Mathematics and Language Skills and the Choice of Science Subjects in Secondary Education*

D. Uerz1, H. Dekkers2, and A.A. Béguin3
1Institute for Applied Social Sciences (ITS), Nijmegen, The Netherlands,
2Department of Education, University of Nijmegen, The Netherlands, and
3Department of Organization and Management, Educational Science and Technology,
University of Twente, Enschede, The Netherlands/Citogroup Arnhem, The Netherlands

ABSTRACT

A growing shortage of technical and science graduates means that the question of determinants of differences in performance and choices in that field based, among other things, on social background characteristics (gender, socioeconomic status, ethnic origin, family composition) remains a topical one. In this study, a rational choice model is proposed in which background-specific comparative advantages in different subjects are in focus. Secondary analyses of a large-scale cohort study (VOCL) show that in addition to the most important explanatory variable, namely gender, the gap between mathematical and language skills makes a relevant contribution in predicting the choice of science subjects in secondary education. This gap can, in turn, be predicted on the basis of gender and family composition.

INTRODUCTION

In recent years, there has been a growing shortage of technical and science graduates on the labor market, not only in The Netherlands but also throughout the industrial and postindustrial world. This shortage is on the one hand caused by society becoming increasingly a technological one, thereby generating the need for more technically trained workers (Sociaal

*An earlier version of this article was presented at the Annual Meeting of the American Educational Research Association, April 1999, in Montreal.
Address correspondence to: Dana Uerz, Institute for Applied Social Sciences (ITS), University of Nijmegen, PO Box 9048, 6500 KJ Nijmegen, The Netherlands. E-mail: D.Uerz@its.kun.nl
Accepted for publication: December 5, 2002.
Economische Raad [SER], 1995), and on the other hand by a diminishing interest for technical and science subjects among secondary school students. In order to meet this shortage, both governments and industries largely carry out measures or policies on an ad hoc basis. At the same time, there is a need to focus on potential technical and science graduates, that is on students who are likely to choose these fields. There is obvious potential among girls and women since differences in choosing subjects are primarily based on gender. Recently, possible reserves of technological potential among ethnic minorities were explored, also in The Netherlands, since differences based on ethnic origin have also been established. Other background variables like socioeconomic status (SES) and family composition were studied to see whether they – possibly interactively – influenced the choice of technical and science subjects.

These investigations obviously also led to the question which factors play a role in linking gender, ethnic origin, SES, and family composition with science and technical choices. Several individual, family-, and school-related factors were looked at, mainly in gender studies. An important determinant was shown to be students’ achievement scores in mathematics. Because it is apparent that – besides gender – socioeconomic and ethnic background characteristics of students are also related to the choice of science subjects, it is useful to study the influence of other, more culturally based indicators, like language attainment, in linking the aforementioned background variables to the choice of science subjects in secondary school.

In Rational Choice explanations (see Coleman, 1990) of gender, socioeconomic, and ethnical differences in the choice of science subjects, previous achievements are considered to be of great influence on the actual subject choice. Not only the absolute achievements per subject, but especially the relative achievements in one subject compared to another are of interest. These relative advantages in subjects have been hypothesized to influence subject choices through their effects on the expected probabilities of success in the different fields of study.

These explanations emphasize the aspect of individual choice in the actual choice of final examination subjects in secondary education. Besides the earlier mentioned group effects of gender, SES, ethnic origin, and family composition, the choice of examination subjects is then influenced by a process of individual rational choice in which previous achievements, and especially relative achievements in one subject compared to another, play an important role.
In this study, we therefore examine both the aforementioned group effects (gender, SES, ethnic origin, and family composition) on the choice of science subjects in secondary education, and the effect of comparative advantage in mathematics. In the research model, especially the relative achievements in one subject compared to another are in focus; we have replaced the absolute mathematic attainment score as predictor of subject choice by a score representing the comparative advantage in mathematics, namely the gap between the mathematics and language score.

The databases of the Dutch VOCL’89 cohort (Secondary Education Student Cohort, 1989) provide the opportunity to test hypotheses on differences in participation in science subjects between groups of students with a combination of the four previously mentioned background variables (gender, socioeconomic status, ethnic origin, and family composition). The databases also contain data to analyze the correlation of these variables with the difference between mathematics and language achievement rather than just the two individual measurements for language and mathematics.

LITERATURE REVIEW

Until the late 1980s, gender inequality played a central role in research into the choice of, and achievement in, science subjects at secondary school level. Both in The Netherlands and abroad, much has been published on gender differences concerning school career. It is also well documented that, while large differences exist between various subgroups, the choices and achievement levels of secondary school students significantly influence the ultimate choice of a tertiary education in science for all groups (Dekkers, 1993; Maple & Stage, 1991). In this theoretical overview, we summarize in what way achievement levels differ, not only according to gender, but also according to students’ ethnic origin, socioeconomic status, and family composition.

The Third International Mathematics and Science Study (TIMSS) demonstrated that attainment in and attitudes towards science subjects correlate highly with gender. Boys do better in physics and chemistry and also enjoy these subjects more than girls. However, for mathematics the differences in attainment and attitudes among boys and girls have declined in recent years (Bos, Knuver, & Kuiper, 1997).

In a longitudinal study of determinants for choosing science subjects, it was found that for both girls and boys the same factors were involved in deciding
to choose these subjects (Dekkers, 1993). In particular, students’ aptitude (according to relevant others), enjoyment, interest, and the level of easiness of a subject were linked to choice of subjects. Thus, the factors involved are not different for boys and girls, but the relations are found on a different attainment level for boys and girls. Boys have higher scores on the determinants for mathematics and physics, and girls on those for French.

In a survey by Johnson (1996), a review was given of several large-scale national and international educational assessment programs in which, among other things, gender differences were revealed. From the description of the findings of these assessments we can learn that, despite large differences in population, design, and research method, there were many similarities regarding final conclusions in relation to gender differences in attainment and attitude towards various subjects. The overall conclusion is that girls achieve better than boys in reading and writing in primary school, but these differences disappear during secondary education. Boys score better than girls, on average, in mathematics and science, an edge that only increases during the school career. Interest in and attitudes to these subjects correspond with the score results.

From 1990 onwards, studies of differences in school careers – partly influenced by the (impending) shortage of science and technical personnel – have looked increasingly at the position of sociocultural groups. Gender-based differences in the choice of science subjects and in attainment in these subjects were increasingly studied in relation to socioeconomic and sociocultural background variables. For instance, the review of Maple and Stage (1991) describes the extent to which the choice of a science training is influenced by students’ background variables (gender, SES, and ethnic origin), students’ aptitude, and previous experiences and choices at the secondary level. As we have mentioned, while large differences exist between the various subgroups, secondary school choices and achievement have a significant influence for all groups on the later choice for a science training in tertiary education. Moreover, Ainley and Daly (1997) reported on a study in which socioeconomic and ethnic origin, as well as gender, were examined in relation to the choice of mathematics and attainment levels therein. After correction for previous achievements, gender proved to be the only significant variable. Other student background variables (like educational level of parents and ethnic origin) only indirectly – that is via earlier achievement scores – influenced the choice of mathematics and achievement levels at secondary school level.
In a follow-up study of the student cohort included in the study “School Matters,” Sammons (1995) examined the size and stability of gender- and ethnic-based and socioeconomic differences in (absolute) educational achievement and the (relative) progress in reading and mathematics from primary school until the end of secondary schooling. Results show that gender-based and socioeconomic differences in achievement at secondary level remained consistent for both subjects or even increased, thus mirroring differences between the groups (boys and girls and higher and lower social classes) already found at primary level. Patterns of ethnic differences, however, changed considerably during this period, both in absolute attainment as well as relative progress. While ethnic minorities showed less progress in primary school, the situation was reversed in secondary education. By the end of secondary school at age 16, ethnic minorities turned out to have progressed more quickly during this period and generally obtained better results in both subjects than their indigenous peers.

There is little Dutch research into the precise relationship between gender and ethnic origin and various aspects of school careers. Where that does occur, research is mainly restricted to certain aspects of a school career like dropping out or grade retention. In the cohort studies of the national evaluation of the Dutch Education Priority Policy (EPP), the progress of pupils through primary and secondary education has been followed since 1988. Findings from these studies which are important here may be summed up as follows: Girls achieve lower mathematics scores than boys, especially in the early secondary school years; these gender differences were found in all ethnic groups, but were largest in the immigrant population (Driessen & Dekkers, 1997; Suhre, De Wit, & Mulder, 1996). It is also often assumed that differences between immigrant and indigenous students regarding the choice of science and technical subjects and achievement levels, are irrelevant when socioeconomic differences are taken into account (De Fraiture, Dronkers, & Van Erp, 1997). Various other studies (i.e., Van ’t Hof & Dronkers, 1994; Wolbers & Driessen, 1996) show that differences in school success between the main ethnic groups in secondary education can be explained using the same factors as for Dutch students. For these students, too, socioeconomic status, based on the education and professional background of parents, is the most important factor for a successful school career.

In a follow-up study to the Secondary Education Cohort (VOCL), in which the effects of the background variables gender, social class, and ethnic background on school success were examined, it appears, however, that
complex interactive effects do exist between the three factors (Dekkers, Bosker, & Driessen, 2000). Findings show that school success cannot be simply predicted by expected additive or multiplicative effects of the stated background variables. For instance, the situation for immigrant girls is better than expected. While they do choose science and technical subjects less often, they generally perform as well on these subjects as Dutch girls. Dutch boys from low social backgrounds, however, do not achieve as well as expected. After 6 years of secondary schooling, they have fallen behind compared to other groups, even compared to immigrant boys from the same socioeconomic background.

In their study of the Dutch Primary Education Cohort (PRIMA), De Fraiture et al. (1997) looked at the differences among groups regarding the gap between arithmetic and language attainment in primary school. Gender and ethnic background were the most important determinants. Parental educational level only explained a scant percentage of the gap between arithmetic and language attainment. Family composition also appeared to be an explanatory factor. As the findings of McNab and Murray (1985) on the influence of family composition on mathematics achievement showed, the presence of a father in a family seems to stimulate the arithmetic skills of both sons and daughters. It was concluded that language and arithmetic attainment are influenced by various factors. Arithmetic is far less a culturally linked skill than language, so that the score-gap between language and arithmetic differs for various groups of pupils. The gap between language and arithmetic scores depends on the degree of integration within a certain society, the mother’s educational level, and the presence of a father in the family.

The question arises as to how far this difference between language and arithmetic scores indicates later preferences for languages, humanities, science, and technical training programs. When making educational choices, various alternatives are considered instead of choosing one specific direction.

One specific group of theories that attempts to explain choice consists of theories of rational choice. Central point in these theories is that, based on his (or her) preferences, an individual chooses to pursue that alternative, that course of action, which he believes to maximize his utility, his or her well-being (e.g., Hargreaves Heap, Hollis, Lyons, Sugden, & Weale, 1992). For students choosing school subjects this would mean that they choose those subjects that maximize their utility (Jonsson, 1999).

An important factor in the weighing of several alternatives, is the student’s subjective probability of success (e.g., Eccles et al., 1985; Jonsson, 1999).
One of the factors influencing the probability of success is the student’s previous achievement; low achievement in a subject means a low probability of success in that subject. These previous achievements can be divided into absolute achievements per subject, and relative achievements in one subject compared to another. Jonsson (1999, p. 395) states

it is reasonable to assume that an individual, in addition to considering his or her ability in absolute sense, prefers to specialize in a subject area in which his or her comparative advantages are high, e.g. in the subject for which they obtain their highest school marks.

When investigating study choice, it is therefore inaccurate to focus on the attractiveness of one of the alternatives (i.e., a science or language course). Both alternatives should be analyzed at the same time – the reason why in this study the relation between mathematics and language aptitude plays a central role. We analyze this relation as the difference between language and mathematics achievement (scores).

**RESEARCH QUESTIONS**

The aforementioned differences in language and mathematics achievement and the gap score between these two subjects lead to two research questions related to the size of the gap scores in secondary education for the various (sub)groups and the link these have with science choices. The gap score is thus the extent to which a score is higher for mathematics than for language (since opting for science is predicted), irrespective of the score level. The score can either be a positive or a negative number; it is negative when the language score is higher than the math score.

The research questions are:

1. Are there differences in the mathematics-language gap score (i.e., higher scores for mathematics than for language) in the 3rd year of secondary school in relation to gender, ethnic background, SES, and family composition or to combinations of these variables?
2. To what extent do differences in scores between mathematics and language in the 3rd year of secondary school influence the choice of subjects in secondary education. And is effect also found after differences in background variables and gap scores in the 1st secondary year have been taken into account?
METHOD

Database
In order to answer the two research questions, data were used that were originally gathered for the VOCL cohort in 1989. This is a national longitudinal study of students’ secondary school careers. The first measurements for the study were carried out in the 1989/90 school year among just under 20,000 first-year students from 381 schools – representing every type of secondary education apart from special schooling. The sample broadly represents all types of students in that particular school year (Driessen, Dekkers, & Bosker, 1997).

In the 1st year of the study, information on students’ gender, socioeconomic status, and ethnic background was gathered via a questionnaire given to schools and parents. Every year, students’ position within the education system (school year and school type) was established, their language and mathematics test scores in the 1st and 3rd years recorded, and information gathered on their choice of subjects for their final school-leaving examinations.

The Structure of Dutch Secondary Education
In order to comprehend the description of the dataset and the relevant variables, and for a good interpretation of the results, a short description of the structure of Dutch secondary education might be helpful.

Nearly all children start their education at the age of 4, even though the jure compulsory education starts at the age of 5. After about 8 years of primary education they will enter secondary education, which is divided into five tracks: pre-vocational education (VBO: 4 years: age 12–16), junior general secondary education (MAVO: 4 years: age 12–16), senior general secondary education (HA VO: 5 years: age 12–17), pre-university education (VWO: 6 years: age 12–18) and senior secondary vocational education (MBO: 4 years: age 16–20). Most secondary schooling takes place within schools offering a variety of the tracks mentioned above. In the higher grades the classes become more and more singularly tracked.

Each type of secondary education has its own subsequent track in tertiary education: a VWO-certificate allows for admission to university education, whereas a HAVO-certificate is sufficient for admission to higher professional

1The information was gathered and made available by the Central Bureau for Statistics (CBS), Heerlen, The Netherlands.
education. The subsequent track for VBO and MAVO is the aforementioned MBO (senior secondary vocational education).

In this study, only general secondary schools (MAVO, HAVO) and pre-university education (VWO) are taken into account. In these school types, students choose their final subjects 1 year before the final school-leaving examinations. The number of subjects chosen varies according to the secondary school type: Students in general secondary schools choose at least six subjects, whereas pre-university students take at least seven subjects.

**Variables**
The following variables are important for the analyses in this study:

**Gender**
This variable requires no further explanation: (1) boys, (2) girls.

**Socioeconomic Status**
A proven good measure for SES is the educational level of parents. The social background of students was therefore indexed by the highest educational level of the parents using the categories: (1) no primary education; (2) primary education; (3) secondary education – lower levels; (4) secondary education – higher levels; (5) higher education – degree level; (6) higher education – postgraduate level.

**Ethnic Origin**
This variable is based on the parents’ country of birth. A student was considered of ethnic origin when one or both parents were foreign. The resulting variable comprised the categories: (1) Turkey/Morocco; (2) Aruba, Dutch Antilles, Surinam and the Moluccas (former Dutch colonies); (3) other nations, mainly in southern Europe, associated with the Dutch Education Priority Policy program, that is, Spain, Greece, Italy; (4) other foreigners from industrialized countries like Belgium, Germany, the UK, and the US; (5) The Netherlands.

**Family Composition**
(1) Both parents present; (2) one-parent family.
**Achievement Scores**

In the 1st and 3rd years, tests were administered. In the 1st year three tests were taken; a language test, an arithmetic test, and an information processing test. All tests consisted of 20 multiple-choice items (Cronbach’s alpha’s >0.80). In the 3rd year, two subjects, Dutch language and mathematics, were tested again. Testing here was limited to students in the 3rd year of the cohort schools who had not had to repeat a year. When these tests were compiled, the differences in the levels between school types were taken into consideration as far as possible (among other things via calibration). The language test consisted of 40 multiple-choice items and the mathematics test of 32 items, some multiple-choice and some open (Cronbach’s alpha’s >0.80). The test results are based on the number of correctly answered items. For the purpose of the analyses the scores have been standardized on a mean of 10 with a standard deviation of 1.

**Gap Scores**

For each student the gap between the (standardized) mathematics and language score was calculated in the 1st and 3rd years:

\[
\text{Gap score} = \frac{\text{Mathematics}}{\text{language score}}
\]

**Final Examination Subjects**

Based on information on the choice of final examination subjects for secondary school students, a cumulative variable was constructed for the number of science subjects – mathematics, physics, and chemistry – chosen (Driessen et al., 1997). The fact that the number of subjects chosen varies according to the secondary school type also had to be taken into account (general secondary school students take at least six subjects, whereas pre-university secondary school students take seven). Thus, in the data analysis a distinction had to be made between the different secondary school types (junior and senior general secondary education and pre-university secondary education).

**Analyses**

To answer the first research question, the relation between the gap between mathematics and language in the 3rd year and the individual background variables (gender, SES, ethnic origin, and family composition) was described.
using frequency divisions and multilevel regression analyses for continuous
data. The multilevel regression analyses with a continuous dependent variable are carried out with MLwiN (Goldstein et al., 1998). In a first analysis, the background variables and interactions between the background variables were used to explain the gap between mathematics and language in the 3rd year. In the second analysis, also the gap between mathematics and language in the 1st year was added to the explanatory variables and interactions of the first analysis. In the final analysis, only those explanatory variables that were significant in the second analysis were used to explain the gap between mathematics and language in the 3rd year.

The second research question was answered using multilevel regression analysis for ordinal data. With these analyses, the relation between background variables, gap scores in the 1st and 3rd years, and choice of science subjects were traced. The decision to apply a multilevel analysis for ordinal data is made due to the nature of the dependent variable, the number of science subjects chosen in the final examination. This dependent variable cannot be modeled as a continuous variable due to two factors. First of all, the possible number of science subjects a student can choose is limited. In junior general secondary schools the maximum number of science subjects a student can choose is three, while in pre-university education the maximum is four. Secondly, not all science subjects are chosen independently; for example, hardly any student chooses physics without mathematics. Modeling the number of science subjects as ordinal, assumes an increasing scale that is not necessarily of interval nature. To carry out the multilevel analyses with the ordinal dependent variable, the program MIXOR (Hedeker & Gibbons, 1994) is applied. After examining the relation between background variables, gap scores in the 1st and 3rd years and choice of science subjects, it was examined whether the gap between mathematics and language achievement could explain the choice of science subjects at secondary school level.

In the analyses used to answer the research questions only students who had taken the aptitude tests in the 1st and 3rd years were included. Also tests in the 3rd year were only given to those students who had not repeated a year. This, however, can give a biased view (see Driessen & Van der Werf, 1994). In order

---

2The same analyses were carried out for the gap scores in the 1st year. Because the links with the background variables are similar for both years, only the results for the 3rd year are reviewed in this article. The results on the 1st-year gap scores can be obtained from the first author of this article.
to control for selective dropout, the group of students who took part in both tests were compared with the total research sample. Results showed that those students who did not take tests in language and mathematics in the 3rd year (repeaters) scored significantly lower in the 1st year than students who had not repeated a year. The differences were greater for mathematics than for Dutch language (mathematics: $t = -7.23; p = .000$; language: $t = -4.60; p = .000$). The group of nonrepeaters scored on average slightly higher in mathematics than in language (mean = 0.02). For the group of students who dropped out before the 3rd year, the reverse was true (mean = −0.03). The differences between the various subgroups were the same for both nonrepeaters and repeaters. When drawing conclusions, the fact that this report concerns the “better” students needs to be taken into account. The students included in the analyses have a greater edge in mathematics compared to language skills than was the case for the group as a whole.

RESULTS

Descriptive Statistics
A short overview of the data is presented in Table 1. In Table 1 percentages and frequencies are given for the categories of each of the explanatory variables that will be used in the multilevel analyses. In Columns 2 and 3, percentages and frequencies are given for junior general secondary education. Columns 4 and 5 present similar data for senior general secondary education, and Columns 6 and 7 for pre-university education. The last two columns give the percentages and frequencies of the totals over junior and senior general and pre-university secondary education. Table 1 shows that the separate data for senior general and pre-university education contained a limited number of students from single-parent families and from each of the ethnic minority groups.

The Gap Between Mathematics and Language Scores in the 3rd Year of Secondary School
To answer the first research question, multilevel regression analyses were carried out for the gap scores in the 3rd year, in which the main and interactive effects of background variables and 1st-year gap scores on the gap between mathematics and language in the 3rd year were assessed stepwise. This means that in the first model only the effects of the background variables and
interactions of these variables were taken into account. Then it was assessed to what extent these effects remain significant after 1st-year gap scores are added (model 2) and every nonsignificant main and interactive effect is left out of the model (model 3). These analyses were carried out for all data at once and also separately for each of the three different types of secondary education studied in this research.

In Table 2, the results of the multilevel analyses for all data at once are summarized by $\chi^2$ test-statistics. For each independent variable and each of the interactions in the analyses the $\chi^2$ statistic is given. Table 3 presents the parameter estimates of the third model, which contains only the significant

<table>
<thead>
<tr>
<th>Family</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Single parent</td>
<td>5.6%</td>
<td>150 f</td>
<td>5.8%</td>
<td>48 f</td>
</tr>
<tr>
<td>Two parents</td>
<td>94.4%</td>
<td>2508 f</td>
<td>94.2%</td>
<td>781 f</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Girl</td>
<td>57.1%</td>
<td>1517 f</td>
<td>55.9%</td>
<td>463 f</td>
</tr>
<tr>
<td>Boy</td>
<td>42.9%</td>
<td>1141 f</td>
<td>44.1%</td>
<td>366 f</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ethnic origin</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Turkish, Moroccan</td>
<td>2.6%</td>
<td>68 f</td>
<td>1.4%</td>
<td>12 f</td>
</tr>
<tr>
<td>Surinam, Antilles, Aruba, Mollu</td>
<td>1.8%</td>
<td>49 f</td>
<td>1.0%</td>
<td>8 f</td>
</tr>
<tr>
<td>EPP south</td>
<td>0.7%</td>
<td>18 f</td>
<td>0.2%</td>
<td>2 f</td>
</tr>
<tr>
<td>Western</td>
<td>4.1%</td>
<td>109 f</td>
<td>5.4%</td>
<td>45 f</td>
</tr>
<tr>
<td>Dutch</td>
<td>90.8%</td>
<td>2414 f</td>
<td>91.9%</td>
<td>762 f</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SES</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary education (not finished)</td>
<td>1.1%</td>
<td>30 f</td>
<td>0.4%</td>
<td>3 f</td>
</tr>
<tr>
<td>Primary education</td>
<td>10.5%</td>
<td>280 f</td>
<td>5.4%</td>
<td>45 f</td>
</tr>
<tr>
<td>Secondary education (lower)</td>
<td>29.2%</td>
<td>775 f</td>
<td>18.7%</td>
<td>155 f</td>
</tr>
<tr>
<td>Secondary education (higher)</td>
<td>41.5%</td>
<td>1103 f</td>
<td>42.0%</td>
<td>348 f</td>
</tr>
<tr>
<td>Higher education (lower)</td>
<td>14.9%</td>
<td>397 f</td>
<td>25.7%</td>
<td>213 f</td>
</tr>
<tr>
<td>Higher education (higher)</td>
<td>2.7%</td>
<td>73 f</td>
<td>7.8%</td>
<td>65 f</td>
</tr>
</tbody>
</table>

Total 2658 829 895 4378
Table 2. The Influence of Background Variables on the Gap Between Mathematics and Language Scores in the 3rd Year Controlled for Gap Scores in the 1st Year.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Model 1</th>
<th></th>
<th></th>
<th>Model 2</th>
<th></th>
<th></th>
<th>Model 3</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\chi^2$</td>
<td></td>
<td>p</td>
<td>$\chi^2$</td>
<td></td>
<td>p</td>
<td>$\chi^2$</td>
<td></td>
<td>p</td>
</tr>
<tr>
<td>Family (single parent)</td>
<td>1</td>
<td>0.81</td>
<td>0.368</td>
<td></td>
<td>0.73</td>
<td>0.392</td>
<td></td>
<td>0.61</td>
<td>0.434</td>
<td></td>
</tr>
<tr>
<td>Gender (girl)</td>
<td>1</td>
<td>75.87</td>
<td>0.000</td>
<td></td>
<td>30.76</td>
<td>0.000</td>
<td></td>
<td>74.81</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Ethnic origin</td>
<td>4</td>
<td>4.86</td>
<td>0.302</td>
<td></td>
<td>5.53</td>
<td>0.237</td>
<td></td>
<td>4.23</td>
<td>0.376</td>
<td></td>
</tr>
<tr>
<td>SES</td>
<td>5</td>
<td>12.58</td>
<td>0.028</td>
<td></td>
<td>10.21</td>
<td>0.069</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family × gender</td>
<td>1</td>
<td>0.26</td>
<td>0.610</td>
<td></td>
<td>0.13</td>
<td>0.723</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family × ethnic origin</td>
<td>4</td>
<td>8.06</td>
<td>0.090</td>
<td></td>
<td>7.27</td>
<td>0.122</td>
<td></td>
<td>11.65</td>
<td>0.020</td>
<td></td>
</tr>
<tr>
<td>Gender × ethnic origin</td>
<td>4</td>
<td>5.54</td>
<td>0.236</td>
<td></td>
<td>4.95</td>
<td>0.293</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender × SES</td>
<td>5</td>
<td>0.97</td>
<td>0.965</td>
<td></td>
<td>1.84</td>
<td>0.871</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnic origin × SES</td>
<td>20</td>
<td>20.89</td>
<td>0.404</td>
<td></td>
<td>19.72</td>
<td>0.476</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gap year 1</td>
<td></td>
<td></td>
<td></td>
<td>317.49</td>
<td>0.000</td>
<td>320.78</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deviance</td>
<td></td>
<td></td>
<td>20397.30</td>
<td>20085.70</td>
<td>20125.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Model 1: all explanatory variables and interactions no covariate.
Model 2: all explanatory variables and interactions and 1st-year distance as covariate.
Model 3: significant explanatory variables, interactions ($p < .05$) and 1st-year distance as covariate.

Table 3. Model 3 With Significant Explanatory Variables and Interactions ($p < .05$).

<table>
<thead>
<tr>
<th>Source</th>
<th>$\beta$</th>
<th>SE($\beta$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.08</td>
<td>0.02</td>
</tr>
<tr>
<td>Gap 1</td>
<td>0.19</td>
<td>0.01</td>
</tr>
<tr>
<td>Gender</td>
<td>Girl</td>
<td>-0.17</td>
</tr>
<tr>
<td>Ethnic origin</td>
<td>Turkish, Morrocan</td>
<td>-0.09</td>
</tr>
<tr>
<td></td>
<td>Surinam, Antilles, Aruba</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>EPP south</td>
<td>-0.15</td>
</tr>
<tr>
<td></td>
<td>Western</td>
<td>0.01</td>
</tr>
<tr>
<td>Family composition</td>
<td>Single Parent (SP)</td>
<td>0.03</td>
</tr>
<tr>
<td>Family Composition × Ethnic origin</td>
<td>SP × Turkish, Morrocan</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>SP × Surinam, Antilles, Aruba</td>
<td>-0.31</td>
</tr>
<tr>
<td></td>
<td>SP × EPP south</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>SP × Western</td>
<td>-0.18</td>
</tr>
<tr>
<td>Random Effects</td>
<td>Level 2</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Level 1</td>
<td>0.62</td>
</tr>
</tbody>
</table>
main effects, significant interactions, and the main effects of the variables involved in the significant interactions.

From Tables 2 and 3 it appears that there is evidence of interaction between the variables ethnic origin and family composition in the final model \( (p < .05) \). The group of students from the former Dutch colonies from one-parent families scored higher in language than mathematics (Gap 3: mean = \(-.20\)), while the students from one-parent families from countries associated with the Dutch Education Priority Policy program had an advantage in mathematics (Gap 3: mean = 0.77). Thus, for these students mathematics achievement did not fall behind that for language. For former Dutch colony students, the lack of a parent – in many cases the father – has a greater influence on their mathematics skills than for the other groups. Students in this group on average score slightly less for language than Dutch students (9.51 and 9.94 respectively), but the difference for mathematics is significantly higher (9.22 versus 9.98).

Next to this interaction effect, a main effect exists for the variable gender, whereby boys score higher in mathematics than in language. Consequently, for girls the reverse is true; they are relatively ahead in language compared to mathematics.

When the three models were compared by performing a test on the model deviances \( (−2 \times \log \text{likelihood}) \), it was found that the second model performed significantly better than the first model \( (χ^2 = 311.60, \ df = 1, p < 0.01) \). Although the third model uses fewer parameters than the second model, the last model does not perform significantly better \( (χ^2 = 40.20, \ df = 39, p = 0.42) \). So, taking into account the gap between mathematics and language scores in the 1st year of secondary school reduces the deviance significantly.

The analyses on the data of the separate school types resulted in similar results as obtained in the analyses with all the data.\(^3\)

**Choice of Subjects and Specialization in Secondary Education**

As mentioned before, different types of secondary education allow for a different number of science subjects that can be chosen. Four analyses were performed on the grouped data of junior general, senior general, and pre-university secondary education. In these analyses, the relationship was assessed of the gap between maths and language scores in the 1st and 3rd years

\(^3\)The results of these analyses can be obtained from the first author.
and the choice for science subjects. In a first preliminary analysis that is not included in Table 4, an empty model was estimated. This model used as independent variables a random intercept and two dummy variables that indicate senior general and pre-university secondary education. The dummy for senior general secondary education was not significant ($Z = 0.49, p = 0.67$). Consequently, it can be concluded that there is little difference between the distribution of number of science subjects chosen in junior and senior general secondary education. The deviance, a measure of the overall fit of the model, was 12192.37. Expanding the empty model with the gap scores in year 1 and year 3 resulted in a model with a deviance of 11837.43. Results of this model are presented in Table 4 (Model 1). The two models could be compared, since the difference between the deviance of two models is asymptotically $\chi^2$ distributed with degrees of freedom equal to the difference in the number of parameters between the models. With this test it was found that Model 1 performed better than the empty model ($\chi^2 = 354.94, df = 2, p < 0.01$). So, it can be concluded that including the gap scores resulted in a significant improvement of the fit of the model.

Next, family composition, sex, ethnic background, SES, and interactions of these variables were included in the model as explanatory variables. Model 2 (Table 4) was obtained by deleting the nonsignificant interactions. Model 2 showed a significant better fit than Model 1 ($\chi^2 = 538.15, df = 8, p < 0.01$). Both the gap scores in year 1 and year 3 were significant (respectively $Z = 7.61, p < 0.01$, and $Z = 8.37, p < 0.01$) and deletion of the gap scores (Model 3) led to a decrease in fit of the model ($\chi^2 = 160.65, df = 2, p < 0.01$).

From Table 4 it appears that student’s gender is an important factor in accounting for differences in the number of science subjects chosen at the final examination level in secondary education; as was to be expected, these subjects are more popular among boys than among girls. At the same time, group effects can be established for both socioeconomic status and ethnic origin; students from higher academic backgrounds tend to choose more science subjects than students with lower educated parents, whereas students from ethnic minority groups are less likely to choose science courses than their Dutch peers.

Beside these group effects, the gap scores in the 1st and 3rd years of secondary education also correlate positively with the choice of science subjects in the final examination year; students with a comparative advantage in mathematics tend to choose more science subjects than students with an advantage in language.
Table 4. The Influence of the Gap Between Mathematics and Language Scores in the 3rd Year of Secondary Education on the Number of Science Subjects Chosen in General Secondary School Types Controlled for Gap Scores in the 1st Year and Background Variables.

| Fixed          | Model 1  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |�

| Model 2  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |�

| Model 3  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |�

Note. All significant main and interactive effects (p < .05) are printed in bold.
The same analyses have also been performed after adding the absolute mathematics achievement score to the model (not included in the table). As was to be expected, the effect of the gap score between mathematics and language on the choice of science subjects was significantly reduced after the inclusion of the absolute mathematics scores. The interaction between the gap score and the mathematics score remained significant, however \( Z = 2.08, p = 0.039 \). So, even after the absolute achievement scores were added to the model, the gap score between mathematics and language resulted in a significant improvement of the fit of the model.

CONCLUSION AND DISCUSSION

In general secondary education, students’ gender is the key factor in accounting for differences in the number of science subjects chosen at the final examination level. Boys generally choose more science subjects than girls. Besides gender, student’s socioeconomic (educational) and ethnic background is also related to the number of science subjects chosen in the final examination year; students from higher academic backgrounds and Dutch students tend to choose more science subjects than their peers.

At the same time, however, previous achievements, more specifically the relative achievements in mathematics compared to language, are also positively correlated to the choice of science subjects. Students with a comparative advantage in mathematics (in the 1st and 3rd years of secondary education) choose more science subjects for their final examination than those where the gap between maths and language is less marked. Including the gap scores resulted in a significant improvement of the research model. Even when the absolute achievement scores for mathematics were added to the model, the gap score between mathematics and language remained significant. The gap score between mathematics and language scores in the 3rd year of secondary education is in its turn predicted by gender and for students from former Dutch colonies by family composition.

What do these findings imply for current research and the ongoing debate on policy measures with regard to the shortage of science and technical graduates in The Netherlands and elsewhere? The most important conclusion from this study is that the variable “gap between mathematics and language

---

4The results of these analyses can be obtained by the first author.
scores” has additional value when accounting for students opting for science subjects.

Besides the earlier mentioned group effects of gender, SES, ethnic origin, and family composition, the choice of examination subjects appears to be influenced by a process of individual rational choice in which previous achievements, and especially relative achievements in mathematics compared to language, play an important role. Also differences in the number of science subjects chosen in (sub)groups of students can at least partly be explained by the comparative advantage in mathematics within these groups. For example, gender differences in subject choice are influenced by the fact that boys have a relative advantage in male-typical subjects, which affects their expectations of the probability of success in these subjects.

This argues in favor of future research and policy aimed at strengthening the choice of science subjects, focusing attention not only on absolute achievement in the most obvious subjects, but also on the relation between (attainment in) various subjects. The degree of enrollment in science and technical subjects is relatively high among students who score higher in mathematics than in language. To increase the enrollment of sufficiently proficient students that currently do not choose science or technical subjects, future campaigns should focus on those who score above average in mathematics but who at the same time score relatively high in language.

REFERENCES


