is to identify which of 5 presented words has the same meaning as a given word.

Data Sufficiency, DS, aims at measuring numerical reasoning.

Reading Comprehension, READ, contains four texts and six multiple choice questions in relation to each text.

Interpretation of Diagrams, Tables and Maps, DTM, consists of 10 collections of tables, graphs and/or maps with two multiple choice questions in relation to each such collection.

General Information, GI, measures knowledge and information that a person may acquire in different contexts such as work and education, and in social, cultural and political activities.

Study Techniques, STECH, consists of a booklet with a number of texts about a subject matter area. The task is to find the answers in the booklet to 20 different questions.

In 1992 *Study Techniques* was replaced by a test in *English Reading Comprehension, ERC.*

For all subtests the items are dichotomously scored and the test score is the number of correctly answered items. The composite score is obtained by summing the results over all subtests without any weighing procedure.

A more comprehensive account of the history of the SweSAT and the content of the test is given by Wedman (1994).

A selection test like SweSAT must satisfy strict demands on fairness in relation to applicants from different demographic groups. Thus the test must not be unfair to anyone because of gender, age, ethnic or socioeconomic background. From the very beginning most discussion concerned the fairness between applicants from different socioeconomic backgrounds, while little interest was paid to the gender differences (Wedman and Henrysson, 1992). Later, most interest has been paid to the gender differences found in favour of men.

Up to now, no information has regularly been collected on the socioeconomic background of the testees. Therefore, little is known about the impact of this variable on the SweSAT results. On the basis of studies in other countries, it was supposed, that the SweSAT would show smaller socioeconomic differences than do marks from upper secondary school (SOU, 1985, p. 85) - an assumption which was not confirmed by the fact that there was a weak tendency of somewhat increased socioeconomic differences on the first occassion of using the new admittance rules (Jansson, 1992).

So far only two empirical studies of the socioeconomic differences in the SweSAT scores have been reported. The first one, presented by Reuterberg, Westerlund and Gustafsson (1992), shows that there are substantial socioeconomic mean differences in test scores. Testees from socioeconomic group I scored 9 points higher than those from group III, and the socioeconomic differences were rather uniform for the different subtests. The main part of these socioeconomic differences could be explained by differences in the average mark from grade nine of compulsory school, the control variable used. However, in this study the correction for initial differences in school achievement was made on group level with the further assumption of no gender differences in school achievement.

In another study Gustafsson and Westerlund (1994) used regression analysis with data at the individual level and a more refined socioeconomic classification (seven categories instead of three). This study, too, showed substantial socioeconomic differences in SweSAT results. The correlation

between the total score on the SweSAT and socioeconomic background amounted to 0.16 which is only marginally lower than the correlation between this background variable and the marks from grade nine of compulsory school (0.18). Just as was found in the previously mentioned study, the differences were rather uniform over the subtests.

However, these group differences cannot be taken as indicators of the real socioeconomic differences in academic ability, since they may have at least two other contributing causes:

- differential selection to the SweSAT

- bias in the SweSAT scores.

By bias is meant that the test score is influenced by irrelevant factors which unduly favour some group of test takers.

In Sweden the debate has concentrated on possible bias in the SweSAT scores, but in order to study whether or not the SweSAT scores are biased we have to keep both the real differences in ability and the effects of differential selection under control. The present study focuses on the effects of differential selection.

Since only those individuals who aspire to an academic education take the SweSAT, the test takers constitute a positively selected sampel out of all individuals. If the selection mechanisms could be expected to work identically within the different groups, valid comparisons on the SweSAT scores could be made, but this is not plausible. Due to selection effects during previous educational stages there are good reasons to expect that there are differential selection effects to the SweSAT among individuals from different socioeconomic backgrounds.

The process of selection to higher education occurs in several successive steps. During grade seven through nine of compulsory school the pupils have to choose between two courses of different levels of difficulty in English and in Mathematics. After compulsory school, the youngsters can choose to enter upper secondary school or to leave school, and those who choose an upper secondary education must make a choice among several different study programmes. These programmes may be divided into two categories, namely the three or four year programmes and the two year programmes, the former of which being the most theoretical ones and providing the broadest qualification for admittance to higher education. Some of the two year programmes qualify for higher education, too, but in this case the youngster gets a more narrow qualification, normally to the shorter and more vocationally inclined programmes of higher education. In the last step the student has to make a judgement of whether he/she wants to enter higher education, and if so, to what extent the SweSAT score might enhance his/her chances to be admitted.

As shown by Härnqvist (1994) these successive choices are not unrelated. The choices of courses in English and Mathematics in compulsory school exert an influence on the choice of upper secondary school programme, partly directly, partly via the marks in grade nine. In the same way the transition to higher education and the choice of programme on this level are influenced by the choices made in upper secondary school.

Social background exerts a great deal of its influence on the choice of upper secondary program and the choice of higher education indirectly via school achievement and intermediate choices. However, there are also direct effects of socioeconomic background on the choice of programme in upper secondary school and also on the transition from the most theoretical programmes of upper secondary school to the longer programmes of higher education.

These results indicate that there may even be differential social selection effects to the SweSAT, but since the test is not compulsory for entrance into higher education, Härnqvist's results give no information on the strength of these effects.

When group comparisons have been made in Sweden on the basis of the SweSAT scores little, if any, consideration has been paid to the differential selection effects. In the USA, on the contrary, the debate has been intense when group comparisons have been made on the basis of the Scolastic Aptitude Test (SAT) scores and many researchers have stated that comparisons between self-selected groups do not lend themselves to any valid generalzations without relevant adjustments (Howe, 1985; Wainer, 1986a; 1986b; 1993; Wainer, Holland, Swinton & Wang, 1985; Linn, 1993). However, the question concerning what variables constitute relevant adjustment factors has caused some controversy.

In two American studies conducted by Powell and Steelman (Powell & Steelman, 1984; Steelman & Powell, 1985) and by Page and Feifs (1985) inter state comparisons were made on the basis of the SAT mean scores. In both these studies the state means were adjusted for differences in proportions of test takers. After having partialled out the effect of the SAT participation rate, Powell and Steelman found that

> "nearly 82 percent of the differences among states appear to be an artifact of the proportion of students electing to to take the SATs and, at best, only 18 percent can be considered 'real' variation." (Powell & Steelman, 1984 p 400).

In the same way Page and Feifs found a correlation between percentage of SAT takers and state mean score amounting to -0.85 (Page & Feifs, 1985, table 2, p 308).

The strategy of using the porportion of test takers as an adjustment variable was critized by Wainer (1986a, b) and by Wainer et al (1985) on several

grounds. Among other things Wainer puts the following question:

> "Does this adjustment miss anything having to do with who decides to take the SAT, rather than merely how many decide to take it?" (Wainer, 1986a, p 11).

As an alternative variable to the proportion of test takers Wainer (1986b) proposes the test takers' self-reported "rank in class". That is to say an indirect measure of school achievement.

In their reply to the critique Powell and Steelman (1987) defend their choice by the fact that there is a very high correlation between the proportion of test takers and rank in class (r=-.945), but at the same time they admit that the proportion of test takers constitutes only a proxy for academic ability.

Thus, the main problem for the American studies has been the lack of valid information from which the character of the self-selection process can be known. In Sweden, however, such data are available in a databank called ETF (Evaluation Through Follow-up). Within the frame of ETF nationally representative samples of pupils in compulsory school have been followed up with successive data collections from the age of 10 and all through the school system. A detailed description of the ETF is given by Härnqvist, Emanuelsson, Reuterberg and Svensson (1994).

Among a lot of other variables the ETF has several indicators on the individuals' academic ability, namely general intelligence, scores on achievement tests in compulsory school and marks from compulsory school and upper secondary school. However, due to fact that the marks from upper secondary school are not comparable between the different programmes they will not be included in the present study.

The aim of the present study

The present study is part of project which aims at studying recruitment to higher education in Sweden. As can be seen from the preceeding section the SweSAT plays an important part in the admission to this educational level and therefore it is natural to start the project by examining the function of this test.

In the present study we will confine ourselves to the process of self-selection to the SweSAT among individuals from different socioeconomic groups. The strength of the self-selection is measured by the differences between the SweSAT takers and the others within each socioeconomic group on marks from compulsory school, achievement tests and intelligence tests. If the strength of the self-selection differ between the socioeconmic groups we have differential selection effects and the aim of this study is to measure these differential selection effects. A parallell study of selection effects among men and women has been conducted by Mäkitalo (1994).

2. METHOD

Variables

The present study is based on data collected from the ETF cohort born in 1972. For this cohort the ETF data have been supplemented by information from the BACE 72 data base (Gustafsson and Westerlund, 1994) concerning the SweSAT scores for those individuals who have taken the test. Consequently, we have information for a nationally representative sample on the SweSAT scores and on a number of previously collected data which can reveal the differential selection effects to the SweSAT.

In Figure 1 we show the collection plan for those variables used in the present study.

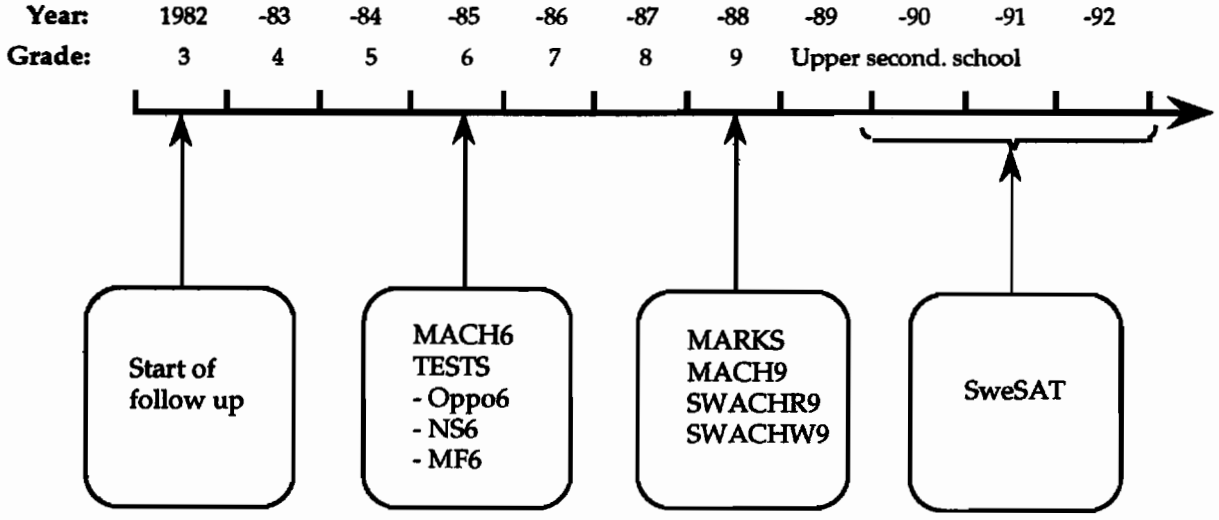

Figure 1. Collection plan for the variables.

In grade six the subjects were tested with three tests representing verbal, spatial and reasoning factors:

Opposites *(OPPO6)* is a traditional test measuring verbal ability. It includes 40 multiple choice items and the task is to select one word out of four, which is the antonym of a given word.

Metal folding *(MF6)* measures spatial ability. The task is to find a three-dimensional figure among four flat pieces of metal with bending lines. The test contains 40 items.

Number Series *(NS6)* measures reasoning ability. In each of the 40 items six numbers are given which are ordered according to a mathematical rule. The respondent's task is to detect the rule and add the two next numbers in the

series. In contrast to OPPO6 and MF6 the correct answers in this test is practically impossible to guess.

The scores from these three tests are combined into a total score *(TESTSUM)*, which constitutes a measure of general intellectual ability. Since the standard deviations are fairly equal the three tests have about equal weights in the total score.

Besides the tests, the students also had to take a mathematical achievement test in grade 6 *(MACH6)*. This test contains 42 multiple choice items covering different aspects of mathematical knowledge.

In grade nine all Swedish pupils have to take standardized achievement tests in Swedish and Mathematics. These tests are reference tests for making the marks comparable all over the country. The tests are administered by the teachers. There are two different standardised tests in Swedish, namely Reading comprehension *(SWACHR9)* and Written composition *(SWACHW9)* and these two tests are common to all students in grade nine.

In Mathematics the students have to choose between a general course and an advanced course in grades seven through nine. Therefore, the standardized achivement test in Mathematics *(MACH9)* has two versions, one for each course. Since the results from these two versions are not directly comparable, an estimated correction factor has been introduced as will be discussed later.

In grade nine, all pupils receive marks in all subjects studied. These marks range from a highest value of 5 to a lowest value of 1. For the whole population the marks should be distributed according to the normal curve with a mean of 3. This principle is valid also for the marks in Mathematics and the marks in English, but in these cases the pupils in the advanced and general courses constitute their own reference groups, and therefore, a correction factor has also been introduced for the marks in these two subjects.

The marks have been grouped according to a nested factor model presented by Gustafsson and Balke (1993). In this model the marks in all subjects load on a general factor, and in addition, most marks also load on one out of four first order factors orthogonal to the general factor. The *General factor (GSA)* measures general scholastic ability and the four first order factors measure scholastic ability in the *Natural science (NATSCI), Language (LANG), Social science (SOCSCI)* and *Practical-Spatial (SPAT)* domains, respectively.

In figure 2 we show the nested factor model in accordance to which the subject marks are grouped.

In the presentation of the results the tests and the achivement tests are also grouped into the relevant domains. However, the marks are not treated separately, but for each domain we have computed an average mark *(AVERMARK)*.

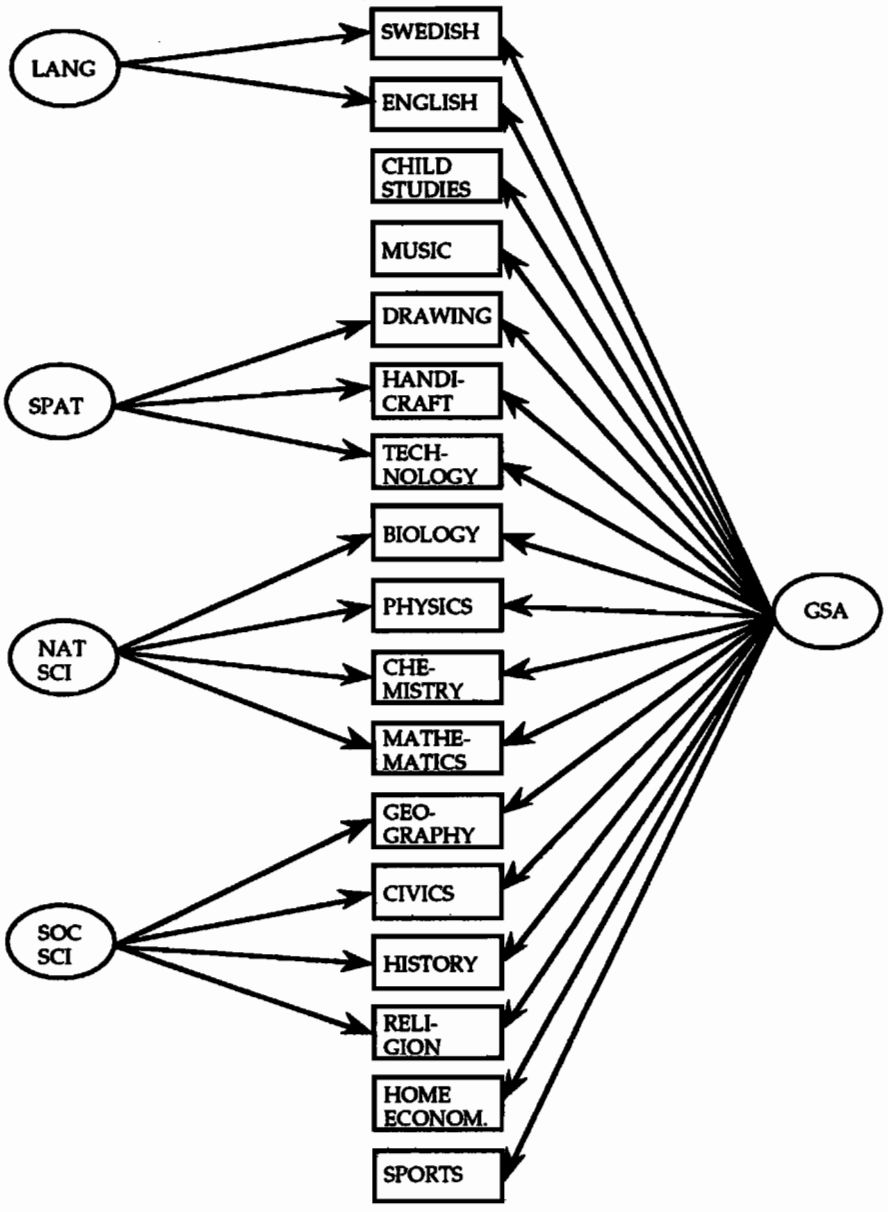

Figure 2. The nested factor model for marks.

As mentioned above, a special problem is caused by the fact that pupils in grade 7 through 9 have to choose between two different courses in Mathematics and in English. Irrespective of the course chosen, the pupils' marks in grade 9 are expressed on the same scale, which means that these marks are not comparable between the two courses. In the same way the standardized achievement tests in Mathematics taken in grade 9 differ between the two groups, so these results are not comparable either. However, a model has been set up which controls for differences in performance in all subjects and on the basis of that model constants are estimated which express the differences in scale means between persons who have chosen advanced and general courses, respectively. The model is shown in figure 3.

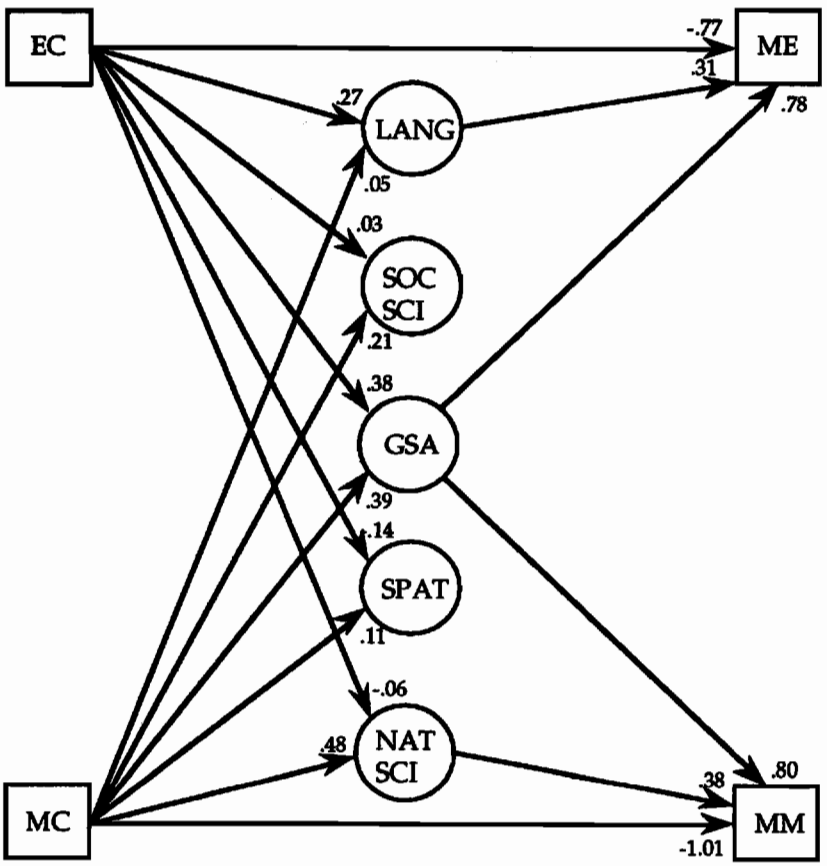

Figure 3. Model for estimation of correction factors for marks in English and Matematics.

In the model all latent variables are influenced by the dummy variables EC (course in English) and MC (course in Mathematics). The latent variables GSA and LANG influence ME (marks in English) and in the same way GSA and NATSCI exert influences on MM (marks i Mathematics). Besides these effects there are also direct effects from EC to ME and from MC to MM. The strength of these direct effects constitute the correction factors, since they show the differences in marks between general and advanced courses when student abilities are kept under control - in English via GSA and LANG and in Mathematics via GSA and NATSCI. By keeping ability under control in this way, the remaining differences in marks between the two courses should be seen as scale deficiencies and the unstandardized regression coefficients constitute the measures of the magnitude of these deficiencies. For MACH9 the correction factor has been computed in the same way, only with MACH9 replacing MM.

The correction factors obtained by this method are 0.77 for marks in English, 1.01 for marks in Mathematics and 12.75 for MACH9. These correction factors have been added to the results obtained by the students in the advanced courses.

In this study *the SweSAT* is handled as a dummy variable with a "1" assigned to those idividuals who have taken the SweSAT at least once from 1990 to 1992 and a "0" to those who have not.

The students are classified into three *Socioeconomic groups (SOC)* according to their parents' occupations. The groups are:

I Upper middle-class
II Lower middle-class
III Working class

Socioeconomic group is handled with two dummy variables with a "0" assigned to group I and a "1" assigned to group II and III, respectively. Two separate analyses are performed in which group I is compared with group II and group III, respectively.

This means that those who belong to SOC I and who have not taken the SweSAT constitute the reference group in the analyses to follow since this group has the value "0" for both SweSAT and SOC. A more detailed explanation is given in section "Statistical method".

Subjects

The present study is based on data collected from a sample of pupils who were in grade three of the Swedish compulsory school in spring 1982. As described by Härnqvist, Emanuelsson, Reuterberg and Svensson (1994) the total sample of some 9,000 individuals constitutes a nationally representative sample of all pupils in grade three. Since the sample is drawn out of pupils in a particular grade it contains individuals of varying ages. However, the great majority (95 per cent) were born in 1972.

From the large data base called 'BACE 72', including everyone born in 1972, the UGU data has been supplemented by the SweSAT scores from the years 1990 - 1992. This set of matched data is available for 8,729 individuals.

The available data imply some restrictions as to the generalizability of the results to be presented. In the first place, those individuals who were not in grade three at the age of ten are excluded, and in the second place we have no SweSAT data available for those individuals who have taken the SweSAT only later than in 1992.

Since the data are longitudinal, there is also some drop out as to separate variables. However, in order to minimize the effects of drop outs the analyses are performed throughout with 'pair-wise' exclusion of individuals. This means that we have included every individual who has information on those variables used in one and the same analysis.

However, there is one variable that causes a general drop-out, namely socioeconomic background. Information on this variable is available for 8,223 individuals, or 94 per cent of those with matched data. Among those who lack information on socioeconomic background, immigrants are overrepresented.

From the first column of table 1, we can see that less than 20 per cent of the total sample belong to socioeconomic group I, while group II constitutes nearly half the sample. In the total cohort, one individual out of four has taken the SweSAT, but this propotion varies a great deal between the socioeconomic groups. The second column shows that every second individual in group I has taken the test, but in group III the proportion of test takers decreases to only one out of eight. From these proportions it is obvious that group I is strongly overrepresented among the test takers while group III is clearly underrepresented. For group II the inclination to take the SweSAT corresponds well to that of the total sample.

Table 1.
The distribution of the sample on socioeconomic background and the proportions of SweSAT-takers in each socioeconomic group. (Per cent.)

	Total sample	SweSAT takers
SOC I	18	51
SOC II	48	28
SOC III	34	12
All	100	26

The socioeconomic differences in frequency of SweSAT taking are not unexpected. As can be seen in Appendix I, there are substantial socioeconomic differences in marks and achievement in compulsory school, and in the test results as well. Whether or not these differences can explain the socioeconomic differences in frequency of SweSAT taking is a question which is outside the scope of this study. However, the differences shown in table 1 indicate that there may be substantial differential selection effects to the test.

Statistical method

The statistical method used in this study is multiple regression analysis. Using this method it had been natural to let the SweSAT taking (SSAT) constitute the dependent variable, while SOC, MARK ,TESTSUM and

achievement scores constitute the independent variables. Taking TESTSUM as an example the regression equation would be:

$$SSAT=C+B_1(SOC)+B_2(TESTSUM)+B_3(SOC*TESTSUM)$$

However, in this equation the dependent variable is dichotomous, which implies that we had been forced to use some kind of logistic measures and these measures are not very easy to interpret. In order to avoid this and in order to receive a direct measure of the selection effects SSAT has been treated as an independent variable together with SOC, while TESTSUM has been treated as the dependent variable. Consequently, the actual regression equation in the analysis can be written:

$$TESTSUM=C+B_1(SOC)+B_2(SSAT)+B_3(SOC*SSAT)$$

With this equation the TESTSUM mean for each subgroup is predicted from SOC, SSAT and the interaction between SOC and SSAT.

The constant C constitutes the TESTSUM mean for the reference group, e.g. those in group I who have not taken the SweSAT.

Since SOC will be treated as a dummy variable with the reference group taken to be socioeconomic group I, the regression coefficients cannot be interpreted in the conventional way. In the first place we receive two B_1-coefficients - one for group II and one for group III. These coefficients show the mean difference between group I and groups II and III respectively, but *only for those individuals who have not taken the SweSAT.* However, these coefficients do not influence the selection effects and therefore they are not of primary interest in this study. Consequently, they will not be presented.

The B_2-coefficient is of greater interest, since it reflects the mean difference between test takers and others in socioeconomic group I. This means that B_2 is a measure of the selection effects to the SweSAT within this socioeconomic group.

In the same way as for the B_1 coefficients, the analyses will result in two different B_3 coefficients, one for group II and one for group III. These coefficients reflect to what extend the selection processes within these groups differ from that in group I. A positive value of B_3 implies that the selection effect is stronger in these groups compared to that in group I and a negative value means that the effect is weaker. In order to obtain an

absolute measure of the selection effects within these two groups their respective B_3 coefficients should be added to that of B_2.

The presentation of results will thus include three regression coefficients: one representing the strength of the selection effects in group I (B_2) and two B_3 -coefficients showing how much stronger or weaker the selection process has been in groups II and III respectively, as compared to that of group I. The sum of these coefficients reflects the selection effects within each socioeconomic group. Finally, in order to make possible comparisons between the different variables, the selection effects will be expressed as standard scores (the selection effects will be expressed as units of the standard deviation for each variable).

An example:

> When analysing the selection to the SweSAT according to TESTSUM we obtained the following regression equation for socioeconomic group II:
>
> TESTSUM= 72.26 -3.73 (SOC) + 11.53 (SSAT) + 1.51 (SOC*SSAT)
>
> and for group III the equation was:
>
> TESTSUM= 72.26 - 9.11 (SOC) + 11.53 (SSAT) + 5.88 (SOC*SSAT)

In these equations the constant (72.26) expresses the mean for those individuals belonging to socioeconomic group I who have not taken the SweSAT. The first regression coefficient (-3.73 and -9.11 respectively) shows that the mean difference between group I and II and group I and III respectively, among those who have not taken the test. In this case we can see that the mean for this category in group II falls 3.73 points below that for group I. The corresponding difference between group I and III amounts to 9.11 points, but, as stated previously, these differences are not of primary interest in the present study.

From the SSAT-coefficient of 11.53 we can see that the test-takers in group I score 11.53 points higher than the others in group I. This value constitutes a measure of the selection effect in this group. The coefficients for the interaction term, finally, show that this effect has been stronger in group II (11.53+1.51=13.04) and even stronger in group III (11.53+5.88=17.41). The standard deviation for TESTSUM amounts to 17.36 and therefore, the selection effects expressed as z-values will be 0.66 for group I, 0.75 for group II and 1.00 for group III.

By inserting "1" or "0" into the appropiate positions in the regression equations we can also reconstruct the means for the different subgroups as shown in table 2.

Table 2
Prediction of TESTSUM means for test takers and others within each socioeconomic group.

Category		Constant	SOC	SSAT	SOC*SSAT	PREDICT. MEAN
SOC I	Testees	72.26	0	11.53	0	83.79
	Others	72.26	0	0	0	72.26
SOC II	Testees	72.26	-3.73	11.53	1.51	81.57
	Others	72.26	-3.73	0	0	68.53
SOC III	Testees	72.26	-9.11	11.53	5.88	80.56
	Others	72.26	-9.11	0	0	63.15

From these predicted means we can confirm the conclusions drawn previously:

-The SweSAT takers constitute a positively selected group within each socioeconomic group.

- The selection effects to the SweSAT increase gradually when we move towards the lower socioeconomic groups.

Furthermore, the predicted means show that due to the fact that there are initial differences in the test results, SweSAT takers from higher socioeconomic groups still have a test mean which exceeds that of the SweSAT takers from group III.

3. RESULTS

Table 3 shows the selection effects to the SweSAT, as measured by intelligence tests, achievement tests and marks. The first column of the table presents the unstandardized regression coefficient for SSAT and, as said previously, it reflects the difference between test takers and others within socioeconomic group I. The interaction terms in the second column show to what extent the corresponding differences differ for SOC II and SOC III, respectively, from that in SOC I. These interaction terms should be added to the regression coefficient in the first column in order to obtain the actual difference between test-takers and others in SOC II and SOC III, respectively. These sums, called selection effect, are presented in the third column. The greater the difference between test takers and others, the stronger the selection effect.

In the fourth column, finally, the selection effects have been expressed as z-values. These z-values are comparable between different variables.

A * for the SSAT regression coefficients in table 3 means that the coefficient is significant on the 5 per cent level, which in turn means that there is a significant difference (selection effect) between test takers and others in socio-economic group I. A significant interaction term (SOC*SSAT) says nothing about the corresponding differences in groups II and III, but instead that the difference between test takers and others is significantly greater in group II or III than it is in group I. However, since groups II and III are bigger than group I and all the groups have about equal variances, a positive interaction term implies that there are significant selection effects in these two groups as well.

From table 3 we can see that all the regression coefficients for SSAT are positive and significant. Consequently, the test-takers in group I are positively selected in all respects studied. Furthermore, all interaction terms for groups II and III are positive, which means that the selection effects to the SweSAT are stronger in these groups than in group I, although not always significantly stronger. A closer look at these coefficients reveals that in no case do the test results give rise to any significant interaction term for group II, nor do the achievement tests within the natural science domain or the average mark within the spatial domain. In these respects we can only say that there is a tendency for the testees in group II to be more strongly selected to the SweSAT as compared to the testees in group I. Within group III the selection effects are in all respects significantly stronger than those in group I.

Table 3
Selection effects to the SweSAT in relation to SOC and ability.

DOMAIN	VARIABLE	SOC	SSAT	SOC*SSAT	SELECT.-EFFECTS	z
GENERAL	TESTSUM	I	11.53*	0	11.53	0.66
		II		1.51	13.04	0.75
		III		5.88*	17.41	1.00
	AVER-MARK	I	0.64*	0	0.64	0.89
		II		0.12*	0.76	1.06
		III		0.28*	0.92	1.28
NATURAL SCIENCE	NS6	I	5.14*	0	5.14	0.63
		II		0.51	5.65	0.69
		III		2.08*	7.22	0.88
	MACH6	I	4.79*	0	4.79	0.68
		II		0.74	5.53	0.78
		III		1.80*	6.59	0.93
	MACH9	I	15.79*	0	15.79	0.82
		II		1.10	16.89	0.87
		III		3.96*	19.75	1.02
	AVER-MARK	I	0.84*	0	0.84	0.88
		II		0.13*	0.97	1.02
		III		0.31*	1.15	1.21
LANG	OPPO6	I	3.70*	0	3.70	0.62
		II		0.59	4.29	0.72
		III		1.89*	5.59	0.94
	SWACH R9	I	9.39*	0	9.39	0.65
		II		1.93*	11.32	0.79
		III		5.76*	15.15	1.06
	SWACH W9	I	0.41*	0	0.41	0.57
		II		0.14*	0.55	0.76
		III		0.16*	0.57	0.79
	AVER-MARK	I	0.73*	0	0.73	0.83
		II		0.12*	0.85	0.97
		III		0.29*	1.02	1.16
SOCIAL SCIENCE	AVER-MARK	I	0.81*	0	0.81	0.90
		II		0.12*	0.93	1.03
		III		0.33*	1.14	1.27
SPATIAL	MF6	I	2.68*	0	2.68	0.36
		II		0.45	3.13	0.43
		III		1.96*	4.64	0.63
	AVER-MARK	I	0.38*	0	0.38	0.59
		II		0.07	0.45	0.70
		III		0.17*	0.55	0.86

Thus, the general conclusions drawn from the results in table 3 are that the SweSAT takers constitute a positively selected group within each socioeconomic stratum and that the selection process to the SweSAT becomes successively stronger when we move from group I to group III.

On the basis on the z-values we can make comparisons between the various variables as to the strength of the selection process. Such a comparison shows that the strongest selection occurs with respect to the different marks except for the spatial domain. This is hardly a surprising result. The marks are normally seen as the best indication of a person's ability for further theoretical education and they also normally constitute the selection instrument for transition from one educational level to the next. The fact that there seems to be a rather weak selection with respect to the marks within the spatial domain could be explained by the fact that the school subjects belonging to this domain are more practically oriented.

A comparison between the average marks from different domains shows an interesting result. Such a comparison namely indicates that the marks within the social science and the natural science domains give rise to somewhat stronger selection effects than do the marks for the language domain and about as strong effects as those for the general average mark.

After the marks follow the achievement tests in mathematics and reading comprehension, while the achievement test in writing seems to be of somewhat less importance. The main explanation for the relatively weak selection effect for the last mentioned variable is its very crude scale properties with only three levels: medium, above or below medium. With such a crude scale a great deal of the interindividual differences are hidden within each scale value.

The separate subtests show rather weak selection effects and especially the spatial one (MF6), although taken together, as in TESTSUM, the selection effects seem to be of about the same strength as those for the achievement tests.

4. DISCUSSION

In Sweden the selection of individuals to higher education is based on marks from upper secondary school or on the Scholastic Aptitude Test (SweSAT) scores. During the last years there has been a debate as to the fairness of the SweSAT scores. Primarily, this debate has focused on the gender differences found in favour of men, but of great importance is also to question whether or not the test is fair to test takers from different socioeconomic groups. The question of fairness, however, cannot be answered only on the basis of the SweSAT scores. Since it is volontary to take the SweSAT, the test takers constitute self-selected groups and consequently, there are no reasons to believe that the test takers constitute representative samples of all individuals in their socioeconomic groups. Nor are there any reasons to expect that the selection effects are of equal strength within the different socioeconomic groups. In other words, there may be differential selection effects and these effects must be known in order to assess the fairness of the SweSAT. The present study aims at investigating the selection effects to the SweSAT among individuals from different socioeconomic groups.

In order to measure the selection effects we need information on relevant variables for representative samples from each socioeconomic group. In this study such information is taken from two big data bases, the BACE 72 and the ETF project. The information used here consists of marks from grade nine in compulsory school, standardized achievement tests in Mathematics and Swedish and scores on three intelligence tests administered at the age of 13 for a nationally representative sample of individuals born in 1972. The selection effects are measured by the differences between test takers and others on these variables within each socioeconomic group.

In the sample as a whole one individual out of four has taken the SweSAT, a proportion which shows a great variation between different socioeconomic groups. In group I, every second individual has taken the SweSAT which should be compared to one out of four in group II and one out of eight in group III. These differences reflect the actual socioeconomic differences in transition to higher education quite well (cf Erikson & Jonsson, 1993, pp 182-185).

As could be expected, those who have taken the SweSAT constitute a positively selected group according to all variables studied. Taking the overall average mark as an example, tables I and V in the Appendix show that the test takers' mean exceeds that of the total group by nearly 90% of a standard deviation unit. However, the degree of selectivity varies with socioeconomic group. In table 4 the selection effects shown in table 3 are summarized in the form of mean z-values for each domain.

Table 4.
Average selection effects to the SweSAT within each domain.

SOC	Domain				
	General	Natural science	Language	Social science*	Spatial
I	0.78	0.75	0.67	0.90	0.48
II	0.91	0.84	0.81	1.03	0.57
III	1.14	1.01	0.99	1.27	0.75

*Average mark only.

As shown by the table there are substantial differential selection effects. The lowest degree of selectivity is found in group I within all the domains. The term "lowest" may be somewhat misleading since even in this group there are big differences between test takers and others. With the exception of the spatial domain, the average selection effects are between 0.7 and 0.9 units of a standard deviation and they are statistically significant throughout. Within socioeconomic group II the differences between test takers and others are greater than in group I, but not always significantly so for the separate vaiables. For group III, however, all variables show a significant interaction which means that the differences between test takers and others are greater in this group than they are in group I. From table 4 we can see that the average selection effects for group III amount to one unit of a standard deviation or more within all the "theoretical" domains. Consequently, there are substantial differential selection effects. Test takers from group III constitute the most positively selected group and the test takers from group I the least positively selected group out of all members of the socioeconomic groups, respectively.

These results clearly support the remark made by Wainer (1986a,b) that when self-selected groups are compared we need not only take into consideration the percentage of test takers in different groups, but also the characteristics of the test takers. If no such data are available comparisons between self-selected groups can cause severely misleading conclusions.

In order to get a more detailed picture of the differential selection effects the proportions of test takers in each socioeconomic group have been related to to the average mark for all school subjects in compulsory school, the variable which have shown the strongest selection effects in this study. These relationships are shown in figure 4.

Up to an average mark of 2 practically no one has taken the SweSAT, and no wonder. For individuals of such low ability levels higher education can hardly be regarded as a realistic alternative and, therefore, they have no reason to take the SweSAT. However, when the average mark exceeds 2 the proportions of test takers begin to increase and most rapidly so for group I and the socioeconomic differences remain big up to an average mark of about 4. On the highest mark levels, however, the proportions of test takers

increase most rapidly for groups II and III, but still there is a difference in favour of group I.

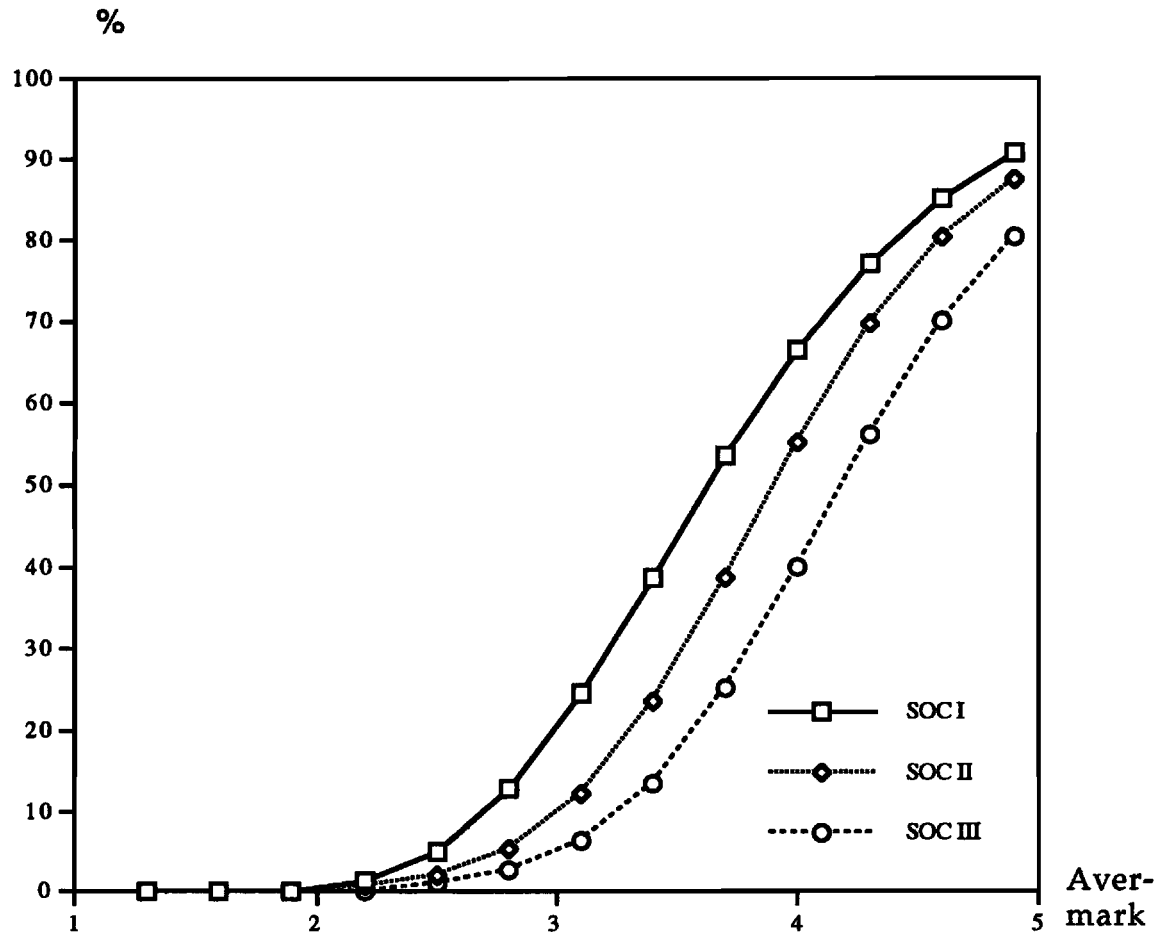

Figure 4. The proportions of SweSAT takers in relation to socioeconomic group and average mark.

Consequently, the differential selection effects to the SweSAT emanate primarily from individuals of medium and somewhat above medium ability, among whom the inclination to take the SweSAT is substantially stronger in group I as compared to the other groups. This is one of the reasons for our finding that the selection effects are weaker in group I than in group III. At the same time, however, the figure shows that especially in group III there are many individuals of very high ability, who do not take the SweSAT. Are these individuals not interested in higher education or do they rely on being admitted on the basis of their marks from upper secondary school? These are questions which will be the subjects of forthcoming studies.

The big socioeconomic differences in proportions of test takers among individuals of medium and somewhat above medium ability levels are interesting also from another perspective. On these ability levels we probably find just those individuals for whom the SweSAT has its most

important function as a second chance for admission into higher education and, as shown by the figure, this chance is most often taken by individuals from higher socioeconomic groups. This finding together with the high overall proportion of test takers in socioeconomic group I may imply that the SweSAT by constituting a second chance for admission to higher education adds to social inequality of educational opportunities. If so this is a consequence that was not expected when the test was introduced in 1977.

The strength of the selection process differs not only between socioeconomic groups, but also between different domains and between variables. The spatial domain gives rise to the lowest degree of selection which is quite natural considering the fact that these measures are less theoretical in nature. As to the different types of variables, the strongest selection occurs with respect to the marks. This result, too, is quite expected even if the marks in this study refer to school achievement in compulsory school. Normally, the marks are regarded as the best indication on the individual's possibilities to succeed in a higher theoretical education and the selection within the educational system is normally based on marks. After the marks follow the various achievement tests, while the separate intelligence tests show comparatively weak selection effects. However, taken together as a measure of a broad intellectual ability, the TESTSUM becomes nearly as important as the standardized achievement tests.

As shown by tables I to IV in the Appendix, within the total sample there are substantial socioeconomic differences according to all the variables, with group I having the highest means and group III the lowest ones. At the same time the selection processes to the SweSAT have worked differentially so that test takers from group III are the most positively selected and those from group I the least positively selected groups. From these two tendencies it is apparent that the socioeconomic differences in all variables must be smaller among the test takers compared to the differences found for the total groups. Have the selection processes then eliminated the differences or are there still any socioeconomic differences among the test takers?

From tables V to VIII in the Appendix we can see that the differences are much smaller among the test takers compared to those among all individuals, but there are still differences with group I obtaining the highest means and group III the lowest ones. These differences imply that we should not expect test takers from different socioeconomic groups to perform equally well on the SweSAT. Rather, some socioeconomic differences are to be expected in favour of test takers from higher socioeconomic groups. An important question which remains to be answered is whether or not the social differences among the test takers in the variables studied are big enough to explain the differences found in the SweSAT scores. Also this question will be the subject of a forthcoming study.

Finally, the differential selection effects found in this study comprise the whole period from compulsory school up to the decision on whether or not

to take the SweSAT. An important intermediate selection process occurs when the individual decides upon his/her upper secondary education - a decision which has a great influence on his/her possibilities for entering into higher education. In a study to be presented in a few months the total selection effects to the SweSAT will be separated on those effects which can be explained by the choice of upper secondary programme and those effects which are at work after the individual has finished upper secondary school.

REFERENCES

Erikson, R. and Jonsson, J. O. (1994) *Ursprung och utbildning. Social snedrekrytering till högre studier. [Origin and education. Social bias in recruitment to higher education.]* SOU 1993:85.

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

Gustafsson, J.-E. and Westerlund A. (1994) Socialgruppsskillnader i prestationer på högskoleprovet. [Social group differences in the Swedish Scholastic Aptitude Test scores.] In Erikson, R. and Jonsson, J. O. (Eds) *Sorteringen i skolan. [The sorting in school.]* pp 264-284. Stockholm: Carlssons bokförlag.

Howe, H. (1985) Let's have another SAT score decline. *Phi Delta Kappan, 66,* pp 599-602.

Härnqvist, K. (1994) Social selektion till gymnasium och högskola. [Social selection to upper secondary school and university.] In Erikson, R. and Jonsson, J. O. (Eds) *Sorteringen i skolan. [The sorting in school.]* pp 95-131. Stockholm: Carlssons bokförlag.

Härnqvist, K., Emanuelsson, I., Reuterberg, S.-E. and Svensson, A. (1994) Dokumentation av projektet "Utvärdering genom uppföljning". [Documentation of the project "Evaluation Through Follow-up".] *Institutionen för pedagogik, Göteborgs universitet,* Rapport 1994:03.

Ingerskog, P. and Stage, C. (1993) Högskoleprovet hösten 1993. Provgruppens sammansättning och resultat. [The Swedish Scholastic Aptitude Test autumn 1993. The composition of the testees and results.] *Pedagogiska institutionen, Umeå universitet,* PM 79.

Jansson, H. (1992) Betyg och prov vid antagning till högskolan. [Marks and tests in admission to the university.] In Betyg och högskoleprov för män och kvinnor. Rapport från ett seminarium. [Marks and the Swedish Scholastic Aptitude Test for men and women. A seminar report.] *Universitets- och Högskoleämbetet.* Rapport 1992:3, pp 24-26.

Linn, R. L. (1993) Educational assessment: Expanded expectations and challenges. *Educational Evaluation and Policy Analysis. Vol. 15,* pp 1-16.

Mäkitalo, Å. (1994). Non-comparability of female and male admission test takers. *Department of Education and Educational Research. Göteborg University.* Report no 1994:06.

Page, E. B. & Feifs, H. (1985) SAT scores and American states: Seeking for useful meaning. *Journal of Educational Measurement. Vol. 22,* pp 305-312.

Powell, B. & Steelman L. C. (1984) Variations in state SAT performance: Meaningful or misleading? *Harward Educational Review, Vol. 54,* pp 389-412.

Powell, B. & Steelman L. C. (1984) On state SAT research: A response to Wainer. *Journal of Educational Measurement. Vol. 24,* pp 84-89.

Reuterberg, S.-E., Westerlund, A. & Gustafsson, J.-E. (1992) Beror gruppskillnader på tidigare selektion i utbildningssystemet? [Are group differences due to previous selection in the educational system?]. In Betyg och högskoleprov för män och kvinnor. Rapport från ett seminarium. [Marks and the Swedish Scholastic Aptitude Test for men and women. A seminar report.] *Universitets- och Högskoleämbetet.*, Rapport 1992:3, pp 14-23.

SOU (1985) *Tillträde till högskolan. Betänkande av tillträdesutredningen. [Admission to the university. Report by the National Commission for admission to higher education.]* SOU 1985:57.

Steelman L. C. & Powell, B. (1985) Appraising the implications of the SAT for educational policy. *Phi Delta Kappan. 66,* pp 603-606.

Wainer, H. (1986a) *Drawing inferences from self-selected samples.* New York: Springer-verlag.

Wainer, H. (1986b) Five pitfalls encountered while trying to compare states on their SAT scores. *Journal of Educational Measurement. Vol. 23,* pp 69-81.

Wainer, H. (1993) Does spending money on education help? A Reaction to the Heritage Foundation and the Wall Street Journal. *Educational Researcher, 9,* pp 22-24.

Wainer, H., Holland, P., Swinton, S. and Wang, M. (1985) On "State education statistics". *Journal of Educational Statistics, Vol. 10,* pp 293-325.

Wedman, I. (1994) The Swedish Scholastic Aptitude Test: Development, use, and research. *Educational Measurement: Issues and practice. Vol. 13, No 2,* pp 5-11.

Wedman, I and Henrysson, S. (1992) Övervägande om könsmässig och social snedrekrytering vid högskoleprovets tillkomst och utformning. [Considerations on gender and social bias in recruitment to higher education at the time of the develpoment of the Swedish Scholastic Aptitude Test.] In Betyg och högskoleprov för män och kvinnor. Rapport från ett seminarium. [Marks and the Swedish Scholastic Aptitude Test for men and women. A seminar report.] *Universitets- och Högskoleämbetet.*, Rapport 1992:3, pp 8-9.

APPENDIX

Table I.
Means and standard deviations for socioeconomic groups.
General domain.

SOC	TESTSUM		AVERMARK	
	M	s	M	s
I	78.16	15.73	3.71	0.64
II	72.15	16.44	3.36	0.69
III	65.21	17.63	3.03	0.70
All	70.80	17.36	3.31	0.72
N	7150		8202	

Table II.
Means and standard deviations for socioeconomic groups.
Natural science domain.

SOC	NS6		MACH6		MACH9		AVERMARK	
	M	s	M	s	M	s	M	s
I	25.51	7.63	27.90	6.55	71.18	17.32	3.89	0.87
II	22.92	7.91	25.09	6.85	62.67	18.54	3.45	0.91
III	20.37	8.19	22.66	6.88	55.22	19.23	3.05	0.92
All	22.49	8.16	24.74	7.05	61.73	19.37	3.39	0.95
N	7170		7183		6656		8193	

Table III
Means and standard deviations for socioeconomic groups.
Language domain.

SOC	OPPO6		SWACHR9		SWACHW9		AVERMARK	
	M	s	M	s	M	s	M	s
I	25.93	5.56	63.98	11.66	2.35	0.67	3.95	0.81
II	23.80	5.64	59.03	13.33	2.06	0.71	3.53	0.83
III	21.52	5.95	53.27	15.42	1.84	0.69	3.17	0.83
All	23.38	5.94	57.92	14.34	2.04	0.72	3.48	0.88
N	7208		6348		6075		8188	

Table IV
Means and standard deviations for socioeconomic groups.
Social science and spatial domains.

SOC	SOCIAL DOM. AVERMARK		SPATIAL DOMAIN MF6		AVERMARK	
	M	s	M	s	M	s
I	3.69	0.79	26.69	6.52	3.52	0.59
II	3.29	0.85	25.42	7.12	3.31	0.62
III	2.87	0.87	23.22	7.74	3.08	0.64
All	3.22	0.90	24.88	7.36	3.27	0.64
N	7058		7181		8192	

Table V
Means and standard deviations for SweSAT takers from different socioeconomic groups. General domain.

SOC	TESTSUM		AVERMARK	
	M	s	M	s
I	83.75	12.95	4.02	0.46
II	81.39	13.33	3.91	0.45
III	80.37	14.03	3.84	0.51
All	82.02	13.37	3.94	0.47
N	1913		2153	

Table VI
Means and standard deviations for SweSAT takers from different socioeconomic groups. Natural science domain.

SOC	NS6		MACH6		MACH9		AVERMARK	
	M	s	M	s	M	s	M	s
I	28.01	6.56	30.21	5.63	78.42	14.59	4.20	0.65
II	26.92	6.86	29.03	5.84	74.52	14.46	4.07	0.65
III	26.66	7.30	28.38	6.26	72.77	15.85	3.98	0.70
All	27.24	6.85	29.32	5.87	75.63	14.86	4.10	0.69
N	1919		1917		1834		2153	

Table VII
Means and standard deviations for SweSAT takers from different socioeconomic groups. Language domain.

SOC	OPPO6		SWACHR9		SWACHW9		AVERMARK	
	M	s	M	s	M	s	M	s
I	27.73	4.83	68.69	7.27	2.55	0.58	4.31	0.64
II	26.84	4.85	67.17	8.40	2.46	0.60	4.14	0.64
III	26.40	5.58	66.37	9.38	2.33	0.65	4.07	0.66
All	27.07	4.99	67.55	8.25	2.47	0.61	4.19	0.65
N	1922		1679		1624		2153	

Table VIII
Means and standard deviations for SweSAT takers from different socioeconomic groups. Social sciencs and spatial domains.

SOC	SOCIAL DOM.		SPATIAL DOMAIN			
	AVERMARK		MF6		AVERMARK	
	M	s	M	s	M	s
I	4.08	0.58	27.99	5.89	3.71	0.52
II	3.95	0.58	27.63	6.22	3.64	0.53
III	3.89	0.65	27.27	6.09	3.57	0.57
All	3.99	0.59	27.70	6.09	3.66	0.53
N	1918		1917		2153	

REPORTS FROM DEPARTMENT OF EDUCATION AND EDUCATIONAL RESEARCH, GÖTEBORG UNIVERSITY

ISSN 0282-2156

To be ordered from Department of Education and Educational Research, Göteborg University, P.O. Box 1010, S-431 26 MÖLNDAL, Sweden

---

Fensham, P & Marton, F. High-school Teacher' and University Chemists' Differing Conceptualisations of thePersonal Activity in Constituting Knowledge in Chemistry. 1991:01

Kärrby, G. 22.000 Minutes in Preschool. 5-6 year old children´s activities, language and group constellations in Swedish preschools. 1991:02

Laurillard, D, Lindström, B, Marton, F & Ottosson, T. Computer simulation as a tool developing intuitive and conceptual understanding. 1991:03

Dahlin, B. Constituting a conceptional world. A phenomenological and existential approach to education. 1991-04

Marton, F, Dall'Alba, G & Beaty, E. Conceptions of learning. 1991:05.

Asplund Carlsson, M. The structure of recall. An empirical study of repeated readings of Franz Kafka's "Before the Law". 1991-06.

Uljens, M. What is learning a change of? 1992:01

Holmer, J, Larsson, S, Thång, P-O, Petersson, B, Stockfelt, T. Working Life, Learning and Adult Education. 1992:02

Uljens, M. Phenomenological features of phenomenography. 1992:03

Hägglund, S. What did they learn? Swedish children's experiences of schooling in Belgium. 1992:04.

Axelsson, H. Environment and school initiatives - ENSI. The Swedish report about the Environmental Education project, initiated at OECD/CERI. 1993:01.

Asplund Carlsson, M, Halász, L, Marton, F. Some Paradoxical effects of attempts to facilitate readers' understanding of a Kafkaesque paradox. 1994:01.

Holmberg L, M. The thoughts and practices of user documentation developers. 1994:02.

Holmberg, L M. A content analysis of nine user documentations on operating systems. 1994:03.

Ahlberg, A. & Csoscán, E. Grasping numerosity among blind children. 1994:04.

Reuterberg, S-E. Selection of Swedish Scholastic Aptitude Test takers as a function of socioeconomic background and ability. 1994:05.

Tryckt & Bunden
Vasastadens Bokbinderi AB
Göteborg 1994

issn 0282-2156