DOCUMENT RESUME

ED 376 969 PS 022 812

AUTHOR Luyten, Hans; Van der Hoeven-van Doornum, Anneke
TITLE Classroom Composition and Individual Achievement: Effects of Classroom Composition and Teacher Goals in Dutch Elementary Education.
PUB DATE 94
NOTE 31p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Ability Grouping; Academic Achievement; *Classroom Environment; Elementary Education; Foreign Countries; High Achievement; Homogeneous Grouping; *Intelligence Differences; Intelligence Quotient; Low Achievement; Socioeconomic Background; Student Characteristics; *Student Educational Objectives; Teacher Attitudes; *Teacher Expectations of Students; Teacher Student Relationship
IDENTIFIERS Netherlands; *Rational Decision Making

ABSTRACT

This study investigated the effects of classroom composition on achievement. The analysis focused on the assessment of classroom homogeneity with respect to the socioeconomic background and general intelligence of the students. In addition, the extent to which classroom composition affects the standards the teachers employ and how these standards can account for the effects of classroom composition were examined. Two different data sets were analyzed: the first, which contains information about a national sample of 212 elementary schools, was used to assess the effects of classroom composition; the second data set, which contained information about a sample of 51 schools from four regions in the East and South of the Netherlands between May 1986 and June 1988, was used primarily to test whether teacher goals can account for the effects of classroom composition. The results suggested that: (1) an increase in classroom homogeneity may lead to a somewhat higher level of average achievement in Dutch elementary education, but also to a considerably larger variation; (2) the achievements of the least talented students are likely to decrease; (3) although teacher goals were found to exert a considerable effect on individual achievement, they could not account for the observed classroom composition effects; and (4) the effect of the classroom average on individual achievement was similar in size to the effect of socioeconomic background. Contains 43 references. (AA)

# CLASSROOM COMPOSITION AND INDIVIDUAL ACHIEVEMENT
## Effects of classroom composition and teacher goals in Dutch elementary education[1]

**Hans Luyten,** *University of Twente, Department of Education*
**Anneke van der Hoeven-van Doornum,** *ITS Nijmegen (Institute for applied social sciences)*

*The research reported in this article deals with the consequences of classroom composition on individual achievement. The main research question is which effects are to be expected when the students in the Dutch system of elementary education become grouped into more homogeneous classes. It seems likely that classroom homogeneity will increase in the near future, as parents of Dutch origin tend to send their children to schools where the students of foreign origin constitute no more than a (small) minority of the entire school population. The investigations specifically focused on two possible effects of more homogeneous classrooms: changes with respect to the variation in student achievement and changes pertaining to the average level of student achievement. The effects of individual cognitive aptitudes, of the classroom average, of the classroom heterogeneity and their interaction effects were examined. It was also investigated to what extent the cognitive goals teachers set for their students can account for classroom composition effects on achievement. The analyses showed that an increase in classroom homogeneity may lead to a somewhat higher level of average achievement in Dutch elementary education, but also to a considerably larger variation. The achievements of the least talented students are likely to decrease. Although teacher goals were found to exert a considerable effect on individual achievement, they could not account for the observed classroom composition effects. The effect of the classroom average on individual achievement was found to be of similar size as the effect of socio-economic background.*

## 1. INTRODUCTION

Most of the empirical research dealing with the effects of classroom heterogeneity in elementary education relates to the American educational system, where tracking students into separate classes according to their perceived cognitive aptitudes is a very common practice (Slavin, 1987). In the Netherlands students are hardly ever deliberately grouped into homogeneous classes. Most elementary schools are even too small to allow for such grouping practices. On average the elementary schools comprise only one class per grade. Because their size is rather small, the number of schools is quite large and it is not unusual to find several elementary schools within very close range, often in one and the



[1]The authors would like to thank Hennie Brandsma, Anja Knuver and Gerry Reezigt for providing part of the data (the "national sample", see section 5) that were analyzed in the present study.

same street and sometimes even in the same building[2]. Together with churches, pubs and bus connections elementary schools are the most widely available facilities in the Netherlands. For 90% of the population between the ages of four and twelve at least two elementary schools are available in their own residence (Blank et al., 1990; pp. 39-66).

Even though students are not deliberately grouped into homogeneous classes, many of them might end up in homogeneous classes, *because* the schools are so small and, as a result, mainly enrol students from a very restricted area. Especially the populations of urban schools are generally quite homogeneous with respect to the socio-economic and ethnic background of their students. Since the cognitive aptitudes that enhance success at school coincide with such background characteristics to a considerable extent, the student populations of urban schools may be quite homogeneous with respect to their students' aptitudes as well. Another important reason for this homogeneity, is the fact that parents of Dutch origin tend to send their children to schools where the students of foreign origin constitute no more than a (small) minority of the entire school population. As a result many students of foreign origin[3] end up in schools where Dutch students are largely absent (Rath, 1991; pp. 177-238).

The fact that in the Netherlands students in elementary schools are not deliberately grouped into homogeneous classes entails some important consequences with respect to the interpretation of the research outcomes to be reported. When students are deliberately tracked into homogeneous low-ability and high-ability classes, this is done to provide the most appropriate level and pace of instruction. In such cases it is no surprise that the teachers in high-ability classes set higher goals for their students than the teachers in the low-ability classes and it is therefore difficult to establish whether the less demanding instruction in low-ability classes is an appropriate response to the students' abilities or that it unnecessarily holds them back (Gamoran, 1992). In Dutch elementary education, however, the same curriculum is supposed to be taught across all classes. If Dutch teachers are found to set consistently lower goals in the low-ability classrooms this can be more straightforwardly interpreted as a response to the average ability level of the students.

---

[2]Rath (1991, pp. 177-238) describes a case of two elementary schools in Rotterdam, housed in the same building, but one with a population of predominantly foreign children (90%) and the other with a student population of predominantly Dutch children (60%). The parents of foreign origin frequently protested against this segregation.

[3]Most of these students originate from Surinam, Turkey, Morocco, Aruba or the Dutch Antilles (mainly Curaçao).

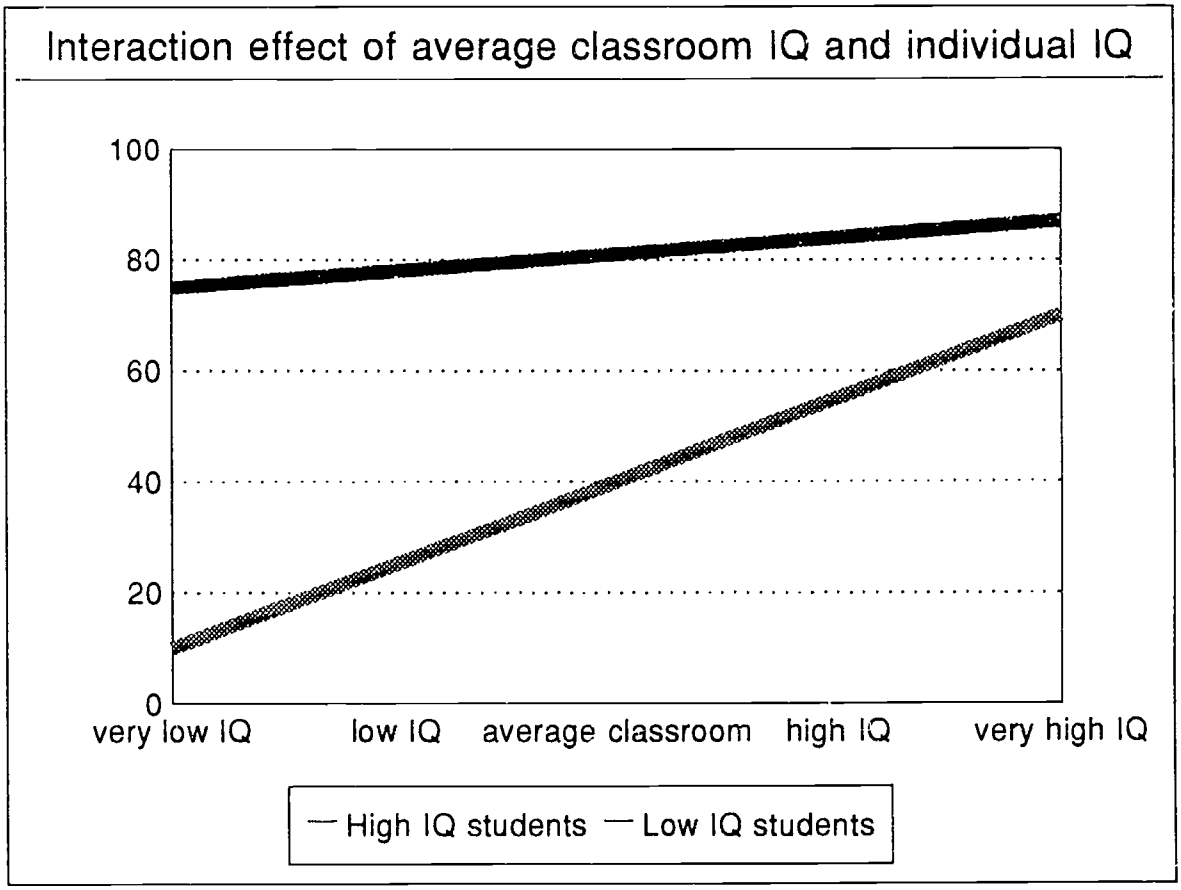
## 2. PRESUMED AND OBSERVED EFFECTS OF CLASSROOM COMPOSITION

Grouping students into heterogeneous classes has frequently been advocated as a means to obtain an equitable distribution of student achievement. This is based on the often reported research finding that student achievement is positively (but moderately) affected by the classroom's average level of ability. Such effects have been reported for both elementary and secondary education in several educational systems, such as the American, Israeli, Dutch, English and Scottish (Beckerman & Good, 1981; Leiter, 1983; Rowan & Miracle, 1983; Dar & Resh, 1986; Kerckhoff, 1986; Willms, 1986; Mensen & Guldemond, 1987; Link & Mulligan, 1991; Resh & Dar, 1992; Reezigt, 1993). Grouping students into homogeneous classes is believed to reinforce already existing inequalities between students with respect to their cognitive aptitudes, because the less talented will end up in low-ability classes, which will affect their achievements negatively, while the more talented experience a positive effect of the high ability level in their classes. Slavin (1987), however, argues that assigning students to separate classes which are homogeneous with respect to their *general* cognitive abilities is less likely to have an effect on achievement than other forms of ability grouping. According to Slavin grouping plans can be expected to produce stronger effects when they reduce the heterogeneity with respect to the specific skill being taught, but classes will remain quite heterogeneous with respect to most subjects, when the same segregation criterion is applied for each subject. Teachers are then still faced with the problem of providing the appropriate level and pace of instruction for all the students in the classroom (Gamoran, 1992). The outcomes of the meta-analyses by Slavin (1987; 1990) dealing with the effects of ability grouping in elementary and secondary education support this view.

Another argument that has been advanced by proponents of heterogeneous grouping is that it might lead to a higher level of student achievement, because the effect of the average classroom ability is believed to be stronger for low-ability than for high-ability students (Oakes, 1985; Hallinan, 1987; De Vos, 1986; 1989). According to this line of reasoning the achievements of the most talented may be slightly impeded, but the achievements of the less talented are expected to rise more substantially. If all students would be grouped into heterogeneous classes instead of homogeneous ones, the talented students would end up in classes with an average ability level that is lower than their own individual level of ability and the reverse would be true for the low-ability students. If the effect of the average classroom ability is the same for the talented and the less talented students, the overall effect of heterogeneous grouping would be zero. However, if the effect is stronger for the less talented students, then the achievement gain of the low-ability students would outweigh the loss of the high-ability students. Apart from the research outcomes presented

by Dar & Resh (1986) and Resh & Dar (1992) which relate to secondary education in Israel there is not much empirical support for this view, which actually presumes an interaction effect of the individual and the classroom ability level on achievement. In figure 1 such an interaction effect is graphically displayed. Two lines are drawn: the upper line represents the effect of classroom composition for the talented students, while the lower line represents the effect for the less talented. In this artificial example the effect of classroom composition on achievement is much stronger for the less talented students.

FIGURE 1: ARTIFICIAL EXAMPLE




It should be noted, that, even if such an interaction effect is actually at work, it does not always make sense to subtract the disadvantage of the talented students from the profit of the low-ability ones. Under certain circumstances more value must be attached to the achievements of specific groups and less to the achievements of the others. Consider the following two situations:

1/ A large percentage of the students leaving elementary education appears to be virtually illiterate;

2/ An increasing number of the students from pre-university education turns out not to be able to succeed in college.

If the number of illiterate students at the end of elementary schooling is unacceptably high, this would call for policy measures aiming to raise the reading achievements of the low-ability students even if this would lead to a considerable decline in the achievements of the more talented ones. In this case an approach resulting in a substantial reduction of the number of illiterate students would be needed, even if this would entail a decreased average achievement level. The second situation is more or less the reverse of the former. This time the achievements of the high-ability students are unacceptably low. Measures producing a rise in the achievements of the most talented students, but at the same time causing a downfall in average achievement may in this case be appropriate.

Another argument against tracking students into homogeneous classes that has frequently been raised is its contribution to the segregation of students from different social and ethnic backgrounds (e.g. Slavin, 1987; Gamoran, 1992).

Opponents of grouping students into heterogeneous classes maintain that teachers in homogeneous classes can more easily provide instruction at the most appropriate pace and level (Kulik & Kulik, 1982; 1984). According to this line of reasoning both high and low-ability students will suffer from heterogeneous grouping. The level and pace of instruction in heterogeneous classes will hardly provide any challenge for the high-ability students, while the low-ability ones are thought to be discouraged when they are confronted with highly talented classmates. Apart from two recent studies dealing with elementary education in the Netherlands (Maas, 1992; Reezigt, 1993) there is hardly any empirical evidence available that confirms the assertion that classroom heterogeneity affects student achievement negatively. Although the meta-analyses by Kulik & Kulik (1982; 1984) provide some support for the idea that ability grouping enhances student achievement in both elementary and secondary education, the outcomes do not allow for the conclusion that grouping students into homogeneous classes affects achievement positively. The Kuliks do not distinguish between several forms of ability grouping, such as assigning students to within-class groups for particular subjects, tracking them into separate classes or offering special programs for the gifted. In the meta-analyses by Slavin (1987; 1990) the effects of several types of ability grouping were separately investigated. Slavin concludes that the overall effect of tracking students into homogeneous classes on achievement is zero, both in elementary and secondary education. This conclusion is mainly based on American research. Positive effects of classroom heterogeneity on

achievement have been reported for students in Israeli secondary education (Dar & Resh, 1986; Resh & Dar, 1992), while in Dutch secondary education classroom heterogeneity does not seem to affect the school careers of the students (De Vries, 1992).

Although the available evidence is not conclusive with respect to the direction of the heterogeneity effect, the findings hardly show any contradiction as far as the size of the effect is concerned. It is not clear whether the effect of classroom heterogeneity on achievement is positive or negative, but is does not seem very strong.

Some other possible effects of classroom composition that are not frequently mentioned in the debate on the advantages and disadvantages of classroom heterogeneity require some attention as well. First of all, the interaction of individual ability and classroom heterogeneity. Grouping students into heterogeneous classes is sometimes believed to result in a higher level of general achievement, because low-ability students are expected to be more sensitive to the average level of ability in their classroom than their more talented classmates. However, if low-ability students are more sensitive to the classroom ability level, they may also be more sensitive to other classroom characteristics, such as the possibly negative effect of classroom heterogeneity. This possibility should be taken into account in an analysis that deals with the effects of grouping students into heterogeneous classes, for it may neutralize the positive effect of the average classroom ability on the achievements of the low-ability students. Secondly, the interaction of the average classroom ability and heterogeneity. The effect of the average ability level may be stronger in homogeneous classes, because in those classes teachers can provide instruction at the same level and pace for the entire classroom. The equalizing effect that is believed to result from heterogenous grouping may then be rather weak. If all students are grouped into maximally heterogeneous classes and the average ability level is identical across all classes, the impact of the average classroom ability on individual achievement may be minimal. Thirdly, there may be a three-way interaction effect of individual ability, classroom ability and classroom heterogeneity, implying that the effect of classroom ability and heterogeneity may be different for low-ability and high-ability students.

## 3. CLASSROOM COMPOSITION EFFECTS AND RATIONAL CHOICE THEORY

The explanations that have been proposed for the effects of classroom composition on individual achievement can be grouped into two main categories. The first one refers to the idea of differential instruction (e.g. Beckerman & Good, 1981; Rowan & Miracle, 1983; Gamoran, 1986; 1992; Slavin, 1987; 1990) implying that high-ability classes receive a more favourable instruction, because teacher behaviour varies depending on the

classroom composition or because the more competent teachers are disproportionately assigned to high-ability classrooms. In the other category of explanations the role of interaction and competition among classmates is emphasized (e.g. Erbring & Young, 1979; Mensen & Guldemond, 1987; Duke, 1993). Students in high-ability classes are believed to be stimulated by the achievements of their peers, whereas students in low-ability classes lack such incentives.

In this paper we will focus on the differential instruction hypothesis in order to explain the effects of classroom composition on individual achievement. De Vos (1986; 1989) has pointed out that the idea of differential instruction can be incorporated into rational choice theory quite easily. Differential instruction can be conceived as resulting from purposeful behaviour of both teachers and students. De Vos has summarized his argument as follows:

> " students' efforts are an important determinant of their achievements, ... the teacher's evaluative behavior affects the amount of efforts students are willing to expend on school work, and ... students' achievements affect the teacher's evaluative behavior." (De Vos, 1989, p. 223)

Both teachers and students are assumed to maximize their "subjective expected utility", which results in striving for social approval and physical well-being. Students can obtain social approval from the teacher through achievement. Achieving, however, requires effort. The efforts, which lead to a decrease in physical well-being, represent the costs of achievement. It goes without saying that the costs of achieving are relatively high for low-ability students and low for high-ability students. According to De Vos students will keep raising their achievements up to the point that the marginal utility of achieving equals its marginal cost. Going beyond this point would be irrational, because raising one's achievements even higher would entail extra costs that are no longer offset by an extra amount of social approval. The amount of utility a student obtains as a result from achievement is largely dependent on the way the teacher responds to the efforts of the students.

Teachers are believed to obtain social approval through the achievements of their students. Like the students, they are faced with a problem of optimization, since raising the student achievements requires effort from the teachers as well. The intellectual capacities of the students also determine the costs for the teachers to a large extent, for it will be easier to obtain high achievements in high-ability classes. It is assumed that teachers behave basically in the same way as the students. Both teachers and students are believed to raise their efforts up to the point that the marginal utility equals the marginal cost. As a result teachers will require higher achievements, i.e. apply a more demanding standard, in classes with a high average ability level and lower achievements in low-ability classes. Students

in high-ability classes, also the less talented, may therefore be expected to reach a higher level of individual achievement, because they are confronted with a more demanding standard.

It should be noted that De Vos strongly emphasizes the effect of standards on the classroom as a whole, although he acknowledges that teachers do not apply a uniform standard to all their students. This would in most cases even be irrational, because the standard would then be too high for some students in the classroom, and too low for others. Students who are faced with standards that are set too high for them, will be strongly discouraged to expend any effort at all, while talented students who are confronted with a very low standard, will hardly be encouraged to raise their achievements, even though they are able to do so. The crucial point, however, is that students of equal ability are believed to be confronted with different standards depending on the average classroom ability level. Teachers in high-ability classes can be expected to set higher individual goals for each student and the students in those classes will be faced with a higher general classroom standard.

De Vos also contends that classroom composition effects are stronger for low-ability students than for the more talented ones and that this can also be conceived as resulting from purposeful behaviour of the teacher, if one is willing to accept the additional assumption that teachers in low-ability classes tend to pay relatively much attention to their best students in order to "prove" that the low achievements in their class are not caused by poor instruction. This implies that teachers in low-ability classes can be expected to set relatively high standards for their most talented students.

## 4. RESEARCH QUESTIONS

Our main research question refers to the effects of classroom composition on achievement. We start our analyses assessing the classroom homogeneity with respect to the socio-economic background and general intelligence of the students. In addition we examine to what extent classroom composition affects the standards the teachers employ and to what extent these standards can account for the effects of classroom composition. The analyses specifically address the following questions:

1/ To what extent are the classrooms in Dutch elementary education homogeneous with respect to the socio-economic and intellectual backgrounds of the students? This question will be dealt with in section 6.1. A substantial amount of classroom

homogeneity is expected, especially with respect to the students' socio-economic backgrounds, because most schools enrol students from a rather restricted area.

2/ In how far do teachers apply a uniform standard for the entire classroom? We will address this question in section 6.2. In the explanation of classroom composition effects proposed by De Vos (1986; 1989) the effect of standards on the classroom as a whole is strongly emphasized. On the other hand, it is not plausible that teachers will apply a perfectly uniform standard for all their students.

3/ Does classroom composition affect the goals teachers set for each individual student? Do teachers set higher goals in high-ability classes independently of the students' individual abilities? According to De Vos teachers set higher goals for their students in high-ability classes. Furthermore, teachers in low-ability classes are believed to set particularly high goals for their most talented students and relatively low goals for the less talented. Section 6.2 presents an empirical assessment of the hypothesized relations between classroom composition and teacher goals.

4/ To what extent do the average classroom ability level and classroom heterogeneity affect the individual achievements of the students independently of their individual ability and their socio-economic family background? We will deal with this question in section 6.3. A positive effect of the average classroom ability level is expected, which would imply that grouping students into homogeneous classrooms reinforces the already existing inequalities with respect to their intellectual capacities. A negative effect of classroom heterogeneity, however, would imply that more equality with respect to achievement entails a somewhat lower level of general student achievement.

5/ Are the effects of classroom composition different for high and low-ability students? Is the effect of the average classroom ability stronger in homogeneous or in heterogeneous classes? In other words: are there any significant interaction effects of individual ability, classroom ability and heterogeneity? These issues are addressed in section 6.3. Investigating the strength of these interaction effects is required for an adequate understanding of the effects of classroom composition. One of the arguments advanced by the proponents of heterogenous grouping is the assertion that heterogeneous grouping will raise the general level of student achievement, because the positive effect of the average classroom IQ is believed to be particularly strong for the less talented students. On the other hand, one should acknowledge that this effect may be counterbalanced by a negative heterogeneity effect which mainly affects the achievements of the low-ability students. It should

also be checked whether the effect of the classroom ability level remains equally strong when classes become more heterogeneous.

6/ To what extent can the goals set by the teachers account for the effects of classroom composition? How strong are the classroom composition effects after controlling for teacher standards? How strong is the empirical basis for the conception of classroom composition effects as resulting from purposeful teacher behaviour? These questions are discussed in section 6.3.

7/ Which changes are to be expected if the students were grouped into perfectly heterogeneous or perfectly homogeneous classes? How will it affect the variation in student achievement? What will be the consequences for high and low-ability students. These questions will be dealt with in section 6.4.

The next section presents a description of the datasets that were used in the analyses. The outcomes are reported in section 6.

## 5. DATA AND METHOD OF ANALYSIS

Two different datasets were analyzed to investigate the effects of classroom composition. The first dataset, which contains information about a national sample of 212 elementary schools, was used to assess the effects of classroom composition, whereas the second dataset was primarily used to test whether teacher goals can account for the effects of classroom composition. The first dataset has been used by several researchers in various studies (e.g. Brandsma, 1993; Knuver, 1993; Reezigt, 1993). The present study focuses on the students who were in their final year of elementary education. The data relate to 3993 students from 221 classes and were collected in 1987 and 1988. To assess the impact of classroom composition on individual achievement information about the following three variables was needed:

- achievement
- intelligence
- socio-economic status (SES)

Classroom characteristics, such as means and standard deviations, could be computed on the basis of this information. To measure achievement a test made up of 88 items relating

to language and mathematics was used[4]. These items were derived from standardized tests ("CITO-tests"), which are administered in most schools a few months before the end of the final year in elementary education. The CITO-tests can be considered to cover the basic curriculum in Dutch elementary education quite adequately (Blok, 1992). Intelligence was measured one year earlier, in 1987, by means of an IQ-test (the "ISI-test") measuring both spatial and verbal intelligence. Information about the socio-economic background of the students, which served as a control variable in our analyses, was obtained from the teachers. The SES-index is based on four indicators: the education of a student's father and mother and their profession.

The second dataset contains information about a sample of 51 schools from four regions in the East and South of the Netherlands (Twente, Arnhem, Tilburg and Zuid-Limburg). The data were originally collected at several points in time between may 1986 and june 1988 by Van der Hoeven-Van Doornum (1990) for a study into the indirect effect of socio-economic status via teacher expectations and aspiration levels on school careers in elementary education. Students were followed during their last two years of elementary education. In the present study we confine ourselves to the effects of classroom composition in the final year of elementary education. The dataset was obtained through disproportionate sampling (Van der Hoeven-Van Doornum, 1990; pp. 48-58). Two stratification criteria were applied: the composition of the schools' student populations with respect to socio-economic and ethnic background and the academic output of the schools in 1985[5]. Schools were grouped into three categories with respect to the first criterion (high, medium and low-input schools). On the basis of the second criterion the schools in each category were classified as high, low or medium-output schools. From each category an equal number of high and low-output schools was included in the designed sample. The medium-output schools were not included in the sample[6]. The schools in the sample did not differ significantly from the total population of schools in the four regions with respect to student enrolment, denomination (public, catholic, protestant or other), degree of urbanization or the socio-economic and ethnic background of their students (Van der Hoeven-Van Doornum, 1990; pp. 59-60). The students in the

[4]Although the test consists of 58 language items and only 30 mathematics items, the mathematics items are not underrepresented. The composite achievement score was obtained by computing the unweighted average of the (standardized) scores on the language and mathematics sub-tests.

[5]The output measure was based on information about the school careers in secondary education of each school's students.

[6]It should be noted that the classification of high, low and medium output was different for the three categories (high, low or medium input). Schools with an "average" output were included in the sample, but these were predominantly high-input schools with a relatively low output or low-input schools with a high output.

sampled schools scored slightly below the national average on the test score that served as the main criterion variable in our analyses - 533.3 versus 535.7 - , while the standard deviation in the sample was somewhat higher than in the national population - 11.25 versus 9.75 (Van der Hoeven-Van Doornum, 1990; p. 94). The analyzed dataset contained information about 698 students from 57 classes. Our analyses required information about only four student level variables. Classroom characteristics were computed on the basis of these four variables, which are:

- achievement
- intelligence
- socio-economic status (SES)
- cognitive goals set by the teacher

Achievement was measured by means of a standardized test (the "final CITO-test") consisting of 180 multiple choice items dealing with language, mathematics and information processing (i.e. reading tables, maps and graphs). This test is administered in the majority of Dutch elementary schools shortly before the end of the final year, usually in february. Intelligence was assessed by means of the same IQ-test (the "ISI-test") as the one in the national sample. This test was administered in september 1987. Socio-economic background was measured by an index constructed on the basis of the education and profession of the students' parents. The information was obtained directly from the parents. The cognitive goals were measured by a nine item index, which serves as an operationalization of their standards. In december, three months after the beginning of the school year, the teachers were interviewed about the cognitive goals they had set for each of their students. The questions were formulated as follows[7]:

Please indicate your approach for this student.

- Set modest standards*
- Emphasize practical work rather than language and mathematics development*
- Teach only topics the student will need later on*
- Let the student work with material from lower grades*
- Teach extensive parsing skills
- Give the student the opportunity to switch to easier subject matter*
- Present enrichment material
- Let the student work ahead
- Demand correct spelling

[7]For the original Dutch formulations see Van der Hoeven-Van Doornum (1990, pp. 107-108).

These nine questions were Likert-type items with five response categories. The internal consistency of the scale (Cronbach's $\alpha$) equalled .86. The items marked with an asterisk (*) were negatively scored.

It should be noted that in the present study classroom homogeneity is conceived as a sociological phenomenon rather than an educational one. In Dutch elementary education students are hardly ever deliberately grouped into homogeneous classes. As far as students do end up in homogeneous classes this is the result of sociological factors. Schools generally enrol students from restricted areas, which are often quite homogeneous with respect to the socio-economic and ethnic backgrounds of the residents and Dutch parents tend to send their children to schools where the students of foreign origin constitute no more than a (small) minority of the entire school population, so that many students of foreign origin end up in "black" schools. We expect that such socio-economic and ethnic homogeneity produces a substantial homogeneity with respect to the cognitive aptitudes of the students as well. The present study aims to establish if how far the classroom composition characteristics with respect to the general cognitive aptitudes of the students affect their individual achievements. It was decided to investigate the effect of the (individual and classroom average) IQ-scores rather than the effect of prior achievement, because IQ-scores are in general more closely related to the student's socio-economic and ethnic background characteristics than actual achievement. IQ-scores are more stable in time and less curriculum-dependent than test-scores which relate to more specific knowledge and skills. The analyses aim to assess the effect of classroom composition on the general achievement scores of the students and do not differentiate between language and mathematics achievement.

Several multilevel analyses were conducted to answer the research questions formulated in section 4. In order to obtain an impression of the effects to be expected if the students were grouped into more heterogeneous or homogeneous classes, some simple simulations were conducted based on the empirical findings with respect to the classroom composition effects.

## 6. RESULTS

The multilevel analyses were conducted using the ML3-software (Prosser et al.; 1991). The findings are presented in the next three sections. The outcomes with respect to the simulations are presented in section 6.4.

### 6.1. Classroom homogeneity with respect to IQ and SES

The homogeneity of the classrooms in the investigated sample with respect to the socio-economic background and the individual IQ-scores of the students was assessed by means of a multilevel analysis. For both variables the total amount of variance was partitioned into student level and classroom level variance. The amount of classroom level variance expresses how much of the total variance can be attributed to differences between the classroom means. The student level variance expresses how much of the total variance is attributable to differences between students within classes. The results are listed in table 6.1, which shows that, although most of the variance is situated at the student level, there are still substantial differences between classes with respect to the socio-economic backgrounds of their students. The differences between classrooms with respect to the IQ-scores were found to be considerably smaller, especially in the national sample. No more than 11.3% of the total variance in IQ-scores appeared to be situated at the classroom level, whereas about a quarter of all variance with respect to the SES-scores turned out to be attributable to differences between classes. This implies that the classes in Dutch elementary education cannot not be made much more heterogeneous than they presently are with respect to the cognitive aptitudes of the students.

TABLE 6.1: *Student and classroom level variance with respect to IQ and SES*

| | National Sample | | Regional Sample | |
|---|---|---|---|---|
| | **SES** | **IQ** | **SES** | **IQ** |
| **Student level variance (standard errors)** | 74.9% (1.7) | 88.7% (1.6) | 73.2% (4.1) | 81.9% (4.6) |
| **Classroom level variance (standard errors)** | 25.1% (2.9) | 11.3% (2.0) | 26.8% (6.5) | 18.1% (4.9) |

The appendix provides additional information about the frequency distributions of these and other variables that were analyzed in the present study. The correlation between socio-economic status and IQ was found not to be very high at the individual level in either sample. The relation between both variables appeared to be stronger at the classroom level. The student level correlation in the national sample is .33 (pearson r), while it equals .49 at the classroom level. In the regional sample the correlation is .36 at the individual and .52 at the classroom level.

In several studies which deal with the effects of classroom composition the researchers have been faced with the problem that the average classroom ability level and the

classroom heterogeneity were very strongly correlated (Guldemond et al., 1987, pp. 61-75). In the present study the standard deviation of the IQ-scores in the classroom served as a heterogeneity measure. The correlations between the average classroom IQ and the standard deviation did not lead to any problems of multicollinearity in either sample. The correlation between both variables equalled -.13 in the national sample (which is not significant for $\alpha < .05$) and -.43 in the regional sample.

### 6.2. Effects of classroom composition on teacher goals

This section deals with the question to what extent teachers set uniform goals for all the students in their classes and to what extent classroom composition variables can account for the goals teachers set for the individual students. The second dataset, the one relating to the regional sample, was used to address these issues. The total variance in teacher goals was partitioned into student level and classroom level variance. Next the effects of individual student characteristics, classroom characteristics and their interaction effects on the teacher goals were examined. The results of these analyses are presented in table 6.2. Before conducting the analyses all variables were transformed into z-scores, so that the reported effects can be interpreted as standardized regression coefficients. The figures in between brackets denote the standard errors. The interaction terms were computed by multiplying the z-score transformations. The effects of these interaction terms can be interpreted as standardized regression coefficients as well (Jaccard et al., 1990).

Model 0 shows the partitioning into student and classroom level variance. The percentages of variance at the student and classroom level are similar to the ones found with respect to socio-economic background and IQ-scores. Most of the variance in teacher goals is situated within the classrooms, which implies that teachers set quite different goals for the students in their classes. Although the amount of classroom level variance is rather modest as compared to the student level variance, the outcomes do reveal substantial differences between classes with respect to teacher goals. Model 1 presents the effects of several independent variables and their interactions on the teacher goals. The regression coefficients that are significant for $\alpha < .05$ in a two-tailed t-test are printed in bold face. Most coefficients, however, are very small and not statistically significant. Individual IQ reveals the strongest effect on the goals set by the teachers. Teachers also set somewhat higher goals for students from a more favourable socio-economic background. Apart from the interaction effect of classroom heterogeneity (measured by the classroom standard deviation for IQ-scores) and individual IQ, no significant effects were found for the classroom composition characteristics. The interaction effect implies that in the more heterogeneous classes teachers set particularly high goals for the most talented students, and particularly low goals for the less talented.

TABLE 6.2: *Effects of individual and classroom characteristics on teacher goals*

*Regional Sample*

| | Model 0 | Model 1 |
|---|---|---|
| **Standardized regression coefficients (standard errors)** | | |
| **1/ Individual characteristics** | | |
| IQ-score | -- | **.511** (.035) |
| Socio-economic status | -- | **.092** (.033) |
| **2/ Classroom characteristics** | | |
| Mean IQ | -- | -.040 (.068) |
| Standard deviation IQ | -- | -.004 (.062) |
| **3/ Interaction effects** | | |
| Individual IQ * mean IQ | -- | -.020 (.033) |
| Individual IQ * std. dev. IQ | -- | **.076** (.031) |
| Mean IQ * std. dev. IQ | -- | -.053 (.045) |
| Ind. IQ * mean IQ * std. dev. IQ | -- | -.003 (.025) |
| | | |
| **Grand Mean (standard errors)** | -.073 (.070) | -.053 (.066) |

| **VARIANCE** | | **VARIANCE EXPLAINED** |
|---|---|---|
| Student level | 79.1% | 37.7% |
| Classroom level | 20.9% | 21.9% |
| Total | 100.0% | 34.3% |

The expectation that teachers set higher goals in high-ability classes was not confirmed. Teachers do set higher goals for the more talented students, but this is not related to the average level of ability in the classroom. Also, the expectation that teachers in low-ability classes set relatively high goals for their most talented students, which would imply a negative interaction effect of individual and classroom ability, cannot be considered to be corroborated. However, the analyses did reveal substantial differences between classrooms with respect to the goals set by the teachers. The average classroom goals may have an effect on student achievement. In the next section it will be discussed whether the average goals per classroom can account for classroom composition effects.

### 6.3. Effects of classroom composition and teacher goals on individual achievement

In this section it will be discussed to what extent classroom composition affects student achievement and whether teacher goals can account for such effects. Both the national and the regional sample are examined. The information provided by the national sample allows for a reliable estimation of the impact of classroom composition in Dutch elementary education. An analysis of the regional sample will show in how far the effects of classroom composition can be attributed to the goals set by the teachers.

The zero models in table 6.3 show that the amount of classroom level variance with respect to student achievement is substantial in both samples. The differences between classrooms in this respect are considerably larger than the classroom differences regarding IQ-scores and teacher goals. They are also somewhat larger than the differences with respect to socio-economic background. The classroom level variance probably provides a close approximation of the school level variance, because only a few schools in both samples comprise more than one class in grade 8. The classroom level variance thus largely coincides with the school level variance. In most studies dealing with school effectiveness in Dutch elementary education the amount of school level variance, however, has been reported to be considerably smaller (Blok, 1992). It is quite likely that the differences between schools are somewhat overestimated in the regional sample, because schools with remarkably high and low outputs in former years are overrepresented to some extent (see section 5). It seems less likely, though, that the relatively large amount of school level variance that was found in the national sample is due to the sampling design. It should be noted that many reports which deal with the differences between Dutch elementary schools are not based on representative samples (Blok, 1992). In some cases the sample only relates to a certain region (Meijnen, 1984; Bosker & Hofman, 1987), in some cases schools with many disadvantaged students are overrepresented (Van der Werf & Weide, 1991; Van der Werf et al., 1991; Weide, 1993; Van der Velden, 1993; Jungbluth, 1993) and in other cases there is no information about the representativeness of the samples (Brandsma & Knuver, 1988; Van de Grift & Akkermans, 1991). Apart from the national sample that was analyzed in the present study the sample analyzed by Blok (1992) may be the only one that can be considered to produce a representative picture of the differences between schools. This sample contains information about student achievement in 179 schools for five consecutive years (1987-1991). Blok reports a declining trend in time with respect to the amount of school level variance. For 1988, the year when the data for the national sample in the present study were collected, he reports a school level variance of 17% for language and 22% for mathematics. These outcomes still deviate to some extent from the findings in the present study, but they also suggest

that the relatively small amounts of school level variance reported in previous studies underestimate the true differences between schools.

For both samples model 1 presents the effects of individual characteristics and classroom composition on student achievement. Again, the reported effects can be interpreted as standardized regression coefficients and the ones which are significant for $\alpha < .05$ in a two-tailed t-test are printed in bold face. The strongest effect was, not surprisingly, presented by individual intelligence. Socio-economic background was found to affect achievement as well. A moderate, but statistically significant effect of the average classroom ability level could be detected in both samples. The national sample also revealed a significant effect of classroom heterogeneity and a significant interaction effect of individual and classroom IQ. Classroom heterogeneity was found to exert a modestly negative effect on individual achievement. The negative sign of the interaction effect implies that the effect of the average classroom IQ is stronger for low-ability students than for the more talented ones. This effect, however, is even smaller than the effect of classroom heterogeneity. Similar interaction effects and heterogeneity effects were found in the regional sample as well, but in that case they were not statistically significant. This is probably due to the limited size of this sample which renders the statistical power of the analyses rather low if one wants to detect effects of a modest size[8]. In general the findings from both samples are fairly consistent. It is difficult, however, to establish which factors may account for the observed inconsistencies. The fact that the samples were obtained through different sampling designs may be the most important explanation for the divergencies. The national sample must be considered to provide a more representative picture, because it relates to the entire system of Dutch elementary education and because in the regional sample schools with high and low output are overrepresented. It should also be noted that two variables, socio-economic status and student achievement, were not measured in exactly the same way in both samples. The achievement scores in the national sample relate only to language and mathematics, whereas the achievement scores in the regional sample are based on items pertaining to language, mathematics and information processing. In the regional sample the information with respect to the students' socio-economic background was obtained directly from the parents, but in the national sample the teachers were asked to provide this information. It can of course not be precluded that

[8]This can be illustrated by considering the figures pertaining to the effect of "Standard deviation IQ" in the regional sample. Given the size of the standard error for this regression coefficient (.046) only an effect larger than .90 would be considered statistically significant, because only coefficients that are about twice as large as their standard error are significant for $\alpha < .05$. If the actual size of the effect is smaller than .90, the chances to obtain a statistically significant result are less than 50 %. In other words: the power of the test is limited when it comes down to detecting small effects.

some of the inconsistencies reflect certain genuine regional peculiarities in the East and South of the Netherlands, but this does not seem very likely.

TABLE 6.3: *Classroom composition, teacher goals and individual achievement*

| | National Sample | | Regional Sample | | |
|---|---|---|---|---|---|
| | **Model 0** | **Model 1** | **Model 0** | **Model 1** | **Model 2** |
| **Standardized regression coefficients (standard errors)** | | | | | |
| **1/ Individual characteristics** | | | | | |
| IQ-score | -- | **.571** (.012) | -- | **.658** (.026) | **.468** (.026) |
| Socio-economic status | -- | **.167** (.012) | -- | **.128** (.024) | **.095** (.021) |
| Teacher goals | -- | -- | -- | -- | **.371** (.025) |
| **2/ Classroom characteristics** | | | | | |
| Mean IQ | -- | **.113** (.024) | -- | **.128** (.050) | **.188** (.055) |
| Standard deviation IQ | -- | **-.063** (.025) | -- | -.048 (.046) | -.043 (.043) |
| Classroom mean teacher goals | -- | -- | -- | -- | -.097 (.055) |
| **3/ Interaction effects** | | | | | |
| Individual IQ * mean IQ | -- | **-.034** (.010) | -- | -.034 (.025) | -.026 (.022) |
| Individual IQ * std. dev. IQ | -- | -.015 (.011) | -- | .024 (.023) | -.004 (.020) |
| Mean IQ * std. dev. IQ | -- | -.003 (.018) | -- | .058 (.034) | **.076** (.032) |
| Ind. IQ * mean IQ * std. dev. IQ | -- | .010 (.008) | -- | .030 (.019) | .031 (.025) |
| | | | | | |
| **Grand Mean (standard error)** | -.057 (.039) | -.000 (.009) | -.094 (.087) | -.005 (.014) | -.012 (.048) |

| **VARIANCE** | | **VAR. EXPL.** | **VARIANCE** | **VARIANCE EXPLAINED** | |
|---|---|---|---|---|---|
| Student level | 71.3% | 49.5% | 64.9% | 61.0% | 70.8% |
| Classroom level | 28.7% | 60.1% | 35.1% | 75.9% | 76.1% |
| Total | 100.0% | 52.5% | 100.0% | 66.2% | 72.7% |

The fact that the other two-way interaction effects were found to be very small and non-significant in the national sample is also worth mentioning. Our concern that the low-ability students might be more strongly influenced by a negative heterogeneity effect than the more talented ones (see section 2) turned out not to be justified. The same is true for our worries that the impact of the classroom average might decrease in heterogeneous

classrooms. The regional sample even displayed (non-significant) effects in the opposite direction. The three-way interaction of individual IQ with classroom IQ and classroom heterogeneity also appeared to be insignificant in both samples.

When controlling for the cognitive goals set by the teacher in the regional sample, some remarkable changes emerge. Model 2 shows a considerable, positive effect of the goal set by the teacher at the individual level, but the classroom average of the teacher goals was found to affect achievement negatively. This effect, however, is not significant for $\alpha < .05$. This finding is consistent with the outcomes presented by Van der Velden (1993) who investigated the effect of a similar variable on language and mathematics achievement. In his analyses the teacher goals related to the classroom as a whole and not to individual students. This variable was not found to affect achievement. Another variable, however, which relates to the minimum goals the teacher sets for his classroom did produce a significant effect. We will return to this question in the discussion section.

The outcomes suggest that at the individual level the teacher goals function as an intermediate variable between individual IQ and socio-economic background on the one hand and achievement on the other. The effects of both variables on achievement in the regional sample are considerably lower in model 2 than in model 1. The effects of individual IQ and socio-economic background in model 1 can be considered as the total effects on achievement, whereas model 2 presents the direct effects, i.e. the effects that are independent of the teacher goals. In the previous section the teacher goals were shown to be dependent on IQ and SES. The difference between the regression coefficients for the interaction effect of classroom heterogeneity with individual IQ presents another indication for the supposition that the teacher goals intermediate between individual ability and achievement. In section 6.2 a significant interaction effect on teacher goals of individual ability with heterogeneity was reported. Model 1 in table 6.3 presents an interaction effect in the same direction on achievement for the regional sample, although it is not statistically significant. When controlling for teacher goals, however, the interaction effect is virtually equal to zero. The effects of the average classroom ability level and its interaction with classroom heterogeneity appear to increase when controlling for teacher goals. In model 2 this interaction effect is even significant for $\alpha <. 05$, which would imply that the effect of the average classroom IQ is *stronger* in heterogeneous classroom instead of weaker. Our worries that the effect of the average classroom ability level might be rather weak in heterogenous classes were therefore not confirmed. The results presented in model 2 clearly indicate that the goals set by the teachers, as operationalized in the present study, cannot account for the classroom composition effects that were found.

### 6.4. Simulated effects of classroom composition

In order to obtain an impression of the effects to be expected if the students were grouped into more heterogeneous or more homogeneous classes some simple simulations were conducted based on the empirical findings with respect to classroom composition effects. The general model that was used to "predict" the effects is shown in table 6.4.1. The simulations are based on model 2. Model 1 is included in the table to show how much of the variance in student achievement can be explained by the classroom composition variables. The extra amount of variance explained by model 2 at the classroom level is substantial (9.3%), but its contribution at the individual level is very small (0.1%). The increase across both levels is therefore quite modest as well (2.7%).

TABLE 6.4.1: *Models for simulating student achievement (based on the national sample)*

| | Model 1 | Model 2 | Means and standard deviations | |
|---|---|---|---|---|
| **Regression coefficients** | | | **Mean** | **Std. Dev.** |
| Individual IQ-score | 1.679 | 1.647 | 11.494 | 1.803 |
| Socio-economic status | 0.133 | 0.129 | 0.000 | 6.695 |
| Classroom mean IQ | -- | 0.821 | 11.494 | 0.714 |
| Standard deviation IQ | -- | -1.020 | 1.673 | 0.323 |
| Interaction Ind. IQ * mean IQ | -- | -0.136 | 0.509 | 1.601 |
| | | | Achievement score | |
| **Grand Mean** | 1.119 | -6.114 | 20.562 | 5.205 |
| **VARIANCE EXPLAINED** | | | | |
| Student level | 49.3% | 49.4% | | |
| Classroom level | 50.7% | 60.0% | | |
| Total | 49.7% | 52.4% | | |

Table 6.4.2 shows the outcomes of the simulations. Three situations were considered:

A/ The present situation

B/ A situation in which students were grouped into completely heterogeneous classes

C/ A situation in which students were grouped into completely homogeneous classes

ad A: First of all, the achievement scores that would be expected in the present situation on the basis of our model were computed. In this way a "yardstick" was obtained to evaluate the outcomes in the other two situations. The expected scores were computed by means of the following equation:

$$CITO_{exp} = -6.114 + 1.647*IQ_{ind.} + .129*SES + .821*IQ_{class} - 1.020*SD_{IQ} - .136*IQ^{c}_{ind.}*IQ^{c}_{class} \quad (1)$$

Where:

| | | |
|---|---|---|
| $CITO_{exp}$ | = | The expected achievement score on the CITO-test |
| $IQ_{ind.}$ | = | The individual IQ-score |
| SES | = | The socio-economic background |
| $IQ_{class}$ | | The classroom mean IQ-score |
| $SD_{IQ}$ | = | The standard deviation of the individual IQ-scores in the classroom |
| $IQ^{c}_{ind.}$ | = | The individual IQ-score centred around the overall mean IQ-score (11.494) |
| $IQ^{c}_{class}$ | = | The classroom mean IQ-score centred around the overall mean IQ-score (11.494) |

The product of the last two variables makes up the interaction term that relates to the interaction effect of individual and classroom IQ. An appropriate procedure to assess a two-way interaction effect in multiple regression models is to centre both variables around their respective means before computing a multiplicative term. In this way the problem that the correlations between the interaction term and the variables $IQ_{ind.}$ and $IQ_{class}$ would be too large (the problem of multicollinearity) can easily be circumvented (Cronbach, 1987, Jaccard et al., 1990)[9].

Table 6.4.2 shows the average and the standard deviation of the expected scores. Note that the standard deviation of these scores is considerably lower than the actual standard deviation. This is because our model accounts for only 52.4% of the total variance in the achievement scores. The table also shows the expected scores for students of different ability. For this purpose the students were grouped into four categories of equal size.

ad B: In this case each class would be a perfectly representative sample of the entire population of students. The average IQ and the standard deviation of IQ-scores in each class would equal the average score and standard deviation in the whole population. The value for $IQ_{class}$ is thus assumed to be 11.494 in all classes. Since the regression effect for $IQ_{class}$ is .821, the effect of this variable should equal 9.437 (.821*11.494) in all classes.

[9] If standardized coefficients are preferred the product of the z-score transformations may serve as the interaction term (see section 6.2).

The value for $SD_{IQ}$ is assumed to equal 1.803 in all classes and the regression coefficient for this variable is -1.020. The expected effect of classroom heterogeneity is thus estimated to be -1.839 (-1.020*1.803) in each and every class. The interaction term is equal to zero in this situation, because the value for $IQ^c_{class}$ is zero in every class. Summing the values for the intercept and the effects of $IQ_{class}$ and $SD_{IQ}$ yields 1.484 (-6.114 + 9.437 - 1.839). The equation to compute the expected scores thus becomes:

$$CITO_{exp} = 1.484 + 1.647*IQ_{ind.} + .129*SES \qquad (2)$$

ad C: In this hypothetical situation the students are assumed to be grouped into classrooms with zero heterogeneity. This implies that the negative effect of $SD_{IQ}$ is totally absent and that for each student the average classroom IQ equals his or her individual IQ. The variable $IQ_{ind.}$ now actually occurs two times in the equation, because $IQ_{ind.}$ and $IQ_{class}$ are assumed to be identical. This can of course be simplified by computing a single term ($2.468*IQ_{ind.}$ instead of $1.647*IQ_{ind.} + .821*IQ_{ind.}$). Since in this situation $IQ_{ind.} = IQ_{class}$, the interaction term changes into a quadratic term. The equation now becomes:

$$CITO_{exp} = -6.114 + 2.486*IQ_{ind.} + .129*SES - .136*(IQ^c_{ind.})^2 \qquad (3)$$

Table 6.4.2 shows that extremely homogeneous grouping would result in much larger deviations from the present situation than extremely heterogeneous grouping. This is not so surprising, because in the present situation the students are already grouped into fairly heterogeneous classrooms (see table 6.1). Homogeneous grouping may be expected to result in a somewhat higher level of average achievement, but also in considerably more variance. The achievements of the least talented would decrease, but the achievements of the others would increase, especially those of the brightest 50%. Grouping students into extremely heterogeneous classes would result in a slightly lower amount of variance and in somewhat higher achievement of the least talented students.

The figures in table 6.4.2 should be considered primarily as an illustration of the trends that were detected in the national sample rather than definite predictions of student achievement under certain conditions. It should be noted that the model on which the simulations are based, leaves nearly 50% of the total variance in student achievement unexplained. Moreover, the simulations are based on the assumption that the trends which were found in the present situation can be extrapolated to conditions of extreme homogeneity or heterogeneity of classrooms. Especially the situation of perfectly homogeneous classes deviates strongly form the prevailing circumstances.

TABLE 6.4.2: *Simulated achievement scores in different grouping conditions*

| | Present Situation (expected scores) | Heterogeneous Classrooms | Homogeneous Classrooms |
|---|---|---|---|
| High IQ students | 25.0 | 24.7 | 27.4 |
| Moderately high IQ | 21.8 | 21.6 | 23.7 |
| Moderately low IQ | 19.4 | 19.3 | 20.3 |
| Low IQ students | 15.5 | 16.0 | 14.8 |
| General Average | 20.5 | 20.4 | 21.6 |
| Standard Deviation | 3.72 | 3.35 | 4.84 |

## 7. DISCUSSION

The main purpose of the present study was to assess the effects of classroom composition on individual achievement in Dutch elementary education. It was specifically investigated what will be the consequences if students become grouped into more homogeneous classes. The research outcomes indicate that classroom homogeneity tends to reinforce the inequalities between students with respect to their cognitive aptitudes. Increasing classroom homogeneity may result in higher achievements for most students, especially the most talented, but it must be expected to impede the achievements of the least talented. Our concerns that the effect of the average classroom IQ on individual achievement might be very weak in heterogeneous classes or that the negative effect of classroom heterogeneity is particularly strong for the low-ability students were not confirmed. Although the classrooms in both samples appeared fairly homogeneous with respect to the socio-economic backgrounds of their students, the classes turned out to be surprisingly heterogeneous with respect to the cognitive aptitudes of the students.

Classroom homogeneity, however, must be expected to increase under the present circumstances. Dutch parents generally seem to prefer schools with no more than a small minority of foreign students for their children, because they suspect that a large number of foreign classmates affects the achievements of the Dutch students negatively. As a result many students of foreign origin end up in classes which are almost entirely made up of other foreign students. Without government intervention this development can be expected to result in a considerably increased homogeneity of the classes, which will coincide with a segregation along ethnic lines. The present government policy, which is aimed at increasing the size of the elementary schools, may provide some opportunities to

counterbalance this trend. The larger the schools, the wider the area from which they enrol their students. The student populations of the schools may thus become more heterogenous as their size increases, although this does not guarantee that the classrooms will become more heterogeneous. Presently most schools are too small to allow for the grouping of students into homogeneous classes. When school size increases, some schools may choose to track students into high and low-ability classes. It seems dubious that the creation of larger schools can on its own turn the tide of increasing classroom homogeneity. More heterogeneous classes and a more equal distribution of student achievement can only be obtained through more specific measures. It should also be noted that, although the classrooms in the national sample were found to be quite heterogeneous with respect to the IQ-scores of their students, this does not preclude that in some regions, especially the large cities, classes may still be very homogeneous in this respect.

Our analyses aimed to test whether the effects of classroom composition can be conceived as resulting from rational behaviour by both teachers and students as proposed by De Vos (1986; 1989). It was investigated if teachers set higher goals for their students in high-ability classes than in low-ability classes. Although our analyses demonstrated that teacher goals exert a considerable effect on student achievement and that the teacher goals are strongly dependent on a student's individual IQ, the evidence clearly failed to confirm the assertion that classroom composition effects result from rational behaviour through the goals set by the teachers. The research outcomes did not reveal a significant effect of the average classroom IQ on the goals set for the individual students (see table 6.2), nor was the classroom average of the teacher goals found to affect student achievement in the way it was expected (see table 6.3). The classroom composition effects appeared even to increase after controlling for the teacher goals, although this effect cannot be considered to be statistically significant. On the other hand, it should be noted that the fundamental assumptions that teachers set higher goals for the intelligent students and that the students' achievements increase when they are faced with higher goals were clearly confirmed *at the individual level*. The present study, however, showed that teachers set quite diverging goals for the students in their class, whereas in the argument by De Vos the emphasis is on the effects of standards for the entire classroom.

On the other hand, the fact that in the present study the classroom composition effects could not be accounted for by the teacher goals may to some extent be due to the way they were measured. The items on which the index measuring the teacher goals is based (see section 5) may have encouraged the teachers to emphasize the variation in approaches they employed for the students in their class. Perhaps, if more specific and objective information about their goals (e.g. through detailed questions about the subject matter they wanted their students to master) had been available, the within classroom differences with

respect to the teacher goals might have appeared smaller. However, the fact that substantial differences between classrooms with respect to the teacher goals were found (see table 6.2) suggests that the index is able to distinguish classes where high goals are set from classes with low goals, apart from detecting differences within the classrooms.

It should also be noted that the index employed in the present study expressed the teachers' maximum aspiration levels rather than their minimum standards. When focusing on the minimum levels of achievement the teachers demand from their students classroom effects on individual achievement may still be (partly) accounted for by the teacher goals. The findings reported by Van der Velden (1993) support this idea. In his analyses the effects of two kinds of teacher goals on student achievement were investigated. Both variables related to the goals for the classroom as a whole and not to individual students. The first variable resembles the operationalization of teacher goals that was employed in the present study quite close. This variable was not found to affect achievement, which is consistent with the findings in this study, as we only detected an effect at the individual level. The other variable, however, which relates to the minimum goals the teacher sets for the classroom did produce a significant effect. These outcomes suggest that at the classroom level the minimum standards set by the teachers are important, whereas our finding indicate that at the individual level the maximum aspiration levels of the teachers enhance achievement. To gain a more clear understanding of the causes and effects of teacher goals further (and probably quite elaborate) research is required.

More accurate measures of teacher goals may be obtained by asking teachers to indicate whether they want a certain student to master some well-defined subject matter. Such questions might even be asked about specific test-items. Separate questions might be formulated to measure the minimum standard and the maximum aspiration level a teacher sets for each student. Explicit questions could also be asked about the goals and standards at the classroom level. The present study was mainly focused on the maximum aspiration levels the teacher had set for each student with respect to not very specifically defined subject matters and it was assumed that the classroom mean of the individual goals would correctly reflect the goals the teachers had set for their class as a whole. Another issue that could not be addressed in the present study refers to the relation between the teacher goals and their actual behaviour in the classroom. We assume, however, that the information provided by the teachers presents a valid picture of their intentions. Furthermore, it should be pointed out that even when teacher goals and their behaviour are not clearly related, the goals may still have an important effect on achievement. Teacher goals reflect what they want their students to achieve and it seems plausible that teachers are able to attain a chosen goal via very different approaches.

Our study has nevertheless demonstrated that a teacher may choose quite different goals for the students in his/her class. Even if the way the teacher goals were measured tends to exaggerate the differences within classes somewhat, which is for the moment no more than an unconfirmed suspicion, this conclusion seems unavoidable. This probably limits the extent to which teacher goals can account for classroom composition effects. The research outcomes suggest that the interactions and competition among classmates may offer more valid explanations of classroom composition effects than the idea of differential instruction, as the effects of classroom composition turned out even to increase after controlling for teacher goals. Apparently, the classroom composition effects did not occur because of the teacher goals, but rather *despite* the teacher goals. In the present study the reinforcement of the inequalities between students which results from homogeneous grouping could certainly not be attributed to the goals set by the teachers. On the contrary, the teachers rather seemed to impede the effects of classroom composition.

## REFERENCES


Beckerman, T.M. & Good, T.L. (1981). "The classroom ratio of high and low aptitude and its effect on achievement". *American Educational Research Journal*, **18**, 317-327.

Blank, J.L.T., Boef-van der Meulen, S., Bronneman-Helmers, H.M., Herweijer, L.J., Kuhry, B. & Schreurs, R.A.H. (1990). *School en schaal* (School and scale). Rijswijk/Den Haag: Sociaal en Cultureel Planbureau/VUGA.

Blok, H. (1992). "De grootte van het schooleffect in het basisonderwijs; een analyse op basis van vijf jaar Eindtoets Basisonderwijs". (The size of the school effect in elementary education; an analysis based on five years Final Elementary Education Test). *Tijdschrift voor Onderwijsresearch*, **17**, 343-354.

Bosker, R. & Hofman, A. (1987). "Dimensies van schoolkwaliteit. De algemene en milieuspecifieke invloed van scholen op prestaties en het keuzegedrag van leerlingen" (Dimension of school quality. The general and background specific effect of schools. In: Scheerens, J. & Stoel, W.G.R. (eds.) *Effectiviteit van onderwijsorganisaties* (Effectiveness of educational organizations). Lisse: Swets & Zeitlinger.

Brandsma, H.P. (1993). *Basisschoolkenmerken en de kwaliteit van het onderwijs* (Characteristics of elementary schools and the quality of education; Dissertation, with a summary in English), Groningen: RION.

Brandsma, H.P. & Knuver, A.W.M. (1988). "Organisatorische verschillen tussen basisscholen en hun effect op leerprestaties" (Organizational differences between elementary schools and their effects on student achievement). *Tijdschrift voor Onderwijsresearch*, **13**, 201-212.

Cronbach, L. (1987). "Statistical tests for moderator variables: Flaws in analysis recently proposed." *Psychological Bulletin*, **102**, 414-417.

Dar, Y. & Resh, N. (1986). "Classroom intellectual composition and academic achievement". *American Educational Research Journal*, **23**, 357-430.

Duke, J.B. (1993). "Estimation of the network effects model in a large dataset". *Sociological Methods and Research*, **21**, 465-481.

Erbring, L. & Young, A.A. (1979). "Individuals and social structure: Contextual effects as endogeneous feedback". *Sociological Methods and Research*, **7**, 396-430.

Gamoran, A. (1986). "Instructional and institutional effects of ability grouping". *Sociology of Education*, **59**, 185-198.

Gamoran, A. (1992). "Is ability grouping equitable?" *Educational Leadership*, **50(2)**, 11-16.

Guldemond, H., Haanstra, F., Laarhoven, P. van & Vries, A.M. de (1987). *Effecten van heterogeen groeperen in het voortgezet onderwijs een voorstudie voor onderzoek in Nederland* (Effects of heterogeneous grouping in secondary education, a preliminary study for research in the Netherlands). Groningen: RION.

Grift, W. van de & Akkermans, W. (1991). "Onderwijskundig leiderschap en leerprestaties in het basisonderwijs" (Educational leadership and student achievement in elementary education) *Tijdschrift voor Onderwijsresearch*, *16*, 244-257.

Hallinan, M.T. (1987). "Ability grouping and student learning". In: M.T. Hallinan (Ed.): *The social organization of schools* (pp. 41-69). New York/London: Plenum.

Van der Hoeven-van Doornum, A.A. (1990). *Effecten van leerlingbeelden en streefniveaus op schoolloopbanen.* (Effects of teacher expectations and aspiration levels on school careers; Dissertation, with a summary in English) Nijmegen: ITS.

Jaccard, J., Turrisi, R. & Wan, C.K. (1990). *Interaction effects in multiple regression.* Newbury Park/London/New Delhi: SAGE Publications.

Jungbluth, P. (1993). "Pygmalion and the effectiveness of 'black' schools; teachers' stereotypes and hidden goal differentiations towards ethnic minorities". *Tijdschrift voor Onderwijsresearch*, **18**, 99-110.

Kerckhoff, A.C. (1986). "Effects of ability grouping in British secondary schools". *American Sociological Review*, **51**, 842-858

Knuver, J.W.M. (1993). *De relatie tussen klas- en schoolkenmerken en het affectief functioneren van leerlingen* (The relation between class and school characteristics and the affective functioning of the students; Dissertation, with a summary in English) Groningen: RION.

Kulik, C.C. & Kulik, J.A. (1982). "Effects of ability grouping on secondary school : a meta-analysis of evaluation findings". *American Educational Research Journal*, **19**, 415-428.

Kulik, C.C. & Kulik, J.A. (1984). "Effects of ability grouping on elementary school pupils: a meta-analysis. *Paper presented at the annual meeting of the American Psychological Association, Toronto.*

Leiter, J. (1983). "Classroom composition and achievement gains". *Sociology of Education*, **56**, 126-132.

Link, C.R. & Mulligan, J.G. (1991). "Classmates' effects on black student achievement in public school classrooms". *Economics of Education Review*, **10**, 297-310.

Maas, C.J.M. (1992). *Probleemleerlingen in het basisonderwijs.* (Pupils with learning and behavioral problems in elementary education; Dissertation, with a summary in English) Utrecht: ICS.

Meijnen, G.W. (1984). *Van zes tot twaalf. Een longitudinaal onderzoek naar de milieu- en schooleffecten van loopbanen in het lager onderwijs* (From six to twelve. A longitudinal study into the school effects and socio-economic background effects in elementary education). Harlingen: Flevodruk.

Mensen, T. & Guldemond, H. (1987). "Groepseffecten op individuele leerprestaties" (Classroom effects on individual achievement). *Mens en Maatschappij*, **62**, 176-199.

Oakes, J. (1985). *Keeping track. How schools structure inequality*. New Haven: Yale University Press.

Prosser, R., Rashbash, J. & Goldstein, H. (1991). *ML3 Software for Three-level Analysis*. London: Institute of Education, University of London.

Rath, J.C. (1991). *Minorisering: de sociale contructie van "etnische minderheden"* (Minorization: the construction of "ethnic minorities"; Dissertation with a summary in English). Amsterdam: SUA

Reezigt, G.J. (1993). *Effecten van differentiatie op de basisschool*. (Effects of within classroom differentiation in elementary schools; Dissertation, with a summary in English). Groningen: RION

Resh, N. & Dar, Y. (1992). "Learning segregation in junior high-schools in Israel: Causes and Consequences". *School Effectiveness and School Improvement*, **3**, 272-292.

Rowan, B. & Miracle, A.W. (1983). "Systems of ability grouping and the stratification of achievement in elementary schools". *Sociology of Education*, **56**, 133-144.

Slavin, R.E. (1987). "Ability grouping and achievement in elementary schools: a best-evidence synthesis." *Review of Educational Research*, **57**, 293-336.

Slavin, R.E. (1990). "Ability grouping and achievement in secondary schools: a best-evidence synthesis." *Review of Educational Research*, **60**, 471-499.

Velden, L.F.J. van der (1993). "The effects of teachers' attitudes and practices on the achievements of dutch low SES pupils". *Tijdschrift voor Onderwijsresearch*, **18**, 85-98.

De Vos, H. (1986). "Een model van het interactiesysteem van de schoolklas" (A model of the interaction system of the classroom). *Mens en Maatschappij*, **61**, 147-168.

De Vos, H. (1989). "A rational choice explanation of composition effects in educational research". *Rationality and Society*. **1**, 220-239.

De Vries, A.M. (1992). *"Hoe breder hoe beter?" De effecten van heterogeen groeperen in het voortgezet onderwijs in Nederland* (The broader the better? The effects of heterogeneous grouping in secondary education in the Netherlands; Dissertation with a summary in English). Groningen: RION.

Weide, M.G. (1993). "Effective provisions for ethnic minorities in primary schools". *Tijdschrift voor Onderwijsresearch*, **18**, 74-84.

Werf, M.P.C. van der & Weide, M.G. (1991). "Effectief onderwijs voor allochtone leerlingen" (Effective education for minority students). *Tijdschrift voor Onderwijsresearch*, **16**, 231-243.

Werf, M.P.C. van der, Weide, M.G. & Tesser, P. (1991). *Het onderwijsvoorrangsbeleid in de school en in de klas. De eerste meting in het basisonderwijs* (The Dutch Educational Priority Policy in the school and in the classroom. The first measurement in elementary education). Groningen/Nijmegen: RION/ITS.

Willms, J.D. (1986). "Social class segregation and its relationship to pupils'examination results in Scotland". *American Sociological Review*, **51**, 224-241.

APPENDIX: *Univariate statistics at the individual and classroom level*

| Variable | Mean | Median | Std. Dev. | Range | Minimum | Maximum |
|---|---|---|---|---|---|---|
| **1/ National Sample (3993 students, 221 classes)** | | | | | | |
| Individual Achievement Score | 20.562 | 21.111 | 5.205 | 25.444 | 4.556 | 30.000 |
| Class Mean Achievement | 20.188 | 20.475 | 3.129 | 15.643 | 10.308 | 25.951 |
| Individual SES | 00.000 | -0.371 | 6.695 | 25.000 | -11.871 | 13.129 |
| Class Mean SES | -0.304 | -0.423 | 3.737 | 20.196 | -11.167 | 9.029 |
| Individual IQ | 11.494 | 11.500 | 1.803 | 12.667 | 5.000 | 17.667 |
| Class Mean IQ* | 11.407 | 11.505 | 0.819 | 5.682 | 8.083 | 13.766 |
| Standard Deviation IQ* | 1.671 | 1.646 | 0.357 | 1.956 | 0.788 | 2.744 |
| **2/ Regional Sample (684 students, 57 classes)** | | | | | | |
| Individual Achievement Score | 534.360 | 536.000 | 10.771 | 49.000 | 501.000 | 550.000 |
| Class Mean Achievement | 533.069 | 534.143 | 6.997 | 30.626 | 514.556 | 545.182 |
| Individual SES | 00.000 | -4.035 | 9.972 | 38.190 | -11.455 | 26.735 |
| Class Mean SES | -1.004 | -2.068 | 6.127 | 27.296 | -9.035 | 18.261 |
| Individual IQ | 12.832 | 13.070 | 2.500 | 15.000 | 3.430 | 18.430 |
| Class Mean IQ | 12.691 | 12.751 | 1.310 | 5.552 | 9.126 | 14.678 |
| Stdandard Deviation IQ | 2.183 | 2.185 | 0.697 | 3.976 | 0.099 | 4.075 |
| Individual Teacher Goals | 3.606 | 3.780 | 0.721 | 4.000 | 1.000 | 5.000 |
| Class Mean Teacher Goals | 3.528 | 3.490 | 0.378 | 2.320 | 2.195 | 4.515 |

* These figures deviate from the ones in table 6.4.1. The figures in table 6.4.1 were obtained as follows: The classroom aggregates were disaggregated to the individual students, whereas in the appendix the figures relate to the frequency distributions across classes. Example: the average of the class mean IQ in the appendix is lower than in table 6.4.1, becuase the class mean IQ is lower in small classes. When computing the statistics in the appendix that relate to classroom aggregates across all classes, large or small, received the same weight. In this sense small classes are overrepresented.