GROUPING AND ITS EFFECT ON 8TH GRADERS' SCIENCE AND MATHEMATICS ACHIEVEMENTS

Ruth Zuzovsky
Center for Science and Technology Education,
School of Education, Tel Aviv University, Israel

Abstract

This study was initially intended to investigate the effect of the organization of learning, namely learning in heterogeneous home classes vs. learning in homogeneous ability groups, in science and mathematics studies in Israel on the "productivity" (overall level of achievement) and on the "inequality" (distribution of achievement within groups/classes) associated with social stratification. Findings showed a positive effect of grouping on the overall attainment and a differential positive effect on students' outcomes from low and medium SES backgrounds in science, while the use of ability grouping in mathematics did not show any significant overall or differential effect on achievement. In the course of the investigation some of my early assumptions were challenged yielding additional research questions and broadening my understanding of the role grouping serves in these two school subjects.

PURPOSE

This study was initially intended to investigate the effect of the organization of learning, namely learning in heterogeneous home classes vs. learning in homogeneous ability groups, in science and mathematics studies in Israel on "productivity" (overall level of achievement) and on "inequality" (distribution of achievement within groups/classes) associated with social stratification.[1] In the course of the investigation some of my early assumptions were challenged yielding additional research questions and broadening my understanding of the role grouping serves in these two school subjects.

THEORETICAL FRAMEWORK AND EDUCATIONAL OR ACADEMIC IMPORTANCE

To many educators ability grouping seems a sensible response to the academic diversity among students in that it allows teachers to attune their instruction to students' capabilities. High achievers are challenged and stimulated and low achievers get more support. Thus all students gain from ability grouping (Gamoran, 1986; Sorensen, 1970). Critics of ability grouping, however, contend that this practice has unintended harmful consequences. Ability grouping usually involves the assignment of poor teachers and inferior instruction to low ability groups (Finley, 1984; Oakes, 1985; Oakes, Gamoran, & Page, 1992; Page, 1991; Talbert, 1990). Moreover, critics assert that due either to the lower expectations of teachers who teach low-ability students (Rosenbaum, 1976), or to the stigmatizing effect ability grouping has on the self-esteem and aspiration of these students the achievement gap between students of high and low ability increases over time (Findley & Bryant, 1971; Gamoran & Berends, 1978; Hallinan, 1992; Murphy & Hallinger, 1989; Rosenbaum, 1980).

When segregation by academic criteria is associated with social and economic characteristics, grouping contributes and perpetuates social inequality (Braddock, 1990; Gamoran, Nystrand, Berends, & lePore, 1995; Oakes, 1990; Persell, 1977; Rosenbaum, 1980).
Research on the effect ability grouping has on achievement has taken two broad forms. One type compares achievement gains of students in grouped units of learning to ungrouped heterogeneous classes and another type focuses on the differential effect of grouping on low and high ability, and low and high social class.

Reviews on the effect of grouping vs. non-grouping have consistently shown that grouping has little or no impact on the overall student achievement in elementary and secondary schools (Slavin, 1987, 1990). Findings from the research on the differential effect of grouping are not conclusive (Hallinan, 1992).

The motivation to deal with the issue of grouping in Israel on the basis of TIMSS-1999 data stemmed from the fact that grouping was found to be quite widely used in Israel in both science and mathematics teaching, especially in the latter. Students from the 139 classes that participated in TIMSS-1999 were found to be taught by more than one teacher per class. For these 139 classes, there were 261 science teachers and 326 mathematics teachers – in most cases teaching separate groups of students. This indicated the spread of a grouping phenomenon at the lower secondary school level in Israel which is in contrast to a declared policy of social integration and of equal provision of educational opportunities for all students which was introduced into the educational system in the late 1960s. As part of this policy, middle (lower secondary) schools with pupils from different neighborhoods and home backgrounds were created. A recommended heterogeneous class composition and suitable instructional methods to deal with this heterogeneity were implemented. It was supposed that this policy would raise the achievement level of all students, close the persistent achievement gaps between students from low and high socio-economic backgrounds and create conditions for social integration. However, these expectations were not realized. The gap between students from an affluent background and students from deprived social and cultural backgrounds continues to exist.

Moreover, the pedagogical difficulties in managing the variation in student background and ability strengthened the tendency for segregation within this integration to re-occur. Public and academic debate regarding the gains and losses of the integration policy has been steady in Israel (Adler, 1984, 1986; Chen, Lewy, & Adler, 1978; Dar, 1981; Dar & Resh, 1986, 1988; Resh & Dar, 1990). Education-related ability grouping or segregation, which, to a significant extent, parallels (and reflects) social segregation phenomena in Israel, was found to have a small negative effect on academic achievement and a differential effect on the achievements of "strong" students who also come from an affluent background versus "weak" students who tend to come from a poor background: The strong students were shown not to have been affected by enhanced segregation while the weak students were apparently harmed (Resh & Dar, 1990). Segregation in Israel thus appeared to widen the achievement gaps and reduced the weak students' chances to enter high-level tracks in their secondary studies.

The publication of TIMSS-1999 data provided an opportunity to revisit the old debate and to study the overall effect of ability grouping on mathematics and science achievement and the differential effect of ability grouping on the achievements of students from different social backgrounds.

This aim was operationalized through four questions:

i. To what extent do different modes of learning organization (whole class, groups formed within or between home classes) enhance the overall level of achievement in science and mathematics studies?
ii. To what extent do the different modes of learning organization act to widen or close achievement gaps associated with social stratification?
iii. Do certain school/class, or group characteristics and especially those related to the organization of learning, act to attenuate the effect of student background variables on students' learning outcomes?
iv. Are there any interaction effects between the organization of learning and other higher level variables such as instruction that can explain the variability in outcome measures?

METHOD

Initial Phase

Preliminary analyses in the tradition of grouping vs. no grouping studies enabled me to formulate some working hypotheses. Comparing the achievement of individual students studying in groups to those studying in whole class settings revealed significant differences in favor of students studying in groups in both schools subjects. In a school level analysis the differences were found to be significant only in science (see Table 1).

Table 1: Achievement in Mathematics and Science by Type of Learning Organization (Student and School Level Comparison)

	Whole Class			Groups			*t-value* and *Sig*.
	N	Mean	*SD*	*N*	Mean	*SD*	
Mathematics individual scores	1379	458	(96)	2816	474	(87)	-5.60***
Science individual scores	2642	455	(89)	1553	499	(87)	-15.09***
Mathematics school mean	48	457	(66)	91	472	(50)	-1.33
Science school mean	90	453	(65)	49	498	(41)	-5.04***

In order to control the possible effect of students' socio-economic background on these differences, two-way analyses of variance were conducted separately for each school subject. The analyses were carried out both at the student level and at the school level. "Grouping" and "school SES" were the two independent variables, and the mean plausible[2] scores in mathematics and science were the dependent outcome variables. Results of the student level analyses in the two school subjects revealed significant main effects of both *grouping* and *socio-economic background* as well as a significant interaction effect between them. However, the effect of grouping is much smaller in mathematics than in science and at the school-level analysis, *grouping* in mathematics had no significant main effect at all (see Table 2), nor was there a significant interaction effect with the school's s*ocio-economic index*.

Table 2: Main Effects and Interaction Effects of Grouping and School SES on Variability in Mathematics and Science Achievements

Source of Variation	Student Level Analysis			School Level Analysis	
	DF	Mathematics *F & Sig*	Science *F & Sig*	Mathematics *F & Sig*	Science *F & Sig*
Main effects	3	106.8**	193.2**	10.4***	19.8***
Grouping (0-whole class; 1-grouping)	1	36.1***	285.0***	2.4	28.9***
School SES index (1-high; 3-low)	2	146.2**	150.2**	14.5***	15.1***
Interaction effect	2	8.0***	10.3***	0.6	1.3

Breakdown of achievements of subpopulations sorted according to the type of learning organization they were exposed to (whole class or groups) and according to the school's SES index (high, medium, or low) reveal that while grouping enhances the overall level of attainment in both school subjects it tends to favor more students from low and medium SES schools, especially in science (see Tables 3a, b).

Findings from these preliminary analyses supported the hypothesis that grouping enhances the overall level of attainment in both mathematics and science and at least in science it clearly acts in closing achievement gaps associated with social stratification.

Further Analyses

Acknowledging the hierarchical organizational structure of the educational system and the need to study simultaneously the effect of both student level variables and group/school level variables on the variability in school outcomes, I decided to apply hierarchical linear modeling (HLM) for this purpose. This modeling enables the simultaneous representation of outcome variation at each level of the educational hierarchy in one model and the identification of the important variables "explaining" this variation at each level in terms of the magnitude of their regression coefficients (for a more detailed account on HLM, see Bryk & Raudenbush, 1992). Allowing the regression coefficients of student level variables (slopes) to vary randomly over groups or schools can detect variation in these regression coefficients that can be explained by the effect of higher level variables in a "slopes as outcomes" model (Raudenbush & Bryk, 1986). This "slopes as outcomes" model is equivalent to an interaction model where student level variables with random slopes and group/school level variables interact. If we specify, in the random models, additional possible interaction effects among the group/school variables themselves, this will cause the hierarchical linear modeling of school effects to become more interactive (Aitkin & Zuzovsky, 1994). In such a model educational outcomes are not considered simply as outcomes of additive main effect models, but rather as products which are non-consistently dependent on a particular combination of pupil characteristics and school characteristics, thus generating a more sensitive and accurate picture of the schooling phenomenon.

At the student level, I estimated the effect of socio-economic background variables (number of books at home, parental education) gender, ethnic affiliation, self esteem and academic aspiration, on the variance in an individual's achievement scores.

At the school-class or group level, I estimated the effect of the student-body characteristics, the school – SES index, the use of grouping technology, types of instruction, and disciplinary climate – on the variability in school means (intercepts) and on the variability in the slopes of one selected student level variable (self esteem as a learner of sciences or mathematics) on achievement. In order to facilitate a straightforward interpretation none of the variables used in this analysis were centered.

Analysis

After selecting and constructing the relevant variables and ensuring a complete data set for the analyses, we allowed all student level variables to vary among the second level units. The small variation found convinced me to specify a fixed two-level model of students nested within groups or schools, which dealt simultaneously with the "within-units-of-analysis" and "between-units-of-analysis" variances in mathematics and science achievement. For each school subject, a null model (Model (a)) enabled me to estimate, simultaneously, the variance in achievements at each of the model's levels. I then specified several models and estimated both their explanatory power and the regression coefficients of their specified variables on the dependent variables: the mean plausible scores of science and mathematics achievement. In Model (b), only student level variables were introduced to the regression equation. In Model (c) aggregates of these variables describing the student body composition were added. In Model (d) schooling variables describing disciplinary climate, the implementation of ability grouping and indices describing instruction were added to the previous sets of variables. Models (d) and (c) were considered alternatives and were compared with Model (b) (individual background variables only). The last model, Model (e) included all effective significant variables found at both levels in previous analyses.

Table 3a: Achievement by Grouping and School SES (Student Level)

Schools	Mathematics Achievement							Science Achievement						
	Whole Class			Grouping				Whole Class			Groups			
	N	Mean	*SD*	*N*	Mean	*SD*	Diff.	*N*	Mean	*SD*	*N*	Mean	*SD*	Diff.
All	1379	458	(96)	2816	474	(87)	16	2642	455	(89)	1553	499	(87)	44
High SES	646	488	(83)	943	492	(83)	4	1003	481	(87)	586	513	(83)	32
Medium SES	211	440	(95)	1131	475	(93)	35	834	454	(92)	508	512	(83)	58
Low SES	463	423	(101)	650	441	(91)	18	654	406	(107)	459	466	(87)	60

Table 3b: Achievement by Grouping and School SES (School Level)

Schools	Mathematics Achievement							Science Achievement						
	Whole Class			Grouping				Whole Class			Grouping			
	N	Mean	*SD*	*N*	Mean	*SD*	Diff.	*N*	Mean	*SD*	*N*	Mean	*SD*	Diff.
All	48	457	(66)	91	472	(50)	15	90	453	(65)	49	498	(42)	45
High SES	21	490	(45)	29	490	(42)	0	32	481	(43)	18	511	(37)	30
Medium SES	8	479	(71)	36	475	(40)	4	28	446	(53)	16	512	(32)	66
Low SES	17	420	(72)	23	439	(60)	19	25	404	(79)	15	468	(42)	64

In separate analyses I allowed the slopes of all individual background variables to vary randomly across groups. Only one variable ("self-esteem" as a science or mathematics learner) was found to have significant, although small slope variation. This variation enabled me, thereafter, to specify a random model and to trace some interaction effects between self-esteem as science and mathematics learners and some higher level variables. The last step of the analysis was to specify in the group-level random models, two theoretically important interactions between grouping and the two indices of instruction, and look for their effects.

With these additional analyses I followed claims that attribute the effect of grouping to two types of mechanisms: "institutional" – usually mediated through symbolic properties that affect self-esteem, expectations and motivations of students and teachers in the different learning organizations, and "instructional" which is supposed to be carried out differently in the different types of learning organizations (Gamoran, 1986).

DATA SOURCE

The data source consists of student responses and information collected in TIMSS-1999 Trend Study in Israel. Organization of learning in small groups occurred, the case of mathematics in 91 schools, and in science in only 49 schools (out of 139 participating schools). Schools differed in the number of groups formed from one home-class and also in the number of students attending each group. Some groups hosted a very small number of students. Due to statistical constraints, I excluded from the HLM analysis small groups with less than seven students and were left with 203 mathematics groups of which 44 were home-class units and 157 small groups the product of grouping policy. In science there were 199 groups of which 92 were home-class units and 107 small groups. Each of these groups was taught by a different teacher. I restricted my analysis to a sample of students with full data –3,159 students in mathematics and 3,242 students studying science. Most of the student variables derived from student questionnaires describing student characteristics were of a general nature, except for student self-esteem which was specified separately for science and mathematics studies.

Group level variables determining class/group composition, disciplinary climate and instruction were also derived from aggregated or averaged student responses. This enabled me to create a file with no missing data for the higher level variables of the model. Variables describing instruction which were built on student perceptions were specified separately for science and mathematics. Out of their responses, two factors for each school subject – one describing *teacher-led* instruction and the other *student centered learning* – were constructed. Another group level variable, an index of positive school-climate, was also based on student perception.

Science and mathematics achievement scores were the dependent variable, the mean of five estimations of plausible scores in each of these subject areas. A full description of the variables employed in the analyses appears in the appendix.

RESULTS AND INTERPRETATIONS

We now present the findings from the second phase of the analysis in which hierarchical linear modeling of school and schooling effects (Sorensen & Morgan, 2000) was employed. This analysis was conducted separately, once decomposing the variance between and within groups of students taught by different teachers (n = 203 in mathematics and n = 199 in science) and another time when the variance was decomposed into between and within school components (n = 134 in both mathematics and science). I will first present findings from the analyses of the fixed models (where the effect of lower level variables on achievement is assumed to be constant over the higher level units of analysis). (For details see Tables A, B, C, D. in the appendix.)

The Between-Group/Between- School/ Variance in Achievements and the Percentage Thereof Explained by the Different Models

- The percentage of *between-groups* variance out of the total variance in outcomes is higher for mathematics (44%) than for science achievement (36%). However, the *between schools/classes* variance component is similar in the two school subjects (38% of the total variance in mathematics and 36% in science). It should be noted here that while the between-schools variance in mathematics found in Israel is similar to the international average (Martin, Mullis, Gregory, Hoyle, & Shen, 2000), in science the between-schools variance component found in Israel is much higher than the international average (23%) that was estimated by the same source.
- Several models were used to explain this variability. First I estimated and controlled for the effect of student-level variables. These models were found to explain about half of the between-groups or -schools variation in achievement in both school subjects. The other specified models contained in addition to the individual student level variables, their aggregates or means at the group or school level, representing the student body composition variables. Another model introduced, in addition to the individual student-level variables, only the group or school indices representing a variety of school factors such as: disciplinary climate, SES, type of instruction and the type of learning organization employed. The full model contained, beyond individual student-level variables, all the significant variables that were found in the previous analyses. Each of these models have their own explanatory power. Table 4 presents the amount of between-groups or schools outcome variance that was explained by each of the alternative models (see Table 4 based on Tables A, B, C, D in the Appendix). Accordingly, the data in the table indicate that:
  - The explanatory power of the combined full models is the highest. They explain, in almost all of the cases, about three-quarters of the between-group or -school variance in achievement.
  - Group composition variables add to the already explained variance due to individual student level variables – about 16% to 22% more of the unexplained variance that is left.
  - The net contribution of school or group variables that represent the conditions for learning set by the school and the type of instruction that takes place in groups or classes is minor, reaching no more than 8% of the unexplained between-groups/schools variance in achievement left beyond the effect of the individual student variables.

Variables Explaining the Between Groups/Schools Variance in Achievement

- Pupil-level variables in order of the magnitude of their effects are: ethnic origin; number of books at home; self-esteem as a pupil; educational academic aspiration; gender, and parental education. Findings in Tables 5, 6, and 7 are based on data in Tables A, B, C, & D in the Appendix.

Table 4: Explanatory Power of Several School Effectiveness Models

	Mathematics		Science	
	Groups Model $n = 203$	Schools Model $n = 134$	Groups Model $n = 199$	Schools Model $n = 134$
Between Groups/Schools Variance Component	44%	36%	36%	38%
Model Specifications	Percentage of Between Groups/Schools Variance Explained by the Different Models			
Model 1: Individual student variables	50%	55%	49%	52%
Model 2: Model 1+ group composition variables	72%	73%	65%	71%
Model 3: Model 1+ group/school variables	58%	63%	57%	63%
Model 3 Model 1+ all significant variables	75%	76%	69%	74%

Table 5 presents the regression coefficients of these variables in the fixed science and mathematics models at the group and school level models. The effects of student-level variables are similar in both school subjects:

- Jewish students achieve about one standard deviation higher than Arab students.
- Students in homes with more than 200 books achieve between one-third to half standard deviation more than students from homes with less than 10 books.
- Students with high self-esteem achieve about one-third standard deviation more than students with low self-esteem.
- Students with high academic aspirations achieve around a quarter of a standard deviation more than students with no academic aspirations at all.

Table 5: Regression Coefficients and Sig. of Student Level Variables in the Group and School – Fixed Models

Student Variables	Mathematics		Science	
	Group Model $n = 203$	School Model $n = 134$	Group Model $n = 199$	School Model $n = 134$
Ethnic origin 0 = Arabic; 1 = Jewish	92.7***	88.8***	100.8***	101.7***
Number of books at home 1 = less than 10; 5 = more than 200	6.9***	8.2***	10.9***	10.3***
Self-esteem as a pupil 0 = low; 1 = high	34.0***	33.9***	36.1***	36.4***
Educational academic aspiration 0 = low; 1 = high	21.9***	26.8***	27.4***	28.4***

- Table 6 presents the regression coefficients of the student-body composition variables in all models studied; among them are: percentage of *parents with academic education* and the groups or school average of the *number of books in the students home*, both usually associated with students' socio-economic background are the strongest and most significant predictors of achievement.
  - In groups or classes where all students have at least one parent with academic education, the group mean achievements are almost one standard deviation higher than in groups or classes where none of the students have parents with academic education.
  - In groups or classes where all members have more than 200 books at home the group mean achievements are about one standard deviation higher in mathematics than those in groups where all members come from homes with little books. This group composition variable is somewhat less effective in science.

Table 6: Regression Coefficients (and sig.) of Group Composition Variables in the Group and School Fixed Models

Group Composition Variables	Mathematics				Science			
	Group Model *n* = 203		School Model *n* = 134		Group Model *n* = 199		School Model *n* = 134	
No. of books at home/Mean	**22.4**	**(6.4)*****	**19.5**	**(8.8)***	**14.7**	**(7.2)***	16.5	(10.4)
At least one parent with academic education (%)	**81.0**	**(15.5)*****	**100.6**	**(21.4) *****	**92.5**	**(16.3)*****	**121.7**	(23.1) ***
Students with high academic aspirations (%)	32.3	(15.3)	-4.3	(26.6)	3.2	(20.2)	-8.7	(29.8)
Students with high self-esteem as learners (%)	-0.4	(14.1)	-23.6	(24.9)	0.5	(17.7)	-20.6	(23.7)

- School or group level variables describe the following characteristics:

Disciplinary climate (average student responses to a battery of questions about disciplined climate or lack of violence they experience in school (1 = *low discipline* – 4 = *high*).
Learning in groups (0 = *whole integrated class* or *school that implements integration*; 1= *group or school that uses grouping technology*).
Types of instruction, only used in the group models – factors describing frequency of student centered vs. teacher-led instruction (1= *seldom*; 4 = *often*).
"School SES index", only used in the school models (1 = *high*; 3 = *low*).

The regression coefficients of these variables in the fixed-group and school models are presented in Table 7.

- In mathematics and science, student-centered instruction (negatively associated with achievement), plays a significant role in explaining the between-group variability in outcomes. Wherever this type of instruction is frequent, students achieve less (about half of a standard deviation for each unit of the index).
- Disciplinary climate (positively associated with achievement) explains the variance in outcomes only in the case of mathematics. The achievement gap between students studying in high disciplinary climate groups and students studying in low disciplinary climate groups reaches one standard deviation in favor of the high *disciplinary climate*.
- Organizing learning in groups is positively associated with achievement only in science. Students studying science in groups achieve between one-fifth to one-quarter of a standard deviation more than students studying in whole-class settings.
- As expected, the school SES index is also significant. Students studying in high-SES schools achieve about a quarter of a standard deviation more than students in low-SES schools.

Table 7: Regression Coefficients of Significant School or Group Variables in the Group and School Fixed Models

Group/School Variables	Mathematics		Science	
	Group Model *n* = 203	School Model *n* = 134	Group Model *n* = 199	School Model *n* = 134
Disciplinary climate (1=low; 4 = high)	**20.4 (9.4)***	16.8 (12.2)	14.2 (11.6)	16.5 (13.2)
Grouping (0-1)	-3.5 (6.9)	-3.6 (6.8)	**23.9 (6.4)*****	**20.2 (7.5)****
Pupil centered instruction (1=low; 4 = high)	**-58.6 (11.9)*****	- -	**-41.6 (10.8)*****	- -
Teacher led instruction (1=low; 4 = high)	-19.9 (11.8)	- -	-13.2 (10.7)	- -
School SES index (1=high; 3=low)	- -	**-5.3** (1.3)***		**-6.3 (1.4)*****

Group and School Variables and their Effect on the Association Between Ethnic Origin and Achievement

The initially intense but fading strength of the association between ethnic origin and achievement as a result of introducing higher-level variables to the alternative models of school effectiveness points to the role school and group variables can play in reducing the high achievement gap found between Hebrew-speaking students and Arabic-speaking students.

These group/school-level variables are not associated with socio-economic factors rather they represent the ways schools operate. Their effects are consistent in both Arabic and Hebrew speaking populations. *Disciplinary climate* and *learning in groups* are positively associated with achievement while *student-centered instruction* is negatively associated with achievement.
Comparing the two populations in terms of these variables (see Table 8) shows that all variables that are positively associated with achievement are lower in the Arabic-speaking groups, while those negatively associated with achievement are higher in these groups. Without ignoring the fact that there are significant differences in socio-economic characteristics between Arabic-speaking and Hebrew-speaking populations which also can explain the achievement gap between the two populations, it is evident that schools can play a role in closing these gaps by providing adequate conditions for learning. With less frequent student-centered instruction, high discipline in mathematic classrooms and more small-group learning in science, there is a chance of narrowing these achievement differences.

Table 8: Comparing Group Statistics of Arabic and Hebrew Speaking Groups

	Arabic (Math *n*-38; Science *n*-28)	Hebrew (Math *n*-171; Science *n*-171)	*t* & *Sig.*
Disciplinary climate – science	1.7 (.19)	3.3 (.34)	-23.1***
Disciplinary climate - math	1.8 (1.8)	3.2 (.38)	-33.9***
Studying in groups - science	7% -	61% -	417.5**
Studying in groups - math	33% -	72% -	395.2***
Pupil centered instruction - science	2.2 (.25)	1.6 (.27)	9.8***
Pupil centered instruction - math	2.1 (.27)	1.6 (.27)	10.7***
Teacher-led instruction - science	3.1 (2.0)	3.2 (.27)	-.18
Teacher-led instruction – math	3.3 (.18)	3.5 (.26)	-6.0***
Percent of parents with academic education	26% (.13)	48% (.22)	-7.6***

Findings from the Random Models

Running a random model was meant to trace some interaction effects between student-level variables and upper-level group or school variables. Such interaction effects can demonstrate how certain group or school characteristics act to attenuate the effects of lower-level variables, primarily those associated with socio-economic factors, on student outcomes. However, the effects of most student-level variables used in this analysis on the variability in outcome measures were found to be constant across groups or schools, except for the effect students self-esteem as a science or mathematics learner. The slope of this variable, which is not necessarily associated with socio-economic factors, showed a small but significant variability among the higher units of analysis. Thus, I specified a set of models similar to those used in the previous analyses allowing the regressions coefficients of science and mathematics achievement on students' self-esteem as learners in these subjects to vary randomly across groups or schools (in addition to groups' school intercept).

Academic self-esteem as a science or mathematics learner can be regarded either as a realistic measure of one's own level of attainment or as a symbolic property associated with the group the individual student is affiliated to. It is positively associated with achievement (Shavelson & Bolus, 1981) in both school subjects. This association was found to be stronger in groups or classes with a high disciplinary climate. In science this relationship is also stronger in groups with high mean of teacher-led instruction (see random model Tables F and H in the Appendix). Since disciplinary climate is higher in the Hebrew-speaking than in the Arabic-speaking groups or classes studying science and mathematics (see Table 9), the association between academic self-concept and achievement in both subject areas is stronger in Hebrew-speaking classes. This occurs in spite of the fact that there are more students with high self-esteem in the Arab classes than in the Hebrew ones (55% vs. 34% in science and 69% vs. 71% in mathematics). Disciplinary climate in both school subjects is also higher in a group organization of learning (see Table 10), which is more frequent in the Hebrew-speaking population, thus, the advantage of Hebrew-speaking students over Arabic-speaking students gets even larger.

Table 9: Differences in Group Means of Disciplinary Climate and Group Aggregates of Students with High Self-Esteem as Learners in Science and Mathematics

Group Characteristics	Arab Speaking			Hebrew Speaking			*t-value* and *Sig.*
	Mean	(*SD*)	*N*	Mean	(*SD*)	*N*	
Teacher led instruction – math	3.3	(.18)	34	3.5	(.26)	171	-6.0***
Teacher led instruction – science	3.1	(.19)	28	3.2	(.77)	171	-1.8
Disciplinary climate in mathematics	1.8	(.19)	34	3.2	(.38)	171	-33.9***
Disciplinary climate in science	2.2	(.25)	28	1.6	(.27)	171	-23.1***
Percent of students with high self-esteem as mathematics learners	51%	(16)	34	44%	(17)	171	2.1*
Percent of students with high self-esteem as science learners	55%	(14)	28	35%	(15)	171	6.7***

Table 10: Differences in Group Means of Disciplinary Climate and Group Aggregates of Students with High Self-Esteem in Science and Mathematics

Group Characteristics	Whole Class			Groups			*t-value* and *Sig.*
	Mean	(*SD*)	*N*	Mean	(*SD*)	*N*	
Teacher led instruction – Mathematics	3.4	(.22)	47	3.5	(.28)	156	-1.1
Teacher led instruction – Science	3.2	(.25)	92	3.2	(.28)	107	-2.2*
Disciplinary climate in mathematics	2.6	(.79)	47	3.1	(.56)	156	-3.8***
Disciplinary climate in science	2.7	(.71)	92	3.3	(.38)	107	-6.6***
Percent of students with high self-esteem as mathematics learners	46%	(14)	47	45%	(18)	156	0.6*
Percent of students with high self-esteem as science learners	39%	(16)	92	37%	(16)	107	0.8

Findings from the Interactive Models

Among the possible interaction terms between the group or school level variables, I decided to focus on the interaction between *grouping* and *instruction.* I followed Gamoran's response to Slavin's (1987) "Best Evidence Synthesis" on grouping, where he comments:

The study of grouping alone provides little information of value. What appear to be positive, negative or neutral effects of grouping may have little to do with grouping in and of itself, but may derive from how grouping is used to provide appropriate or inappropriate instruction. (p. 342)

Instruction was also found to be an important factor of the explanation for widening achievement gap among students assigned to groups at different educational levels (Gamoran, Nystrand, Berends, & Lepore, 1987; Oakes, 1985; Page, 1999).

I then ran another alternative random model for the specified variable adding two interaction terms between *grouping* and the *two indices of instruction*: *student-centered* and *teacher-led*.

The random interactive models were found to have maximal explanatory power regarding the between-group achievement variance (See Tables E and F (model f) in the Appendix.) They explain, in the case of mathematics, an additional 26% of the between-group yet unexplained variance to the already 48% of between-variance explained by the individual student-level variables (48% + 25% = 73%). In science, this model explains 28% more of the between group variance in addition to the 40% explained by the student variables model (40% + 28% = 68%).

Only one significant and positive interaction effect in science was found between *grouping* and *instruction*. Students working in small groups in science achieve about half a standard deviation more for each unit of the student centered type of instruction they receive than if they study in a whole-class setting. Such interaction effects do not occur in mathematics.

QUESTIONS, FURTHER INVESTIGATION AND CONCLUSIONS

The main purpose of this study was to look for the effect of grouping that was found to be very common in mathematics classes and somewhat less so in science classes in Israel

Supported by the results of my preliminary analyses that showed the positive effect of grouping on the overall attainment in both school subjects and a differential positive effect on students from low and medium SES schools, especially in science, I was puzzled by the contrasting findings from the HLM analyses that did not show any significant overall or differential effect of grouping in mathematics. This led me to further investigate my assumption of the common nature of grouping in both subject areas.

I conducted a set of analyses of variance in all schools where grouping was implemented in both school subjects using the mean plausible score in mathematics as my dependent variable (serving as a proxy for ability). I indeed found that while in mathematics the groups were found to be statistically different from each other in their mathematics scores (ability measure), in science, this was not the case. Science groups formed from one home-class did not differ in their mathematics (ability) mean. This led me to believe that while grouping in mathematics is basically ability grouping, a reaction to initial individual (ability or academic) differences that exists among students in the heterogeneous lower secondary schools, grouping in science is not associated with initial ability differences among students, but rather results from other considerations. These considerations are related to the empirical nature of science and the collaborative working habits of scientists that consequently lead to a preferred mode of learning in science, i.e., small-group learning which enables intensive hands-on learning, foster helping relations, intensive peer interaction and increased information processing (Springer, Stanne, & Donovan, 1999).

Organizing learning groups in science can thus be viewed as a school factor open to manipulation, that can advance attainment. Findings from my study indeed support this view as *grouping* was found to affect both the productivity and equality of science outcomes.

The use of ability grouping in mathematics, on the other hand, seems to be in line with the individual difference paradigm that accepts differential outcomes and is led by the belief that individual differences in ability account for the vast range of performance existing both within and between schools (Miller & Brookover, 1986; Persell, 1977). Grouping in mathematics is a responsive technique necessary to conduct meaningful instruction that only provides a decent opportunity for every student to learn to their potential; thus it does not advance students beyond what they can attain, nor does it close the gaps between high and low ability students.

References

Adler, H. (1984). School integration in the context of the development of Israel's educational system. In Y. Amir, S. Sharon, & R. Ben Ari (Eds.), *School desegregation* (pp. 21-46). New Jersey: Erlbaum.

Adler, H. (1986). Israeli education addressing dilemmas caused by pluralism: A sociological perspective. In D. Rothermund & J. Simon (Eds.), *Education and integration of ethnic minority* (pp. 64-87). London: Frances Pinter.

Aitkin, M., & Zuzovsky, R. (1994). Multilevel interaction models and their use in the analysis of large-scale school effectiveness studies. *School Effectiveness and School Improvement, 5* (1), 45-73.

Braddock, J.H. (1990). Tracking the middle grades. National patterns of grouping for instruction. *Phi Delta Kappan,* (Feb.), 445-449.

Bryk, A.S., & Raudenbush, S.W. (1992). *Hierarchical linear models.* Newbury Park, CA: Sage.

Chen, M., Lewy, A., & Adler, C. (1978). *Processes and outcomes in the educational system.* School of Education, Tel Aviv University, Tel Aviv and The NCJW Research Institute for Innovation in Education, School of Education, The Hebrew University, Jerusalem. (Hebrew)

Dar, Y. (in collaboration with Resh, N.) (1981). *Homogeneity and heterogeneity in education.* The NCJW Research Institute for Innovation in Education, School of Education, The Hebrew University of Jerusalem, pub. No. 73. (Hebrew)

Dar, Y., & Resh, N. (1986). Classroom intellectual composition and academic achievement. *American Educational Research Journal, 23,* 357-374.

Findley, W.G., & Bryant, M.M. (1971). *Ability grouping: 1970.* Athens, GA: Center for Educational Improvement.

Finley, M.K. (1984). Teachers and tracking in a comprehensive high school. *Sociology of Education, 57,* 233-243.

Gamoran, A. (1986). Instructional and institutional effects of ability grouping. *Sociology of Education, 57,* 233-243.

Gamoran, A., & Berends, M. (1987). The effect of stratification in secondary schools: Synthesis of survey and ethnographic research. *Review of Educational Research, 57* 415-435.

Gamoran, A., & Mare, R.D. (1989). Secondary school tracking and educational inequality: Compensation, reinforcement, or neutrality? *American Journal of Sociology, 94,* 1146-1183.

Gamoran, A., Nystrand, N., Berends, M., & Le Pore, P.C. (1995). An organizational analysis of the effects of ability grouping. *American Educational Research Journal, 32* (4), 687-715.

Hallinan, M.T. (1992). The organization of student for instruction in the middle school. *Sociology of Education, 65,* 114-127.

Martin, M.O., Mullis, I.V.S., Gregory, K.D., Hoyle, C., & Shen, C. (2000). *Effective schools in science and mathematics – IEA's TIMSS.* Chestnut Hill, MA: TIMSS International Study Center, Boston College.

Miller, S.K., & Brookover, W.B. (1986). *School effectiveness vs. individual differences paradigmatic perspectives on the legitimation of economic and educational inequalities.* Paper presented as part of Symposium at the AERA Annual meeting, San Francisco.

Murphy, J., & Hallinger, P. (1989). Equity as access to learning: Curriculum and instructional treatment differences. *Journal of Curriculum Studies, 19,* 341-360.

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

Oakes, J. (1990). *Multiplying inequalities: The effect of race, social class and tracking on opportunities to learn mathematics and science.* Santa Monica: Rand.

Oakes, J., Gamoran, A., & Page, R.N. (1992). Curriculum differentiation: Opportunities, outcomes, and meanings. In P.W. Jackson (Ed.), *Handbook of research on curriculum* (pp. 570-608). Washington, DC: American Educational Research Association.

Page, R.N. (1991). *Lower track classroom: A curricular and cultural perspective.* New York: Teachers College Press.

Persell, C.H. (1977). *Education and equality: A theoretical and empirical synthesis.* New York: Free Press.

Persell, C.H. (1977). *Education and equality: The roots and results of stratification in America's schools.* New York: The Free Press.

Raudenbush, S.W., & Bryk, A.S. (1986). A hierarchical linear model: A review. *Sociology of Education, 59,* 1-17.

Resh, N., & Dar, Y. (1990). *Segregation within integration. Educational separation in junior high schools. Factors and implications*, Publication No. 125, The NCJW Research Institute for Innovation in Education, School of Education, The Hebrew University, Jerusalem. (Hebrew)

Rosenbaum, J.E. (1976). *Making inequality: The hidden curriculum of high school tracking.* New York: Wiley & Sons.

Rosenbaum, J.E. (1980). Social implications of educational grouping. *Review of Research in Education, 8,* 361-401.

Shavelson, R.J., & Bolus, R. (1981). Self concept, interplay of theory and methods. *Journal of Educational Psychology, 74,* 3-17.

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). Achievement effects of ability grouping in secondary schools. A best evidence synthesis. *Review of Educational Research, 60,* 471-499.

Sorensen, A.B. (1970). Organizational differentiation of students and educational opportunity. *Sociology of Education, 43,* 355-376.

Sorensen, A.B., & Morgan, S.L. (2000). School effects theoretical and methodoical issues. In M.T. Hallinan (Ed.), *Handbook of the sociology of education.* Kluwer/ Plenum.

Springer, L., Stanne, M.E., & Donovan, S.S. (1999). Effects of small group learning on undergraduates in science, mathematics, engineering and technology: A meta-analysis. *Review of Educational Research, 69* (1), 21-52.

Talbert, J. (1990). *Teacher tracking: Exacerbating inequalities in the high school.* Paper presented at the AERA annual meeting. Boston.

Yamamoto, K., & Kulick, E. (2000). Scaling methodology and procedures for the TIMSS mathematics and science scales. In M.O. Martin, K.D. Gregory, & S.E. Stemler (Eds.), *Technical report.* Chestnut Hill, MA: ISC, Boston College.

NOTES

1. The terms *productivity* and *inequality* were used by Gamoran and Mare (1989) to reflect the two sides of the main debate concerning ability grouping.

2. Estimation of students proficiency scores in science and mathematics based on IRT scaling (Yamamoto, & Kulick, 2000).

APPENDIX

LEVEL 1: STUDENT VARIABLES

Social background variables, self-esteem and academic aspiration, gender and ethnic affiliation.

BSBGBOOK: No. of books in student's home (1=0-10; 2=11-25; 3=26-100; 4=101-200; 5=200+books)

SEX: (0 = boy; 1= girl)

ISRARB: (1-Hebrew speaking; 2-Arabic speaking)

ACADP: Parental academic education (One parent at least with academic education = 1; none = 0)

S/MGOOD-D: Self-esteem as science or mathematics learner – a dichotomous variable (I usually do well in Science/Mathematics. 0 = do not agree; 1 = strongly agree)

BSBGEDSE-D: Academic aspiration – a dichotomous variable (How far do you expect to go in school? (Reversed) 0 = less than university; 1 = complete university)

LEVEL 2: GROUP/SCHOOL VARIABLES

Disciplinaric class/group climate, class organization variable, instructional variables, class/group composition variables.

BEHAVE: Index - How often the following occurred the last month in school? Skipping a class, something was stolen, hurt by student, friend skipping a class, friends had something stolen, friends were hurt (l = more than 5 times; 4 = never)

GROUPING: Class organization in science or mathematics: (0 = whole class; 1 = ability groups)

FACTR1-S: Student centered learning in science class (mean index: 1 = never; 4 = always)
FACTR1-S: Teacher led instruction in science class (mean index – 1 = never; 4 = always)

FACTR1-M: Student centered learning in mathematics class (mean index: 1 = never; 4 = always)
FACTR2-M: Teacher led instruction in mathematics class ((mean index: 1 = never; 4 = always)

GOOD-P: Percent of students with high self-esteem in group or class

BGEDSE-P: Percent of students with academic aspiration in group or class

ACADP-P: Percent of families with at least one parent with academic education in group or class

BGBOOK: Index – Number of books at home (mean: 1 = few; 5 = more than 2000)

Level 1 – Students' Descriptive Statistics

Variable	Science			Mathematics		
	N	Mean	*SD*	*N*	Mean	*SD*
BSBGBOOK - No. of books at students' home (1 = few; 5 = more than 2000)	3864	3.39	(1.17)	3463	3.41	(1.17)
ACADP - One parent at least with academic education	3480	0.44	(0.50)	3159	0.44	(0.50)
SEX (1=girl; 2= boy)	4061	0.50	(0.50)	3625	0.51	(0.50)
ISRARB (1=Hebrew; 2=Arabic)	4061	0.80	(0.40)	3626	0.79	(0.41)
GOOD-D - Usually does well in subject (0=do not agree; 1=agree)	3877	0.38	(0.49)	3495	0.47	(0.50)
BGEDSE-D - How far you expect to go at school (0=less than university; 1=university)	3242	0.69	(0.46)	2938	0.71	(0.45)

Level 2 – Group Descriptive Statistics

Grouping Variable Name	Science			Mathematics		
	N	Mean	*SD*	*N*	Mean	*SD*
Percent of small groups	199	0.54	(0.50)	203	0.77	(0.42)
BEHAVE - Index of disciplinary climate (1=negative ; 4=positive)	199	3.05	(0.62)	203	2.98	(0.65)
FACTR 1 Student centered instruction (1=never; 4=always)	199	1.72	(0.32)	203	1.68	(0.34)
FACTR 2 Teacher centered instruction (1=never; 4= always)	199	3.21	(0.27)	203	3.48	(0.24)
GOOD-P - Percent of high self-esteem students)	199	0.38	(0.16)	203	0.46	(0.17)
BGEDSE-P - Percent of students with academic aspiration	199	0.69	(0.16)	203	0.70	(0.19)
ACADP-P - Percent of students with academic education)	199	0.45	(0.25)	203	0.44	(0.24
BGBOOK- Group mean	199	3.38	(0.46)	203	3.48	(0.27)

INSTRUCTIONAL FACTORS
(Extraction Method: Principal component/Analysis
Rotation Method: Varimax with Kaiser Normalization)

SCIENCE

FACTR 1 (Main variables)

Student Centered Learning in Science Lessons		
Variable Name and Id. No.	Description	Loading
SUSOS (20)	Student use overhead projector in your science lessons	.78
SIDEA (22)	Teacher uses computer to demonstrate ideas	.72
SHWFC (13)	Students check each other's homework	.66
SCALC (6)	Students use calculators	.65
SCOMP (7)	Students use computers	.65
SUSOT (18)	Teachers use overhead projector in science lessons	.64

Standardized α coefficient = .80

FACTR 2 (Main variables)

Teacher Led Instruction in Science Lessons		
Variable Name and Id. No.	Description	Loading
SHWGV (10) SHWTC(12)	Teacher gives homework in science lesson	.67
SNOTE (2)	Teacher checks homework in science lesson	.67
SUSBT (17)	Student copies notes from the board	.59
SDEMO	Teacher uses board in science lesson	.60
	Teacher gives demonstrations	.53

Standardized α coefficient = .69

Total Variance Explained

FACTR 1	19.3
FACTR 2	16.1
Total	35.4

Descriptive Statistics

	Mean	SE	SD
FACTR 1	1.7	.01	.73
FACTR 2	3.2	.00	.58

Correlation Matrix

	Science Mean Plausible	FACTR 1	FACTR 2
Science Mean Plausible	1.0		
FACTR 1	-.32***	1.0	
FACTR 2	-.03	.11***	1.0

MATHEMATICS

FACTR 1 (Main variables)

Student Centered Learning in Math Lessons		
Variable Name and (Id.)	Description	Loading
MUSOS (18)	Student uses overhead projector in math lessons	.75
MUSOT (16)	Teacher uses overhead project in math lessons	.72
MCOMP (7)	Students use computers in math lessons	.66
MIDEA (20)	Teachers use computer to demonstrate ideas in math	.64
MPROJ (4)	Students work on math projects in math lessons	.60
MHWFC (13)	Students check each other's homework in math lessons	.60
MSGRP ( )	Students work in pairs or small groups	.50

Standardized α coefficient = .78

FACTR 2 (Main variables)

Teacher Led Instruction in Math Lessons		
Variable Name and (Id.)	Description	Loading
MHWTC (12)	Teacher checks homework in math lessons	.56
MUSIBT (17)	Teacher uses board in math lesson	.54
MHWGV (10)	Teacher gives homework in math lesson	.51

Standardized α coefficient = .56

Total Variance Explained

FACTR 1	17.6
FACTR 2	13.1
Total	31.1

Descriptive Statistics

	Mean	SE	SD
FACTR 1	1.7	.01	.63
FACTR 2	3.5	.00	.61

Correlation Matrix

	Math Mean Plausible	FACTR 1	FACTR 2
Science Mean Plausible	1.0		
FACTR 1	-.37***	1.0	
FACTR 2	-.091***	-.21***	1.0

Table A: Fixed School Effectiveness Models in Mathematics (n = 203)

	a	b	c	d	e
	0 Null	1 Student	1 + Group Student-body	1+ Behave Grouping, Instruction	Full (Behave, Grouping, ACADP-P, BGEDSP-P)
Variance decomposition					
Between groups:	3493 (44%)	1733 (31%)	974 (20%)	1474 (28%)	889 (19%)
Within groups:	4400	3862	3860	3862	3859
Total	7893	5595	4834	5336	4748
Percentage of between group variance explained beyond student level model		50% a-b	44% b-c	15% b-d	49% b-e
Added contributions to the explained between group variance			22%	8%	25%
Between group explanatory power of the models		50%	72%	58%	75%
INTERCEPT		347 (7.9)***	234 (22.4)***	503 (56.8)***	294 (36.6)***
Behave				20.4 (9.4)*	12.2 (7.6)
Grouping				-3.5 (6.9)	-2.9 (6.0)
FACTR 1				-58.6 (11.9)***	-33.0 (9.8)**
FACTR 2				-19.9 (11.8)	
BOOK-P			22.4 (6.2)***		21.5 (6.1)**
BGEDSE-P			32.3 (15.3)		24.5 (14.6)***
ACADP-P			81.0 (15.5)***		74.4 (15.0)***
GOOD-P			-0.4 (14.1)		
No. of books at home (BOOK)		6.9 (1.0)***	6.2 (1.0)***	6.9 (1.0)***	6.2 (1.0)***
Parents academic ed. (ACADP)		6.7 (2.0)***	4.2 (1.9)*	6.6 (1.2)**	4.2 (1.9)*
Sex		-12.8 (2.3)***	-13.2 (2.3)***	-13.1 (2.3)***	-13.5 (2.3)***
Ethnic origin (ISRARB)		92.7 (7.8)***	72.1 (8.7)***	36.6 (16.8)*	39.1 (13.9)**
Self-esteem as pupil (GOOD-D)		34.0 (2.5)***	34.1 (2.5)***	34.0 (2.5)***	34.1 (2.4)***
Academic aspiration (BGEDSE-D)		21.9 (2.5)***	20.8 (2.4)***	21.7 (2.5)***	20.8 (2.4)***

Table B: Fixed School Effectiveness Models in Mathematics (n = 134)

	a	b	c	d	e
	0 Null	1 Student	1 + Group Student-body	1+ Behave Grouping, Instruction	Full (Behave, Grouping, School SES, ACADP-P, BOOK-P)
Variance decomposition					
Between schools:	3038 (36%)	1363 (23%)	813 (15%)	1116 (20%)	744 (14%)
Within schools:	5306	4579	4579	4579	4579
Total	8344	5942	5392	5695	5223
Percentage of between school variance explained beyond student level model		55% a-b	40% b-c	18% b-d	45% b-e
Added contributions to the explained between group variance			18%	8%	21%
Between schools explanatory power of the models		55%		63%	76%
INTERCEPT		334 (8.3)***	261 (34.4)***	335 (24.4)***	279 (26.4)***
Behave				16.8 (12.2)	
Grouping				-3.6 (6.8)	
School SES				-5.3 (1.3)***	-3.1 (1.1)**
BOOK-P			19.5 (8.8)*		15.5 (8.1)
BGEDSE-P			-4.3 (26.6)		
ACADP-P			100.6 (21.4)***		89.5 (21.8)***
GOOD-P			-23.6 (24.9)		
No. of books at home (BOOK)		8.2 (1.1)***	7.8 (1.1)***	8.2 (1.1)***	7.8 (1.1)***
Parents academic ed. (ACADP)		11.1 (2.2)***	9.3 (2.2)***	11.0 (2.2)***	9.4 (2.2)***
Sex		-11.0 (2.3)***	-11.3 (2.3)***	-11.1 (2.3)***	-11.3 (2.3)***
Ethnic origin (ISRARB)		88.8 (8.5)***	64.0 (10.1)***	61.6 19.3**	66.2 (9.2)***
Self esteem as pupil (GOOD)		33.9 (2.6)***	34.3 (2.6)***	33.8 (2.6)***	34.3 (2.6)***
Academic aspiration (BGEDSE)		26.8 (2.5)***	26.6 (2.4)***	26.7 (2.5)***	26.5 (2.4)***

Table C: Fixed School Effectiveness Models in Science (n = 199)

	a	b	c	d	e
	0 Null	1 Student	1 + Group Student-body	1+ Behave Grouping, Instruction	Full (Behave, Grouping, School SES, ACADP-P, BOOK-P)
Variance decomposition					
Between groups:	3253 (36%)	1650 (25%)	1122 (18%)	1392 (22%)	1006 (18%)
Within groups:	5814	4987	4987	4987	4987
Total	9067	6637	6109	6379	5993
Percentage of between group variance explained beyond student level model		49% a-b	32% b-c	16% b-d	39% b-e
Added contributions to the explained between group variance			16%	8%	20%
Between group explanatory power of the models		49%	65%	57%	69%
INTERCEPT		323 (10.3)***	250 (29.6)***	428 (44.9)***	324 (30.7)***
Behave				14.2 (11.6)	
Grouping				23.9 (6.4)***	17.0 (4.8)**
FACTR 1				-41.6 (10.8)***	-32.7 (9.4)**
FACTR 2				-13.2 (10.7)	
BOOK-P			14.7 (7.2)*		15.6 (7.0)*
BGEDSE-P			3.4 (20.2)		- -
ACADP-P			92.5 (16.3)***		78.9 (15.4)***
GOOD-P			0.5 (17.7)		- -
No. of books at home (BOOK)		10.9 (1.0)***	10.4 (1.1)***	10.9 (1.1)***	10.4 (1.1)***
Parents academic ed. (ACADP)		7.4 (2.4)***	4.8 (2.4)*	7.1 (2.4)**	4.8 (2.4)*
Sex		-9.5 (2.5)***	-9.6 (2.5)***	-9.8 (2.5)***	-9.8 (2.5)***
Ethnic origin (ISRARB)		100.8 (10.2)***	78.9 (11.2)***	45.9 (20.5)*	55.7 (11.3)***
Self esteem as pupil (GOOD)		36.1 (2.8)***	36.8 (2.8)***	36.1 (2.8)***	36.7 (2.8)***
Academic aspiration (BGEDSE)		27.4 (2.5)***	26.9 (2.5)***	27.3 (2.5)***	26.9 (2.5)***

Table D: Fixed School Effectiveness Models in Science (n = 134)

	a	b	c	d	e
	0 Null	1 Student	1 + Group Student-body	1+ Behave Grouping, Instruction	Full (Behave, Instruction, Organization, ACADP-P, BGEDSP-P)
Variance decomposition					
Between schools:	3691 38%	1774 26%	1086 18%	1367 21%	982 16%
Within schools:	5907	5047	5046	5046	5046
Total	9598	6821	6132	6413	6028
Percentage of between school variance explained beyond student level model		52% a-b	39% b-c	23% b-d	45% b-e
Added contributions to the explained between school variance			19%	11%	22%
Between group explanatory power of the models		52%	71%	63%	74%
INTERCEPT		319 (10.4)***	253 (40.8)***	324 (26.9)***	313 (12.9)***
Behave				16.5 (13.2)	
Grouping				20.2 (7.5)**	12.3 (8.9)*
School SES				-6.3 (1.4)***	-4.0 (1.3)**
BOOK-P			16.5 (10.4)		
BGEDSE-P			-8.7 (29.8)		
ACADP-P			121.7 (23.1)***		116.8 (17.7)***
GOOD-P			-20.6 (23.7)		
No. of books at home (BOOK)		10.3 (1.1)***	9.8 (1.1)**	10.3 (1.1)***	10.1 (1.1)***
Parents academic ed. (ACADP)		8.3 (2.5)**	6.5 (2.5)*	8.1 (2.5)**	6.5 (2.5)*
Sex		-10.0 (2.5)***	-10.1 (2.5)**	-10.1 (2.5)***	-10.2 (2.5)***
Ethnic origin (ISRARB)		101.7 (10.5)***	70.4 (12.4)***	65.0 (21.0)**	68.9 (10.2)***
Self esteem as pupil (GOOD)		36.4 (2.8)***	37.1 (2.8)***	36.2 (2.8)***	36.6 (2.8)***
Academic aspiration (BGEDSE)		28.4 (2.6)***	28.2 (2.6)***	28.3 (2.6)***	27.9 (2.6)***

Table E: Mathematics Random and Interactive Group Models (n = 203)

	a	b	C	d	e	f
	0 Null	**1 Student**	**1 + Group Student-body**	**1+ Behave Grouping, Instruction (FACTR)**	**1 + Full Model**	**Interactive Model**
Variance decomposition						
Between groups:	3493 (44%)	1607 (32%)	875 (22%)	1206 (26%)	790 (20%)	801 (20%)
Between slopes	-	209	197	124	142	142
Within groups:	4400	3811	3809	3813	3810	3810
Total	7893	5627	4881	5143	4742	4752
Percentage of between group variance explained		48%	41%	27%	49%	50%
		a-b	b-c	b-d	b-e	b-f
Additional explained variance beyond student model level			21%	14%	25%	26%
Between school explanatory power of the models		48%	61%	75%	73%	74%
INTERCEPT		350 (7.8)***	237 (23.0)***	346 (18.2)***	360 (54.9)***	317(12.19)*
Behave				-10.8 (9.2)	-	-
Grouping				-0.1 (6.8)	-	-
FACTR 1				-39.5 (7.5)***	-35.9 (10.7)**	52.3 (12.3)*
FACTR 2				-27.7 (7.2)***	-6.2 (9.9)	-46.8 (20.3)*
Grouping / FACTR 1						13.2 (20.9)
Grouping / FACTR 2						21.9 (28.9)
BOOK-P			24.9 (6.4)***		22.0 (6.3)**	21.8 (6.2)**
BGEDSE-P			23.1 (15.8)			
ACADP-P			76.1 (16.1)***		84.6 (15.5)***	84.5 (15.5)**
GOOD-P			-1.4 (14.2)			
No. of books at home (BOOK)		7.0 (1.0)***	6.4 (1.0)***	7.1 (1.1)***	6.5 (1.0)***	6.5 (1.0)***
Parents academic ed. (ACADP)		6.9 (2.0)**	3.9 (1.9)*	6.2 (2.0)**	3.9 (1.9)*	3.9 (2.0)*
Sex		-12.8 (2.3)***	-13.1 (2.3)***	13.2 (2.3)***	-13.2 (2.3)***	-13.2 (2.4)
Ethnic origin (ISRARB)		89.1 (7.8)***	69.1 (8.6)***	55.3 (16.1)**	45.0 (8.9)***	41.9 (10.1)
Self-esteem as pupil (GOOD)		33.3 (2.4)***	33.3 (19.9)	7.5 (11.9)	5.1 (20.4)	4.9 (20.4)
Academic aspiration (BGEDSE)		21.5 (2.4)***	20.6 (2.4)***	21.3 (2.4)**	21.1 (2.4)***	21.2 (2.4)
For good slope:						
			GOODP 2.9 (15.9) ACADP 17.3 (12.3) BGEDSEP19.9 (14.9) BGBOOK -6.7 (6.2)	BEHAV 11.1 (3.8)** FACTR1 –7.7 (4.7) FACTR2 –0..5 (5.5) GROUPING –9.4 (5.)	BEHAV 12.3 (3.5)** ACAD 13.6 (12.9) BGBOOK -2.6 (5.9)	BEHAV 12.4 (3.5)** ACADP 3.6 (12.9) BOOK -2.7 (5.9)

Table F: Science Random and Interactive Group Models (*n* = 199)

	a	b	C	d	e	f
	0 **Null**	**1** **Student**	**1 + Group** **Student-body**	**1+ Behave** **Grouping, Instruction** **(FACTR)**	**Full Model**	**Interactive Model**
Variance decomposition Between groups: Between slopes Within groups: Total	 3253 (36%) 5814 9067	 1643 (29%) 320 4917 6880	 1059 (21%) 247 4921 6227	 1390 (24%) 136 4913 6439	 968 (18%) 123 4912 6005	 942 (17%) 89 4920 5951
Percentage of between group variance explained Additional explained variance beyond student model level Between group explanatory power of the model		40% a-b 40%	33% b-c 20% 60%	22% b-d 13% 53%	44% b-e 26% 66%	47% b-f 28% 68%
INTERCEPT Behave Grouping FACTR 1 Instruction FACTR 2 Instruction		325 (10.2)***	238 (29.5)	470 (46.6)*** 11.1 (11.6) 21.9 (6.8)** -37.1 (11.0)** -24.3 (11.4)*	337 (44.9)*** - 15.3 (4.9)** -31.9 (9.5)** -5.1 (10.2)	325 (48.9)*** - 47.3 (59.3) -59.5 (115.0)*** 18.5 (13.6)
Grouping / FACTR 1 Grouping / FACTR 2						45.8 (17.1)** -33.8 (18.1)
BOOK-P BGEDSE-P ACADP-P GOOD-P			20.8 (7.1)** 5.6 (20.7) 79.7 (17.3)*** 2.6 (18.9)		19.7 (7.1)** - 73.9 (16.5)*** -	18.6 (6.8) - 78.0 (15.9) -
No. of books at home (BOOK) Parents academic ed. (ACADP) Sex Ethnic origin (ISRARB) Self-esteem as pupil (GOOD) Academic aspiration (BGEDSE)		10.9 (1.1)*** 7.4 (2.5)*** -9.2 (2.5)*** 97.6 (10.1)*** 36.9 (2.8)*** 27.4 (2.5)***	10.4 (1.1)*** 4.7 (2.5) -9.2 (2.5)*** 73.3 (11.1)*** 91.4 (24.9)*** 26.8 (2.5)***	11.1 (1.1)*** 6.8 (2.5)** -9.6 (2.5)*** 40.3 (20.1)* -77.4 (33.4)* 27.2 (2.5)***	10.45 (1.1)*** 4.5 (2.5) -9.5 (2.5)*** 43.5 (11.7)*** -68.5 (40.8) 26.8 (2.5)***	10.5 (1.1)*** 4.5 (2.8) -9.4 (2.5)*** 28.0 (12.4)* -68.9 (4.0.4) 26.7 (2.5)***
For good slope:						
			GOODP -17.7 (18.90) BGEDSEP -4.2 (21.3) ACADP 41.4 (16.3)* BGBOOK –18.2 (7.1)*	BEHAV 11.0 (4.2)** FACTR1 –12.8 (9.5) FACTR2 31.9 (9.9)** GROUPING 4.4 (5.4)	BEHAV 14.7 (4.2)** FACTR2 29.6 (9.6)** ACADP 22.0 (14.7) BGBOOK –12.2 (6.7)	BEHAV15.2 (4.2)*** FAC2 29.4 (9.6)** ACADP21.4 (14.5) BOOK -12.3 (6.6)

Table G: Mathematics Random School Model (n = 134)

	a	b	c	d	e
	0 Null	**1 Student**	**1 + Group Student-body**	**1+ Behave Grouping, School SES**	**Full (Grouping, School SES, Behave, ACADP-P)**
Variance decomposition Between groups: Between slopes Within groups: Total	 3038 (36%) 5306 8344	 1239 (25%) 216 4526 5981	 687 (17%) 209 4524 5420	 1055 (20%) 142 4526 5723	 647 (15%) 152 4526 5325
Percentage of between group variance explained		52% a-b	45% b-c	15% b-d	48% b-e
Additional explained variance beyond student model level			22%	7%	23%
Between school explanatory power of the model		52%	74%	59%	75%
INTERCEPT		337 (8.2)***	255 (33.9)***	350 (24.2)***	280 (26.6)***
Behave Grouping School SES				10.5 (11.9) -0.1 (6.9) 4.9 (1.3)***	 -2.9 (1.1)**
Grouping FACTR 1 FACTR 2					
BOOK-P BGEDSE-P ACADP-P GOOD-P					
No. of books at home (BOOK)		8.3 (1.1)***	7.9 (1.1)***	8.4 (1.1)***	8.0 (1.1)***
Parents academic ed. (ACADP)		10.8 (2.2)***	8.9 (2.2)	10.7 (2.2)***	9.1 (2.2)***
Sex		-11.0 (2.3)***	-11.2 (2.3)***	-10.9 (2.3)***	-11.2 (2.3)***
Ethnic origin (ISRARB)		84.8 (8.4)***	60.7 (9.8)***	61.5 (18.7)**	58.0 (9.0)***
Self-esteem as pupil (GOOD)		33.7 (2.6)***	59.0 (26.4)*	3.8 (13.3)	-4.8 (10.0
Academic aspiration (BGEDSE)		26.5 (2.5)***	26.3 (2.4)***	26.2 (2.4)***	26.0 (2.4)***
			MGOOD1 7.1 (22.6) BGEDSEP 11.9 (20.8) ACAD-P 34.0 (14.1)* BGBOOK -13.6 (7.8)	BEHAV 13.2 (2.8)** GROUPING -6.7 (5.0) ASIRON -0.9 (0.9)	BEHAV 12.6 (3.9)** ACADP 3.4 (12.2)

Table H: Science Random School Model (n = 134)

	a	b	c	d	e
	0 Null	**1 Student**	**1 + Group Student-body**	**1+ Climate Grouping, School SES**	**Full Grouping (Behave, School SES, ACADP-P, BGBOOK)**
Variance decomposition Between groups: Between slopes Within groups: Total	 3691 (38%) 5907 9598	 1775 (29%) 283 4085 7043	 966 (20%) 217 4986 6199	 1415 (24%) 139 4984 6530	 897 (17%) 104 4987 5980
Percentage of between group variance explained Additional explained variance beyond student model level Between school explanatory power of the model		44% a-b 44%	41% b-c 23% 67%	24% b-d 13% 57%	51% b-e 29% 73%
INTERCEPT Behave Grouping School SES		322 (10.3)**	241 (39.6)***	345 (27.6)*** 10.3 (13.4) 17.2 (7.9)* -6.1 (1.4)***	266 (31.5)*** - 10.8 (5.7)* -3.6 (1.1)**
BOOK-P BGEDSE-P ACADP-P GOOD-P			23.3 (9.9)** -8.1 (28.2) 114.3 (23.6)*** -18.3 (23.8)		- - 99.8 (20.6)***
No. of books at home (BOOK) Parents academic ed. (ACADP) Sex Ethnic origin (ISRARB) Self-esteem as pupil (GOOD) Academic aspiration (BGEDSE)		10.3 (1.1)*** 8.3 (2.5)** -9.7 (2.5)*** 99.0 (10.5)*** 36.2 (2.8)*** 28.4 (2.6)***	9.9 (1.1)** 6.3 (2.5)* -9.8 (2.5)*** 60.9 (12.2)*** 108.5 (29.8)** 28.0 (2.6)**	10.3 (1.1)*** 7.7 (2.5)** -9.7 (2.5)*** 62.2 (2.5)*** -8.8 (4.4) 28.3 (2.5)***	9.9 (1.1)*** 6.2 (2.5)** -9.9 (2.5)*** 57.7 (11.2)*** 33.9 (30.0) 28.0 (2.5)***
For good slope:			GOOD-P -23.4 (21.6) BGEDSEP -6.7 (25.6) ACAD-P 39.9 (18.0)* BOOK -22.1 (8.3)***	BEHAV 15.2 (4.4)** GROUPING 7.9 (5.5) ASIRON -0.5 (0.9)	BEHAV 17.8 (9.5)*** ACADP 7.5 (16.2) BOOK-P-15.6 (8.1)*