Reports from the Department of Education and Educational Research

EVALUATION
THROUGH
FOLLOW-UP

WHO TAKES THE SWEDISH SCHOLASTIC APTITUDE TEST?

A study of differential selection to the SweSAT in relation to gender and ability.

Åsa Mäkitalo

Sven-Eric Reuterberg

Report No. 1996:03
Department of Education and Educational Research
Göteborg University

EVALUATION THROUGH FOLLOW-UP

WHO TAKES THE SWEDISH SCHOLASTIC APTITUDE TEST?

A study of differential selection to the SweSAT in relation to gender and ability.

Åsa Mäkitalo **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

Åsa Mäkitalo and Sven-Eric Reuterberg, WHO TAKES THE SWEDISH SCHOLASTIC APTITUDE TEST? A study of differential selection to the SweSAT in relation to gender and ability.

ISSN 0282-2156

Number of pages: 24

The gender differences in SweSAT scores in favour of male test takers have been the subject of a rather intense public debate in Sweden during the last few years. Normally, these differences have been interpreted as a consequence of bias in the test. However, an alternative explanation might well be that the differences in test scores are caused by differential selection effects, which implies that the male and the female test takers are not comparable.

In this study the differential selection effects to the SweSAT are studied for a nationally representative sample of male and female test takers born in 1972. The selection effects are measured by test scores, scores on standardized achievement tests and grades from the compulsory school.

According to all these variables the male test takers are more strongly selected to the SweSAT than are the female test takers. That is to say that the differences between test takers and others in all of these variables are greatest among men. To some extent, these differential selection effects are the result of men being more variable in all the respects studied. A statistical method has been developed for keeping this difference in variability under control. This control implied that the differential selection effects were reduced - for some of the variables up to nearly 50 per cent- but still the male test takers were more positively selected.

Then, another control variable was introduced, namely previous education measured by the programme chosen in upper secondary school, and the differential selection effects were studied separately for those who had finished a theoretical upper secondary programme and for those who had not. When introducting this control variable the differential selection effects disappeared within the theoretical group, but within the nontheoretical group the male test takers remained more positively selected.

Since the great majority of the SweSAT takers belongs to the theoretical group, the results show that the differential selection effects to the SweSAT are mainly due to the differential selection in the transistion from compulsory school to upper secondary school.

Furthermore, it has been shown that even if differential selection effects must be taken into consideration when comparing self selected groups, they cannot by themselves explain the group differences in SweSAT scores. Also the differences in the unselected group have to be taken into consideration before claiming bias in the SweSAT.

INTRODUCTION

Admission tests for entrance into higher education tend to show gender differences in results favouring males. On the Scholastic Assessment Test (SAT) the gender difference on the mathematical part amounts to about half a standard deviation unit. Up to 1972 the verbal part showed differences in the opposite direction, but since then males have outperformed females even on the verbal sections (Wilder & Powell, 1989).

The Swedish counterpart to SAT, the Swedish Scholastic Aptitude Test (SweSAT) has shown gender differences in favour of males ever since its introduction in 1977. However, these differences were rarely the subject of any public debate, mainly because the test used to play a limited role in selection to higher education. Only adult applicants, namely, were allowed to take the test. In 1991 the test was given a much more important role as an alternative selection instrument to the leaving certificate from upper secondary school among all applicants. The new role of the SweSAT resulted in a dramatic extension of its use and the gender differences in scores were discussed more intensely.

The SweSAT contains six time limited subtests. Three subtests are verbal *(Vocabulary, Reading comprehension, English reading comprehension)*, two are more quantitative *(Data Sufficiency and Diagrams, Tables and Maps)* and one is a test of general knowledge *(General information)*. All items are in the format of multiple choice and the total SweSAT score is the sum of the number of correctly answered items. A more comprehensive presentation of the SweSAT, its content and history is given by Wedman (1994).

Up to 1992 the gender differences amounted to 8 points out of a maximum of 144 items (See Stage, 1985; 1988; 1990; 1992). In 1992, when a test of Study Techniques was replaced by the English reading comprehension test, the difference amounted to 10 points out of a total of 148, which corresponds to about half a standard deviation unit (Ingerskog & Stage, 1993). The greatest gender differences in favour of males have always been found on the more quantitative tests Data Sufficiency (DS) and Diagrams, Tables and Maps (DTM), which is in accordance with earlier studies of results on mathematical tests (Hyde, Fennema & Lamon, 1990). In standardized mean differences the gender difference amounts to approximately 0.60 through 0.70 on these tests. However the gender differences go in the same direction also for all the other subtests even if they are smaller on the verbal parts, about 0.20 - 0.25 on the Vocabulary and Reading comprehension tests and about 0.40 on the English reading comprehension test.

As stated by Wilder & Powell (1989) the gender differences on admission tests may be regarded in different ways. They may be regarded as real, and if so the problem is to identify the underlying mechanisms. Another way is to regard them as artifacts of differential treatment of men and women in society. A third way is to question their existence by claiming bias in the test, differential selection of test takers or statistical effects.

Wilder & Powells' review of possible causes of the gender differences show that biological, social, psychological and educational explanations have been considered. Biological explanations have mostly been addressed to differences in spatial ability and the explanations often focus on genetic and chromosomal determinants, sex hormones or differences in brain structure and function (Halpern, 1986). The social and psychological explanations often focus on differential socialization processes or the social construction of gender (Chodorow, 1978; Eagly, 1987; Gilligan, 1982; Lorber & Farrell, 1991), different cognitive styles (Messick, 1976), achievement motivation or self-confidence (Lenney, 1981).

The educational explanations mostly concern differences in educational experiences (Wernersson, 1977; 1988) and course taking (Chipman & Thomas, 1985; Wice, 1985). Fennema & Sherman (1977a) found that variables associated with the female sex-role influenced the election of mathematical courses among tenth- and eleventh grade females.

This sex-role influence worked through factors such as confidence, expected usefulness and the perceived expectations of significant others. These kinds of views and sets of values are to a great extent socio-historical. In Sweden, for instance, females once were denied the formal opportunities to get a higher education, and not until 1927 did they get access to the former male public schools (Florin & Johansson, 1993). Since then young Swedish females have increased their educational investments enormously and in 1991 51% of the admitted applicants to higher education were females (Forneng & Jansson, 1991).

The choices of educational programmes, however, still reflect traditional sex roles (Franke-Wikberg, 1981). Males are clearly overrepresented in the scientific/technical and technical/industrial sectors of upper secondary school, while females usually are found in the social, humanistic and economy sectors (Härnqvist & Svensson, 1981; Wernersson 1991). These gender differences in upper secondary school are also reflected in higher education where females also are overrepresented on the more vocationally oriented programmes (Swedish Ministry of Education and Science, 1992; SCB, 1993)

Since the SweSAT is an admission test for higher education it is obvious that the test takers constitute a positively selected group on the basis of earlier school achievements. It has been stated repeatedly that the results from such unrepresentative groups do not lend themselves to any valid generalizations without relevant adjustments (Howe, 1985; Wainer 1986a; 1986b; 1993). However, considering the great gender differences in educational careers mentioned above, there are not even reasons to assume that male and female test takers have been selected in the same way. On the contrary, there are several studies which indicate differential selection effects among men and women (Fennema & Sherman, 1977b; Hyde & Linn, 1988; Rosenthal & Rubin, 1982; Reuterberg, Gustafsson, & Westerlund, 1992; Mäkitalo, 1994). As to the SweSAT, the male test takers have been shown to be more positively selected than are the females. Thus, differential selectivity is a factor which has to be taken into account when the male and female scores on the SweSAT are compared. It should be pointed out, however, that the stronger selection of males, in itself, does not tell us if males are expected to achieve better on SweSAT scores. Before any conclusion of this kind can be made we have to take into account the differential selection effects as measured by the initial achievement levels of male and female test takers.

Differential selectivity has been given little attention in the public debate in Sweden. Instead, bias in the test has been claimed as the main cause of the gender differences and several studies have been conducted focusing the impact of item content and format, testing time, item position, and problem solving strategies (Henriksson, Stage & Lexelius, 1986; Stage, 1987; Wester-Wedman, 1992a; 1992b; 1992c; Mäkitalo, 1993). However, these studies have not resulted in any greater changes of the SweSAT.

The present study will focus on the differential selection effects among male and female SweSAT takers. In measuring the selection effects we have to take gender differences in variability into account (Becker & Hedges, 1988; Humphreys, 1988; Cleary, 1991 and Feingold, 1992). Cleary (1991), for instance, showed that differences in variability have a great impact on the group comparisons and this impact is different at different points of the score distribution. She found boys to be more variable than girls and the comparison of two positively selected samples - one from each gender - showed that the girls were disadvantaged to a greater extent the more extreme the sample. She also found that the effect sizes in favour of boys increased with more quantitative items and with age. Earlier studies of gender differences in variability have shown greater variability among men in mathematical and spatial abilities (Maccoby & Jacklin, 1974) as well as on standardized aptitude test batteries (Feingold, 1992). However, the result pattern is not invariant across cultures (Feingold, 1994). In our data we also have found greater variability among the males which means that in order to get a 'pure' measure of the differential selection effects the differences in variability have to be controlled for. As far as we know, there is no standard method available for making such a control, and therefore a method will be developed for this purpose. The primary aim of the present study, however, is to study the differential selection effects to the SweSAT among male and female test takers and to investigate to what extent these differential selection effects are influenced by previous educational careers, i.e. participation in theoretical or nontheoretical upper secondary programmes.

METHOD

Subjects

The present study is based on data collected within a Swedish longitudinal project called Evaluation Through Follow-up (ETF). This project has followed up nationally representative samples of pupils born in 1948, 1953, 1967, 1972, 1977, and 1982, respectively, from the age of 10 or 13 and all through the formal school system (Härnqvist, Emanuelsson, Reuterberg & Svensson, 1994). The 9,000 pupils included in this study were in grade three of the Swedish compulsory school in the spring of 1982. 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 ETF data have been supplemented by the SweSAT scores from the years 1990 - 1992. This set of matched data is available for 8,728 individuals. 26 per cent of the total group have taken the SweSAT during the period mentioned, and this proportion is somewhat higher for the females as compared to the males - 28 and 23 per cent, respectively.

The available data imply some restrictions as to the generalizability of the results. 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 design is 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.

Variables

As shown by figure 1 the data collection among those born in 1972 started in 1982.The data used in this study, however, have been collected on two later occasions, namely in grades 6 and 9.

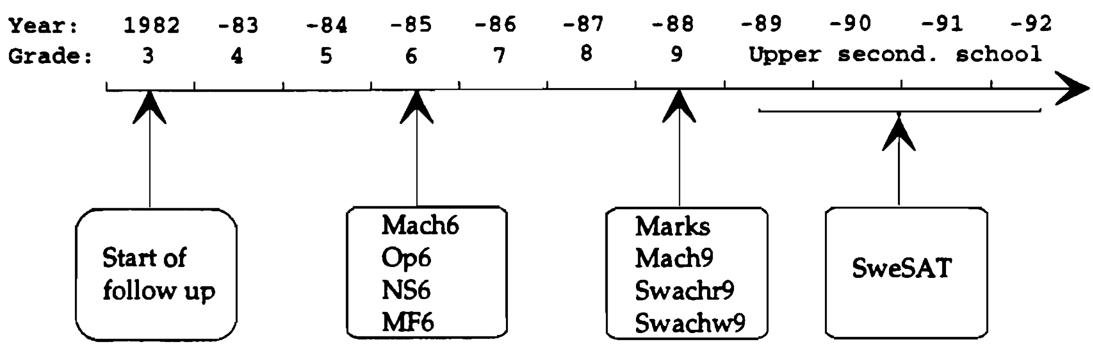

Figure 1. Collection plan for the variables.

In grade 6 the pupils were tested with three tests representing verbal, spatial and reasoning factors:

Opposites *(Op6)* 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 identify 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 Op6 and MF6 the correct answers in this test are 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.

In addition, the students also took 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 pupils had to take standardized achievement tests in Swedish and Mathematics. These tests constitute reference tests for making the marks comparable all over the country. The tests are administered by the teachers. There are two different standardized 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 (Reuterberg, 1994).

In grade nine, all pupils receive marks in all school 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 normally distributed 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 been introduced also for these two variables (Reuterberg, 1994).

The variables have been grouped into three domains:

General domain, which includes Testsum and the average mark from grade 9 of compulsory school *(GSA).*

Verbal domain, which includes Op6, Swachr9, Swachw9 and Verbmark. Verbmark is defined by the average mark for Swedish and English.

Natural science domain includes NS6, Mach6, Mach9 and Natmark. The last mentioned variable is defined by the average mark for Biology, Chemistry, Mathematics and Physics.

The variables belonging to the verbal domain are regarded as indicators of the ability primarily measured by the verbal subtests of the SweSAT and the those belonging to the natural science domain are regarded as indicators of the ability primarily measured by the quantitative subtests of the SweSAT.

The *SweSAT* is handled as a dummy variable with a "1" assigned to those individuals who have taken the SweSAT at least once from 1990 to 1992 and a "0" to those who have not. Also upper secondary education (*USE)* is handled as a dummy variable. In this case a "1" is assigned to those who have finished a theoretical upper secondary programme of at least 3 years of study. All others have been assigned a "0".

Statistical method

The statistical method used in this study is multiple regression analysis with the variable that expresses the selection effects constituting the dependent variable, and SEX, USE and SweSAT constituting the independent variables. All independent variables are handled as dummy variables with 0 assigned to males, to those who have no theoretical upper secondary education and to those who have not taken the SweSAT, respectively.

The case of two independent variables:

The analyses of the total selection effect to the SweSAT comprise only two independent variables, namely SEX and SweSAT, and in this case the mean of each subgroup on the dependent variable (y) is estimated by the following regression equation:

$$y = C + b_1 SEX + b_2 SweSAT + b_3 SEXxSweSAT$$

The y-means are estimated for each of the subgroups by summing those coefficients for which an "x" has been assigned in the tableau below.

Group		C	b_1	b_2	b_3
SEX	SweSAT		SEX	SweSAT	SEX x SweSAT
0	0	x			
0	1	x		x	
1	0	x	x		
1	1	x	x	x	x

However, the aim of the study is not to predict the y-means, but the *Selection Effects (SEff)*, that is to say the differences between SweSAT takers and others within each gender. In this case the coefficient C is of no importance for the selection effects to the SweSAT since it is a constant for all the four subgroups. Neither is b_1 of importance, since it differs only between males and females but not between SweSAT takers and others within each gender. Then only b_2 and b_3 remain.

Focusing on the selection effects to the SweSAT within each gender we can see from the tableau above that SEff is obtained in the following ways:

For males: $SEff_m = b_2$

For females: $SEff_f = b_2 + b_3$

The *Differential Selection Effects (DSEff)* is defined as the difference between the selection effects for males and females, respectively. Thus:

$$DSEff = b_3$$

Since each regression coefficient is subjected to test of significance, this method of computing the selection effects also implies a direct statistical test of the significance of

DSEff.

Both SEff and DSEff are expressed as unstandardised regression coefficients which means that they are directly related to the standard deviation of the y-variable. Therefore, they are not comparable between variables with differing standard deviations. In order to make them comparable over the various y-variables they should be divided by the standard deviation for the y-variable. However, in order not to make the standard deviation influenced by the mean differences between groups the standard deviation has been computed on the basis of the pooled within-group variation S_{yw}.

Thus, the *Standardized Selection Effects (SSEff)* are obtained by the following expressions:

For males: $$SSEff_m = \frac{SEff_m}{S_{yw}} = \frac{b_2}{S_{yw}}$$

For females: $$SSEff_f = \frac{SEff_f}{S_{yw}} = \frac{b_2 + b_3}{S_{yw}}$$

and the *Differential Standardized Selection Effects(DSSEff)* is obtained by:

$$DSSEff = \frac{DSEff}{S_{yw}} = \frac{b_3}{S_{yw}}$$

SSEff and DSSEff could also be obtained by first transforming the raw scores of y to a scale with a standard deviation of 1 and then perform the regression analysis on the basis of these transformed values.

Since SSEff and DSSEff are obtained by the pooled within-group variation, they do not take into consideration the fact that the variability may differ for males and females. In order to adjust for this gender difference two more selection effects are computed, namely the *Adjusted Selection Effect (ASEff)* and the *Differential Adjusted Selection Effect (DASEff)*. These effects are obtained by using each subgroup's own standard deviation instead of using the pooled within-group standard deviation:

For males: $$ASEff_m = \frac{SEff_m}{S_{ym}} = \frac{b_2}{S_{ym}}$$

and

for females: $$ASEff_f = \frac{SEff_f}{S_{yf}} = \frac{b_2 + b_3}{S_{yf}}$$

The differential adjusted selection effect (DASEff) constitutes the difference between males' and females' adjusted selection effects and it is obtained by the following expression:

$$DASEff = ASEff_m - ASEff_f = \frac{b_2}{S_{ym}} - \frac{b_2 + b_3}{S_{yf}}$$

ASEff and DASEff can also be computed by first transforming the raw scores to a scale which has a standard deviation of 1 for each subgroup and after that the regression analysis is performed.

The case of three independent variables:

As we have discussed previously, the choice of upper secondary education (USE) is supposed to be of at least some importance for the selection to the SweSAT. Therefore, the selection effects will be studied also with this variable included as an independent variable. Then the y-mean for each of the eight subgroups is estimated by the following regression equation:

$$y = C + b_1(USE) + b_2(SEX) + b_3(SweSAT) + b_4(USExSEX) + b_5(USExSweSAT) + b_6(SEXxSweSAT) + b_7(USExSEXxSweSAT)$$

As before, SEff is estimated from only those regression coefficients which refer to terms including SweSAT. Thus, for each of the four subgroups obtained by crossing USE and SEX, SEff is estimated by summing those regression coefficients which have an "x" assigned in the tableau below.

Subgroup		b_3	b_5	b_6	b_7
USE	SEX	SweSAT	USE x SweSAT	SEX x SweSAT	USE x SEX x SweSAT
0	0	x			
0	1	x		x	
1	0	x	x		
1	1	x	x	x	x

Keeping in mind that

- DSEff stands for the gender difference in selection effects within each educational group
- DSSEff is the differential selection effect divided by S_{yw}.
- DASEff is obtained by dividing the selection effect within each subgroup by the group's own standard deviation and after that the gender differences are computed within each educational group.

Then we will have the following expressions:

For USE=0:

$$DSEff = b_6 \qquad DSSEff = \frac{b_6}{S_{yw}} \qquad DASEff = \frac{b_3}{S_{y00}} - \frac{b_3 + b_6}{S_{y01}}$$

For USE=1:

$$DSEff = b_6 + b_7 \qquad DSSEff = \frac{b_6 + b_7}{S_{yw}} \qquad DASEff = \frac{b_3 + b_5}{S_{y10}} - \frac{b_3 + b_5 + b_6 + b_7}{S_{y11}}$$

RESULTS

Selection effects to the SweSAT in relation to gender and ability

In Table 1 we present those regression coefficients and standard deviations which are required for computing the various selection effects mentioned above.

Table 1.
Regression coefficients and standard deviations for the analysis of selection effects to the SweSAT.

Domain		Regression coefficients		Standard deviations		
		SweSAT	SweSAT x SEX	Females	Males	Within groups
		b_2	b_3	S_{yf}	S_{ym}	S_{yw}
General	GSA	0.952*	-0.184*	0.701	0.732	0.714
	Testsum	17.238*	-2.775*	17.027	17.806	17.428
Natural sciences	NS6	7.577*	-1.693*	7.858	8.484	8.182
	Mach6	7.513*	-1.968*	6.743	7.371	7.070
	Mach9	23.604*	-6.916*	18.371	20.310	19.390
	Natmark	1.251*	-0.297*	0.913	1.000	0.959
Verbal	Op6	5.340*	-0.375	6.092	5.826	5.958
	Swachr9	15.346*	-4.173*	13.023	15.377	14.268
	Swachw9	0.661*	-0.159*	0.671	0.702	0.686
	Verbmark	1.057*	-0.213*	0.833	0.868	0.849

*) Significant at the 5 per cent level

SweSAT takers constitute a positively selected group with respect to all the variables. However, the coefficients for the interaction between SweSAT and SEX (b_3) are all negative, and with one exception statistically significant. This means that the female SweSAT takers are less positively selected out of all females than are the male test takers out of all males. In other words, there are substantial differential selection effects between males and females.

This result must not be interpreted to mean that the female test takers have lower means on the variables than have the male test takers. Gender differences in favour of the females in the total sample may be so large that the female test takers still outperfom the male test takers in spite of their weaker selection effects to the SweSAT.

The standard deviations in Table 1 show a greater variability for men. There is only one exception, namely Op6, and in this case the females have only a slightly higher standard deviation. The greatest differences in favour of men are found for Swachr9, Mach9 and Natmark with standard deviation ratios of 1.10 or more.

According to our discussion in the previous section, all the different selection effects can be computed on the basis of these regression coefficients and standard deviations. However, in order to facilitate reading, the selection effects (SEff) for both sexes are shown in Table 2 together with the differential selection effects (DSEff). Since the males constitute the reference group, the selection effects for them are identical to the b_2-coefficients in Table 1, and the b_3-coefficients in this table correspond to the DSEff-values in Table 2.

Table 2
Selection effects (SEff) and differential selection effects (DSEff) to the SweSAT.

Domain	Variable	$SEff_f$	$SEff_m$	DSEff
General	GSA	0.768	0.952	-0.184*
	Testsum	14.463	17.238	-2.775*
Natural science	NS6	5.884	7.577	-1.693*
	Mach6	5.545	7.513	-1.968*
	Mach9	16.688	23.604	-6.916*
	Natmark	0.954	1.251	-0.297*
Verbal	Op6	4.965	5.340	-0.375
	Swachr9	11.173	15.346	-4.173*
	Swachw9	0.502	0.661	-0.159*
	Verbmark	0.844	1.057	-0.213*

As mentioned before the selection effects are expressed in the raw score scales and therefore, they are not comparable between the different variables. However, the standardized selection effects (SSEff) shown in Table 3 are comparable.

Table 3.
Standardized selection effects (SSEff) and differential standardized selection effects (DSSEff) to the SweSAT.

Domain	Variable	$SSEff_f$	$SSEff_m$	DSSEff
General	GSA	1.076	1.333	-0.257*
	Testsum	0.830	0.989	-0.159*
Natural science	NS6	0.719	0.926	-0.207*
	Mach6	0.784	1.063	-0.279*
	Mach9	0.861	1.217	-0.401*
	Natmark	0.995	1.304	-0.309*
Verbal	Op6	0.833	0.896	-0.063
	Swachr9	0.783	1.076	-0.293*
	Swachw9	0.732	0.964	-0.232*
	Verbmark	0.994	1.245	-0.251*

In commenting on the differential selection effects consideration is taken only into the strength of the effect. The signs give supplementary information about whether females or males show the strongest selection effect.

As shown by Table 3 the female SweSAT takers outperform the other females with between 0.7 and 1.0 standard deviation units, while the corresponding differences among men are between 0.9 and 1.3 standard deviation units. Accordingly, the DSSEff-values normally fall between 0.2 and 0.4 standard deviation units, the males being most strongly selected. The only exceptions are Testsum and Op6 which both have DSSEff-values lower than 0.2. Moreover, the DSSEff for Op6 of 0.063 is the only differential standardised selection effect which is not significant.

While the test variables (Testsum, NS6 and Op6) have the lowest standardized selection effects, the highest ones are found for the marks with the very highest value for the over all average mark (GSA). This result is to be expected since this variable usually is regarded as the best indicator of academic ability and it constitutes a selection instrument for admittance into higher educational levels.

A comparison between the natural science and verbal domains in Table 3 shows that the DSSEff-values on the whole are higher within the first mentioned domain and the very highest differential standardized selection effect is found for Mach9 and Natmark. The last mentioned fact is quite interesting in the light of gender differences in variability. As shown by Table 1 both these variables have substantially greater variability for men than for women.

A further indication on the possible impact of gender differences in variability is found within the verbal domain where Swachr9 shows the greatest differential standardized selection effect. This was the very variable showing the greatest gender differences in variability within this domain. Are gender differences in variability the reason why these variables show the greatest DSSEff? This question will be answered when we now turn to the adjusted differential selection effects.

Table 4.
Adjusted selection effects (ASEff) and differential adjusted selection effects (DASEff) to the SweSAT.

Domain	Variable	$ASEff_f$	$ASEff_m$	DASEff
General	GSA	1.096	1.301	-0.205
	Testsum	0.849	0.968	-0.119
Natural science	NS6	0.749	0.893	-0.144
	Mach6	0.822	1.019	-0.197
	Mach9	0.908	1.162	-0.254
	Natmark	1.045	1.251	-0.206
Verbal	Op6	0.815	0.917	-0.102
	Swachr9	0.858	0.998	-0.140
	Swachw9	0.748	0.942	-0.194
	Verbmark	1.013	1.218	-0.205

A comparison between the DSSEff-values in Table 3 and the DASEff-values in Table 4 shows that taking into account the gender differences in variability really matters. For the only variable on which the females are more variable (Op6) the DASEff value exceeds that of DSSEff. In contrast, for all other variables showing a greater male variability the change goes in the opposite direction with lower DASEff-values. We can also see that the change is most pronounced for those variables which have shown the greatest gender differences in variability, i.e. Mach9, Natmark, and Swachr9. The last mentioned variable, for instance, had a DSSEff-value of 0.293 but by taking into account the differences in variability, the selection effects are reduced to 0.140, that is to say that, in this case, at least half the standardized differential selection effect can be explained by gender differences in variability.

However, the greater male variability cannot explain all the differential selection effects. As shown by Table 4, there are also 'pure' such effects which means that men are more influenced by their ability than are women when deciding on whether or not to take the SweSAT. Could this imply that the male SweSAT takers are of higher ability than are the female test takers? As mentioned before, such a conclusion is not justified only on the basis of the selection effects, but we have also to take into account the gender differences within the total group. We will return to that question.

We will finish this section by summarizing to what extent the selection effects are influenced by taking into account the gender differences in variability. This is done by showing the average standardized selection effects and the average adjusted selection effects for each domain and for all the ten variables taken together.

When the variables are combined within each domain the gender differences in variablity explain between 20 and 30 per cent of the differential standardized selection effects and the greatest influence is found for the natural science domain. However, still the greatest differential selection effects are found within this domain. Combining all variables as is done in the last line of Table 5 the greater variability among men explains 26 per cent of the differential standardized selection effects.

Table 5.
The average standardized selection effects and the average adjusted selection effects by gender and domain.

Domain	SSEff		DSSEff	ASEff		DASEff
	Females	Males		Females	Males	
General	0.953	1.161	-0.208	0.973	1.135	-0.162
Natural	0.840	1.128	-0.288	0.881	1.081	-0.200
Verbal	0.836	1.045	-0.209	0.859	1.019	-0.160
All	0.861	1.101	-0.240	0.890	1.067	-0.177

Thus, it is obvious that differences in variability should be taken into account when differential selection effects are studied. However, as shown in Table 5, the stronger selection effects among men are not only a consequence of greater variability. There are also 'pure' differential selection mechanisms and one such mechanism might be the previous educational career. To what extent this factor influences the differential selection to the SweSAT will be studied in the next section.

Selection effects to the SweSAT in relation to gender, ability, and upper secondary education

In this section the primary interest is to explain the causes of the selection effects. Do these selection effects occur at the time when the person decides on the SweSAT taking, or are they a consequence of previous selection effects within the educational system? In order to clarify this question the various selection effects are studied separately for those who have finished a theoretical upper secondary eduction and for those who have not.

Of all individuals in the sample 37 per cent have finished a theoretical upper secondary education and among them 61 per cent have taken the SweSAT. The frequency of SweSAT taking among those who have no such education is only 4 per cent. Thus, an overwhelming majority of the test takers have finished a theoretical programme of upper secondary school.

Table 6 shows those regression coefficients which determine the various selection effects to SweSAT in relation to gender, ability, and upper secondary education, and Table 7 presents the selection effects (SEff) for each subgroup and the differential selection effects (DSEff), which all are based on the regression coefficients shown in Table 6.

Table 6.
Regression coefficients for the analyses of differential selection effects to the SweSAT in relation to gender and upper secondary education.

Domain	Variable	SweSAT	USE x SweSAT	SEX x SweSAT	USE x SEX x SweSAT
		b_3	b_5	b_6	b_7
General	GSA	0.602*	-0.375*	-0.141	0.127
	Testsum	12.091*	-7.221*	-3.396	4.592
Natural science	NS6	5.021*	-2.777*	-1.791	2.050
	Mach6	4.498*	-2.030*	-1.364	1.129
	Mach9	14.053*	-5.671*	-5.347*	3.543
	Natmark	0.745*	-0.410*	-0.228*	0.202
Verbal	Op6	4.104*	-2.119*	-0.744	0.761
	Swachr9	12.481*	-8.863*	-5.958*	5.778*
	Swachw9	0.359*	-0.167*	-0.050	0.005
	Verbmark	0.738*	-0.448*	-0.297*	0.270*

*Significant on the 5% level

Table 7
Selection effects (SEff) and differential selection effects (DSEff) to the SweSAT among persons with different educational background.

Domain	Variable	No theoretical upper secondary education			Theoretical upper secondary education		
		$SEff_f$	$SEff_m$	DSEff	$SEff_f$	$SEff_m$	DSEff
General	GSA	0.461	0.602	-0.141	0.213	0.227	-0.014
	Testsum	8.695	12.091	-3.396	6.066	4.870	1.196
Natural science	NS6	3.230	5.021	-1.791	2.503	2.244	0.259
	Mach6	3.134	4.498	-1.364	2.233	2.468	-0.235
	Mach9	8.706	14.053	-5.347	6.578	8.382	-1.804
	Natmark	0.517	0.745	-0.228	0.309	0.335	-0.026
Verbal	Oppos6	3.360	4.104	-0.744	2.002	1.985	0.017
	Swachr9	6.523	12.481	-5.958	3.438	3.618	-0.180
	Swachw9	0.309	0.359	-0.050	0.147	0.192	-0.045
	Verbmark	0.441	0.738	-0.297	0.263	0.290	-0.027

As shown previously the SweSAT coefficients constitute a direct measure of the selection effects (SEff) to the SweSAT within the reference group i.e. *males without a theoretical upper secondary education*. According to Table 6 they are all positive and statistically significant. Therefore, we can conclude that there are strong positive selection effects to the SweSAT among those males who have no theoretical upper secondary education.

The SEXxSweSAT coefficients indicate the differential selection effects (DSEff) within the nontheoretical group and they are all negative although not always statistically significant. These negative coefficients imply that there is a general trend of females being less strongly selected than males to the SweSAT within the nontheoretical group. These effects are significant only for Mach9, Natmark, Swachr9, and Verbmark, however.

Even the USExSweSAT coefficients are negative and they are all statistically significant. This means that the selection effects among men become substantially weaker within the theoretical group as compared to the nontheoretical group. The same is true for women as well, but for them the change is less pronounced as shown by the positive USExSEXxSweSAT coefficients.

By summing the SEXxSweSAT and the USExSEXxSweSAT cofficients we obtain a measure of DSEff within the theoretical group and for most variables the two coefficients are of about the same magnitude but with different signs. This means that the differential selection effects are small within this educational group. For some variables the sum even reaches a positive value, which means that women are somewhat more strongly selected to the SweSAT than are men. However, on the whole the effects are so small that it seems justified to speak of no differential selection at all within the theoretical group.

Finally, it is also worth noting that the USExSEXxSweSAT coefficients constitute a direct measure of difference between the differential selection effects within the nontheoretical group and those within the theoretical group. Since they are all positive we can conclude that there is a general trend of weaker differential selection effects among those with a theoretical upper secondary education. Two variables indicate significant differences between the two educational groups in this respect, namely Swachr9 and Verbmark.

Thus, we can sum up the results in Tables 6 and 7:

- SweSAT takers in both educational groups constitute a positively selected group

- the selection effects are strongest among persons who have no theoretical upper secondary education

- within the nontheoretical group male test takers are more positively selected than are the female test takers, but not significantly so for all variables studied

- among those with a theoretical upper secondary education, the differential selection effects (DSEff) are throughout so small and of varying signs that there is no reason to speak of any differential selection effects

The last mentioned conclusion is interesting in the light of our previous finding, namely that the selection effects were stronger for male test takers than for female test takers when upper secondary education was not taken into consideration. In other words, keeping educational background constant these differential selection effects are removed for those with a theoretical upper secondary education. This implies that the differential selection effects to the SweSAT found for the total group must be mainly ascribed to those selection effects which are working when the individuals decide upon their upper secondary programme.

The selection effects discussed so far are directly influenced by varying standard deviations between variables and also by differences in variability between the various subgroups. As shown previously, these influences can be eliminated by taking the standard deviations into account. These are shown in Table 8.

Table 8
Standard deviations for the total sample and for different subgroups.

Domain	Variable	No theoretical upper secondary education		Theoretical upper secondary education		Pooled within-group stand. deviat.
		Females	Males	Females	Males	
General	GSA	0.615	0.579	0.424	0.424	0.538
	Testsum	16.526	16.657	13.160	13.038	15.362
Natural science	NS6	7.485	7.958	6.807	6.756	7.390
	Mach6	6.343	6.541	5.605	5.908	6.193
	Mach9	17.045	17.986	14.157	14.126	16.308
	Natmark	0.794	0.795	0.647	0.674	0.748
Verbal	Op6	5.772	5.436	5.133	4.831	5.376
	Swachr9	13.131	15.160	8.296	8.396	12.330
	Swachw9	0.655	0.601	0.537	0.625	0.608
	Verbmark	0.727	0.699	0.606	0.630	0.678

Those who have chosen a theoretical upper secondary education constitute a more homogenous group with respect to practically all variables in comparison with those who have no such education. This result is expected considering the fact that the students in theoretical upper secondary programmes normally have been selected on the basis of their leaving certificates from compulsory school. This selection process also implies mean differences between the two educational groups, and these differences cause the pooled within-group standard deviations to be lower in Table 8 than those in Table 1 where upper secondary education was not taken into account.

The gender differences in variability within the two educational groups are on the whole small and in some cases the males are more variable and in other cases the females have a higher standard deviation. There are two exceptions, however, Swachr9 among those who have no theoretical upper secondary education, and Swachw9 among those with such an education. In both cases the males are substantially more variable than are the females.

In Table 9 we show the standardized selection effects and the differential standardized selection effects to the SweSAT. This table shows that among those who have no theoretical upper secondary school most DSSEff values are between 0.1 and 0.5 with a highest value for Swachr9 and a lowest one for Swachw9. Within the theoretical group only one DSSEff reaches the value of 0.1, namely that for Mach9.

Table 9.
Standardized and differential standardized selection effects (SSEff and DSSEff) to the SweSAT among persons with different upper secondary education.

Domain	Variable	No theoretical upper secondary education			Theoretical upper secondary education		
		$SSEff_f$	$SSEff_m$	DSSEff	$SSEff_f$	$SSEff_m$	DSSEff
General	GSA	0.857	1.119	-0.262	0.396	0.422	-0.026
	Testsum	0.566	0.787	-0.221	0.395	0.317	0.078
Natural science	NS6	0.437	0.679	-0.242	0.339	0.304	0.035
	Mach6	0.506	0.726	-0.220	0.361	0.399	-0.038
	Mach9	0.534	0.862	-0.328	0.403	0.514	-0.111
	Natmark	0.691	0.996	-0.305	0.413	0.448	-0.035
Verbal	Op6	0.625	0.763	-0.138	0.372	0.369	0.003
	Swachr9	0.529	1.012	-0.483	0.279	0.293	-0.014
	Swachw9	0.508	0.590	-0.082	0.242	0.316	-0.074
	Verbmark	0.650	1.088	-0.438	0.388	0.428	-0.040

When we now turn to the question about the effects of previous education we will present only the average SSEff and DSSEff for each domain and for all the variables. In order to facilitate this comparison we also give the DSSEff values for the total group.

Table 10.
Average standardized selection effects (SSEff) and differential standardized selection effects (DSSEff) for the two educational groups and for the total group.

Domain	No theoretical upper secondary education			Theoretical upper secondary education			Total group
	$SSEff_f$	$SSEff_m$	DSSEff	$SSEff_f$	$SSEff_m$	DSSEff	DSSEff
General	0.712	0.953	-0.241	0.396	0.370	0.026	-0.208
Natural	0.542	0.816	-0.274	0.379	0.416	-0.037	-0.288
Verbal	0.578	0.863	-0.285	0.320	0.352	-0.032	-0.209
All	0.590	0.862	-0.272	0.359	0.381	-0.022	-0.240

Taking upper secondary education into account implies a substantial decrease of the standardized selection effects and particularly so among those who have finished a theoretical upper secondary education (cp. Table 3). Even if the standardized selection effects have decreased, the differential standardized effects remain at least as high within the nontheoretical group as for the total group. The DSSEff values for the theoretical group, on the other hand, are all much lower.

Dividing the group according to previous education chosen also implies that each subgroup becomes more homogenous and we have also shown that this effect is most pronounced for the theoretical group. Thus, differences in variability ought to be one reason why the nontheoretical group shows the highest SSEff values. On the other hand, there are only small gender differences in variability within each educational group. Therefore, we cannot expect any great changes in the differential selection effects when we adjust for differences in variability.

In Table 11 the adjusted selection effects are shown for each subgroup and each variable. These values are summarized as means for each domain in Table 12 and in order to show the effect of taking previuos education into account the corresponding values for the total group are given, as well.

Table 11.
Adjusted and differential adjusted selection effects (ASEff and DASEff) to the SweSAT among persons with different upper secondary education.

Domain	Variable	No theoretical upper secondary education			Theoretical upper secondary education		
		$ASEff_f$	$ASEff_m$	DASEff	$ASEff_f$	$ASEff_m$	DASEff
General	GSA	0.750	1.040	-0.290	0.502	0.535	-0.033
	Testsum	0.526	0.726	-0.200	0.461	0.374	0.087
Natural science	NS6	0.432	0.631	-0.199	0.368	0.332	0.036
	Mach6	0.494	0.688	-0.194	0.398	0.418	-0.020
	Mach9	0.511	0.781	-0.270	0.465	0.593	-0.128
	Natmark	0.651	0.937	-0.286	0.478	0.497	-0.019
Verbal	Op6	0.582	0.755	-0.173	0.390	0.411	-0.021
	Swachr9	0.497	0.823	-0.326	0.414	0.431	-0.017
	Swachw9	0.472	0.597	-0.125	0.274	0.307	-0.033
	Verbmark	0.607	1.056	-0.449	0.434	0.460	-0.026

Table 12.
Average adjusted selection effects (ASEff) and differential adjusted selection effects (DASEff) for the two educational groups and for the total group.

Domain	No theoretical upper secondary education			Theoretical upper secondary education			Total group
	$ASEff_f$	$ASEff_m$	DASEff	$ASEff_f$	$ASEff_m$	DASEff	DASEff
General	0.638	0.883	-0.245	0.482	0.454	-0.028	-0.162
Natural	0.522	0.759	-0.237	0.427	0.460	-0.033	-0.200
Verbal	0.540	0.808	-0.268	0.378	0.402	-0.024	-0.160
All	0.552	0.803	-0.251	0.418	0.446	-0.028	-0.177

After adjusting for differences in variability the ASEff values are more equal for the two educational groups. Thus, the greater homogenity of the theoretical group is one contributing factor to the finding that the selection effects are substantially weaker among those who have finished a theoretical upper secondary education. However, the greater homogenity is not the only explanation. Even when the differences in variability are taken into account there still are substantially weaker selection effects within the theoretical group.

As expected, adjusting for gender differences in variability has only small effects. Still the males are more positively selected to the SweSAT than are the females among those who have not finished a theoretical upper secondary education and within the theoretical group the differential selection effects remain very low.

The results in this section can be summarized in the following way:

- Irrespective of upper secondary education those who have taken the SweSAT constitute a highly positively selected group.

- However, the selection effects are substantially reduced by taking upper secondary education into account which implies that the selection mechanisms to upper secondary school are important causes to the selection effects to the SweSAT found for the total group.

- Among those who have finished a theoretical upper secondary education the strength of the selection effects is equal for male and female test takers, and therefore, the differential selection effects are small.

- Within the nontheoretical group males remain more strongly selected to the SweSAT than females.

- Since there are practically no gender differences in variability when previous education is controlled for this factor is of minor importance to the differential selection effects for males and females.

- The test takers within the nontheoretical group are more strongly selected than are those in the theoretical group.

- To some extent, this is an effect of a greater variability within the nontheoretical group as compared to the theoretical one.

DISCUSSION

Like many other admission tests for entrance into higher education the Swedish Scholastic Aptitude Test (SweSAT) shows substantial gender differences in favour of male test takers. However, these differences attracted little attention in the public debate prior to 1991, when the admission rules were changed and all applicants could be admitted on the basis of their SweSAT scores. Since then, however, the gender differences have been the subject of an intense debate and mostly they have not been regarded as 'real' but as 'artifacts' to use the terms of Wilder and Powell (1989). The main explanation proposed is test bias in favour of men. Little attention has been paid to the fact that the test takers constitute self selected groups and that there might be differential selection effects in relation to gender - a fenomenon that has been pointed out in many previous studies (Fennema & Sherman, 1977b; Hyde & Linn, 1988; Rosenthal & Rubin, 1982; Gustafsson, Reuterberg & Westerlund, 1992; Mäkitalo, 1994).

In order to study such differential selection effects we need information on large and representative samples. In the Swedish databank ETF (Evaluation Through Follow-up) this kind of data is available. However, even with that kind of data at hand there are methodological problems to be solved. One such problem is that of differences in variability. Earlier studies of gender differences in variability have shown greater varibility among men i mathematical and spatial abilities as well as on standardized aptitude tests (Maccoby & Jacklin, 1974; Feingold, 1992a). As a consequence the mean differences between boys and girls differ between different parts of the distribution. Within a positively selected group, for instance, the greater male variability will cause mean differences in favour of males even if there are no gender differences at all within the unselected group (Cleary, 1991).

Since we are studying the selection effects to an admittance test to higher education, we have to do with a positively selected group. Furthermore, all variables included in this study have shown gender differences in variability, males being the more variable group on nine out of ten variables studied. Therefore, gender differences in variability have to be taken into account.

Quite naturally, the results show that the SweSAT takers constitute a positively selected group according to all variables studied. There are also substantial differential selection effects, male test takers being more positively selected than are female test takers. The differences between test takers and others vary between 0.7 and 1.1 units of a standard deviation among the females and between 0.9 and 1.3 units among the males.

Taking the gender differences in variability into account does not change the general conclusion of men being more positively selected, but the differential selection effects decrease by more than 25 per cent. The average selection effect as expressed by Differential Standardized Selection Effects (DSSEff) namely amounts to -0.245, and the average adjusted effect as expressed by the Differential Adjusted Selection Effects (DASEff) to -0.177.

Thus, this result gives strong support to Cleary's conclusion that differences in variability should be taken into account when making gender comparisons based on self selected groups (Cleary, 1991). Furthermore, the results also support Cleary's finding that the differential selection effects are greater within the Natural science domain as compared to the Verbal domain. To some extent this is caused by the fact that the gender differences in variability are greatest within the first mentioned domain. Adjusting for these differences implies a reduction of the differential selection effects by approximately one third in the Natural science domain and by about 25 per cent in the Verbal domain, but within both domains the male test takers are more strongly selected to the SweSAT.

The general conclusions to be drawn from these results are that

- the SweSAT takers constitute a strongly selected group and therefore, generalizations cannot be made from the test takers to all males and females respectively

- the male test takers are more positively selected than are the female test takers and especially so on variables within the Natural science domain

- to a substantial extent this is due to the fact that males are more variable with respect to most of the variables studied. However, differences in variability do not explain all the differential selection effects. Thus, there are also 'pure' differential selection effects.

These conslusions give rise to the following two questions:

-Are the differential selection effects the cause of the gender differences in favour of men on the SweSAT scores?

-What are the causes of the differential selection effects?

It is not possible to address the first question exhaustively in this paper, since it is connected with complicated methodological problems. Still, it is worthwhile to touch upon the question, because the differential selection effects do not tell us anything about the absolute differences between male and female test takers, but only the relative differences. If there are large gender differences in the unselected group the less positively selected group of test takers might still outperform the more positively selected group.

Table 13 shows the mean differences between females and males among the test takers and within the total group on all variables studied.

Table 13.
Mean differences expressed as z-values between females and males for the SweSAT takers and for the total sample. Positive values indicate female superiority.

Domain	Variable	SweSAT takers	Total sample
General	GSA	0.15	0.41
	Testsum	-0.15	0.01
Natural science	NS6	-0.25	-0.12
	Mach6	-0.34	-0.10
	Mach9	-0.45	-0.16
	Natmark	-0.14	0.16
Verbal	Op6	0.03	0.12
	Swachr9	0.02	0.30
	Swachw9	0.42	0.64
	Verbmark	0.29	0.54

The differential selection effects to the SweSAT have substantially reduced the females' superiority in GSA, and for Testsum, where there are practically no gender differences in the total sample, male test takers excel.

Within the Natural science domain the females in the total sample have a higher mean on only one variable, Natmark. However, the differential selection effects turn this difference into the opposite direction among the SweSAT takers. As to the other variables in this domain, the differential selection effects reinforce the total gender differences in favour of males so they become substantial among the test takers.

In the verbal domain, finally, there are differences in favour of the females within the total group, but these differences are diminished by the selection effects. However, still the female test takers have slightly higher means than the male test takers.

The results in Table 13, thus indicate that, on the mathematical part of the SweSAT the great differences in favour of male test takers are to at least some extent explained by initial gender differences within the Natural science domain among the test takers. Also, on the verbal part of the SweSAT the score differences are in favour of male test takers, but this is not to be expected taking into account the initial gender differences among those who take the SweSAT. Thereforere, it seems likely that the SweSAT favours the male test taker unduly on its verbal parts.

It is obvious that previous educational choices play an important role for the differential selection. Among those who have finished a theoretical upper secondary programme there are practically no differential selection effects to the SweSAT. Furthermore, there are only small gender differences in variability within this group. Thus, these results indicate that the differential selection effects to the SweSAT found for the total group mainly are caused by differential selection effects not to the test itself but to upper secondary education.

Among those rather few test takers, who have not finished a theoretical upper secondary education, there are still strong differential selection effects to the SweSAT - the males being more positively selected. Even if a great deal of these effects can be ascribed to gender differences in variability, the adjusted differential selection effects, too, are substantial. There might be several reasons for this result. One such reason is that we based the definition of upper secondary education on the programme which the individual have finished. When our data were collected, there were certainly a number of individuals in the nontheoretical group who were still following a theoretical programme of upper secondary school and they are probably overrepresented among the SweSAT takers.

Consequently, there are probably some selection effects to upper secondary education which in this study have emerged as selection effects to the SweSAT for the nontheoretical group.

Another reason for the strong differential selection effects within the nontheoretical group may be that there are gender differences in educational choices in secondary school. For instance, it has been shown that many high achieving girls enter the more vocationally oriented upper secondary programmes (Härnqvist & Svensson, 1980; Arnman & Jönsson, 1986). Since the transition rate to higher education is lower from these programmes as compared to that from the theoretical ones, many high achieving females from the nontheoretical programmes do not take the SweSAT. Furthermore, there are indications that females more often than males apply for higher education only on the basis of their leaving certificates from upper secondary school (Forneng & Jansson, 1991) and this ought to hold above all for the high achieving females. The effect of this selection to the SweSAT is shown by Mäkitalo (1994) who found that the probability of taking the SweSAT is substantially higher among males than among females on the high GSA levels, while there are small gender differences on the medium and low levels.

The differential selection effects to upper secondary school in relation to gender is a most interesting result in itself. It implies, namely, that these effects must be taken into consideration when making any kind of gender comparison on this educational level, and as far as we know, this is not any common practice.

As to the SweSAT, there are several important research areas in connection with differential selection. One such area is to study these effects in relation to socioeconomic background. The socioeconomic differences in SweSAT scores were a main issue at the time when the SweSAT was introduced in Sweden, but during the last years it has attracted little attention. The main question, however, is that of bias in SweSAT scores. Are certain groups of test takers unduly favoured by the test, or can group differences in SweSAT scores be explained by initial ability differences in combination with differential selection effects?

REFERENCES

Arnman, G., & Jönsson, I. (1986). *Olika för olika. Aspekter på svensk utbildningspolitik. (Different for different pupils. Aspects on Swedish educational policy).* Lund: Arkiv.

Becker, B. J. & Hedges, L. V. (1988). The effects of selection and variability in studies of gender differences. *Behavioral and Brain Sciences, 11,* pp 183-184.

Chipman, S.F., & Thomas V.G. (1985) Womens participation in mathematics: Outlining the problem. In Chipman et al (Eds.), *Women and mathematics: Balancing the equation.* New Jersey: LEA.

Chodorow, N. (1978). *The reproduction of mothering: psychoanalysis and the sociology of gender.* Los Angeles: University of California Press.

Cleary, A. (1991). Gender differences in aptitude and achievement test scores. In J.Pfeiderer (Ed.), *Sex equity in educational opportunity, achievement, and testing. Proceedings of the 1991 ETS invitational conference, pp 51-90*. Princeton, NJ: Educational Testing Service.

Eagly, A. (1987). *Sex differences in social behavior: A social role interpretation.* New Jersey: LEA

Feingold, A. (1992). Sex differences in variability in intellectual abilities: A new look at an old controversy. *Review of Educational Research.* 62, pp 61-84.

Feingold, A. (1994). Gender differences in variability in intellectual abilities: A cross-cultural perspective. *Sex roles* 30, pp 81-92.

Fennema, E., & Sherman, J. (1977a). Sex-Related Differences in Mathematics Achievement, Spatial Visualization and Affective Factors. *American Educational Research Journal,* 14, pp 51-71.

Fennema, E., & Sherman, J. (1977b). The Study of Mathematics By High School Girls and Boys: Related Variables. *American Educational Research Journal,* 14, pp 159-168.

Florin, C., & Johansson, U. (1993). *"Där de härliga lagrarna gro..." Kultur, klass och kön i det svenska läroverket 1850-1914. ("Where the glorious laurels grow..." Culture, class and gender in the Swedish secondary grammer school 1850 - 1914)*. Kristianstad: Tidens förlag.

Forneng, S., & Jansson, H. (1991). Betyg och prov. Antagningen till högskolan 1991. (Marks and tests. The admission to higher education in 1991). *UHÄ-rapport* 1992:25.

Franke-Wikberg, S. (1981). *Arbete, familj och utbildning. Bakgrundsanalyser för en analys av skolans studie- och yrkesorientering. (Work, family and education. Background analyses for an analysis of educational and vocational guidance in school)*. Stockholm: Liber Förlag.

Gilligan, C. (1982). *In a different voice: Psychological theory and womens development.* Cambridge: Harvard University Press.

Gustafsson, J.-E., Reuterberg, S.-E., & Westerlund, A. (1992). Beror gruppskillnaderna på tidigare selektion i utbildningssystemet? (Are the group differences due to earlier selection in the educational system?). *UHÄ-rapport* 1992:3.

Halpern, D. (1986). *Sex differences in cognitive abilities.* New Jersey: LEA Publishers.

Henriksson, W., Stage, C., & Lexelius, A. (1986). Samma uppgift men olika innehåll - en studie av NOG-provet. (The same item but a different content - a study of the Data Sufficiency subtest). *Pedagogiska institutionen, Umeå universitet.*PM 10.

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

Humphreys, L. G. (1988). Sex differences in variability may be more important than sex differences in means. *Behavioral and Brain Sciences,* 11, pp 195-196.

Hyde, J.S., Fennema, E., & Lamon, S.J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin,* 104, pp 153-169.

Hyde, J.S., & Linn, M.C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin,* 104, pp 153-169.

Härnqvist, K., Emanuelsson, I., Reuterberg, S.-E., 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.

Härnqvist, K., & Svensson, A. (1980). *Den sociala selektionen till gymnasiestadiet. (The social selection to upper secondary school). Stockholm: Statens offentliga* utredningar 1980:30.

Ingerskog, P., & Stage, C. (1993). Högskoleprovet hösten 1993. Provdeltagargruppens sammansättning och resultat. (The Swedish Scholastic Aptitude Test in the autumn of 1993. The characteristics and the results of the group of test takers). *Pedagogiska institutionen,Umeå universitet.* PM 79.

Lenney, E. (1981). What's fine for the gander isn't always good for the goose: sex differences in self-confidence as a function of ability area and comparison with others. *Sex Roles,* 7, pp 905-924.

Lorber, J., & Farell, A. (1991). *The social construction of gender.* Newbury Park: SAGE.

Messick, S. (1976). *Individuality in learning.* San Fransisco, CA: Jossey-Bass Publishers.

Mäkitalo, Å. (1993). Gender differences in performance on the DTM subtest of the Swedish Scholastic Aptitude Test as a function of item position and cognitive demands. Submitted for publication in Scandinavian Journal of Educational Research.

Mäkitalo, Å. (1994). Non-comparability of female and male admission test takers. *Department of Education and Educational Research, Göteborg University.* Report 1994:06.

Reuterberg, S.-E. (1994). Selection of Swedish Scholastic Aptitude Test takers as a function of social background and ability. *Department of Education and Educational Research, Göteborg University.* Report 1994:05.

Rosenthal, R., & Rubin, D. B. (1982). Further Meta-Analytic Procedures for Assessing Cognitive Gender Differences. *Journal of Educational Psychology,* Vol. 74, No. 5, pp 708-712.

Stage, C. (1985). *Gruppskillnader i provresultat. Uppgiftsinnehållets betydelse för resultatskillnader mellan män och kvinnor på prov i ordkunskap och allmänorientering. (Group differences in test scores. The impact of item content on differences in scores between men and women on the Word and the General Information subtests).* Akademisk avhandling nr 17. Pedagogiska institutionen, Umeå universitet.

Stage, C. (1987). Analys av NOG-uppgifter med avseende på könsskillnader i resultat. (Analysis of Data Sufficiency items with respect to gender differences in scores). *Pedagogiska institrutionen, Umeå universitet.* PM 12.

Stage, C. (1988). Gender Differences in Test Results. *Scandinavian Journal of Educational Research,* 36, pp 101-111.

Stage, C. (1990). Könsskillnader i resultat på högskoleprovet våren 1990.(Gender differences in the SweSAT scores in the spring of 1990). *Pedagogiska i nstitutionen, Umeå universitet.* PM 42.

Stage, C. (1992). Högskoleprovet våren 1992. Provdeltagargruppens sammansättning och resultat. (The SweSAT in the spring of 1992. The characteristics and the results of the group of test takers). *Pedagogiska institutionen, Umeå universitet.* PM 63.

Statistics Sweden (1993). *Trender och prognoser '94. Befolkningen, utbildningen, arbetsmarknaden. (Trends and forecasts '94. Population, education and labour market in Sweden).* Stockholm: Statistics Sweden.

Swedish Ministry of Education and Science (1992). *The Swedish Way Towards a Learning Society.* Stockholm: Ministry of Education and Science.

Wainer, H. (1986a). The SAT as a Social Indicator: A Pretty Bad Idea. In H. Wainer, (Ed), *Drawing Inferences from Self-selected Samples.,* pp 7-21. 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 22-24.

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.

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

Wernersson, I. (1977). *Könsdifferentiering i grundskolan. (Gender differentiation in the compulsory school).* Göteborg: Acta Universitatis Gothoburgensis.

Wernersson, I. (1988). Olika kön samma skola? En kunskapsöversikt om hur elevernas könstillhörighet påverkar deras skolsituation. (Different gender the same school? A review of research on the impact of gender on the pupils' school situation). *Skolöverstyrelsen, Vad säger forskningen.* F 89:1.

Wernersson, I. (1991). Könsskillnader i gymnasieskolan. En kunskapsöversikt. *(Gender differences in upper secondary school. A review of research).*

Skolöverstyrelsen, Vad säger forskningen F 91:2.

Wester, A. (1992a). Item Bias with Respect to Gender Interpreted in the Light of Problem-Solving Strategies. *Department of Education, University of Umeå.* EM 3.

Wester-Wedman, A. (1992b). Ett försök med öppna frågor i DTK-provet. En jämförelse mellan öppna frågor och flervalsfrågor avseende könsskillnaden i prestation på DTK-provet. (An experiment with open-ended questions in the DTM subtest. A comparison between open-ended questions and multiple choice questions with respect to the gender difference in performance on the DTM subtest). *Pedagogiska institutionen, Umeå universitet.* PM 54.

Wester-Wedman, A. (1992c). Förlängd provtid på DTK-provet. En studie av effekten av förlängd provtid på könsskillnaden i prestation på DTK-provet. (Prolonged testing time on the DTM subtest. A study of the impact of prolonged testing time on the gender difference in performance on the DTM subtest). *Pedagogiska i nstitutionen, Umeå universitet.* PM 56.

Wice, L. L. (1985) Project talent: Mathematics course participation in the 1960 and its career consequences. In Chipman et al (Eds.), *Women and mathematics: Balancing the equation*. New Jersey: LEA.

Wilder, G., & Powell, K. (1989). Sex Differences in Test Performance: A Survey of the Literature. Collage Board Report. No 89-3

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

---

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. & Csocsán, 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.

Mäkitalo, Åsa. Non-comparability of female and male admission test takers. 1994:06.

Rosén, M. Gender differences in hierarchically ordered ability dimensions: the impact of missing data. 1995:01.

Holmberg, L.M. The design of a learning environment for users of an accounting information system. Theoretical foundations and an empirical study. 1996:01.

Holmberg, L.M. Design of learning environments for IT-users. 1996:02.

Mäkitalo, Å & Reuterberg, S-E. Who takes the Swedish scholastic aptitude test? A study of differential selection to the SweSAT in relation to gender and ability. 1996:03.