The Effects of Grouping and Mastery Learning Method of Instruction on English Achievement Levels of Lycée One Students in Kuleli Military High School*

Güzver Yıldıran and İlker Tuğal

Abstract

The aim of the study is to assess the effects of instruction (Mastery Learning versus conventional learning) and type of grouping (heterogeneous, homogeneous, and non-grouping of low, average, and high-achieving students) on English achievement levels of lycée one students at Kuleli Military High School in İstanbul. The results show that Mastery Learning and type of grouping had significantly positive and additive effects on summative achievement, the effect of Mastery Learning being 18.8 times greater than type of grouping. Heterogeneous grouping was significantly more effective in comparison to non-grouping under both Mastery Learning and conventional instruction, and with the exception of low-achieving students in conventional classes, it was more conducive to achievement as a trend under both instructional methodologies.

Keywords: Mastery learning, grouping, English achievement level

Statement of the Problem

Because the effects of type of grouping based on a global and rather stable characteristic like I.Q. does not lead to conclusive results, it was thought that if grouping (heterogeneous or homogeneous) was based on more task related or task specific factors rather than global characteristics, its impact could be more visible. Thus, research was carried out in a military academy in Istanbul, Turkey.

The research aims to respond to two questions. The first question is related to the effect of homogeneous or heterogeneous groupings of students (based on a standardized achievement test) on their subsequent achievement levels in the same area. The second major question aims to investigate whether grouping has an additive effect to Mastery Learning in explaining achievement levels of students.

This project was the Master's thesis of my graduate student, now Colonel İlker Tuğal, carried out in Kuleli Military High School in Istanbul, which he completed under my supervision in the Department of Educational Sciences of Boğaziçi University in 1995. The thesis entitled The Effects of Grouping and Mastery Learning Method of Instruction on English Achievement Levels of Lycée One Students in Kuleli Military High School shared the "Best Thesis in Social Sciences**" award in 1995, given yearly to a thesis of exceptional quality in social sciences as well as to another one in natural or physical sciences by the Boğaziçi University Research Fund. The award in social sciences was shared by my other graduate student Emin Aydın the same year.*

Güzver Yıldıran, Prof., Boğaziçi University, Faculty of Education, Department of Educational Sciences, Istanbul, yildiran@boun.edu.tr

İlker Tuğal, Retired Colonel, Head of the Department of Foreign Languages, Bursa Işıklar Air Force High School

One can possibly posit that ability grouping began to sweep across different cultural settings, starting from the time of Binet and Simon (1905a, b). They were commissioned by the French Ministry of Public Instruction in 1904 to develop a valid andreliable instrument which would objectively separate children who could profit from public instruction given in Paris, from those who could not. Through the years that followed, the interest in separating students into various different categories became more and more refined. Slavin defined ability grouping as "some means of grouping students for instruction by ability or achievement so as to reduce their heterogeneity", (Slavin,1987a, p. 294). Among various kinds of ability groupings, ability grouped class assignment (called "tracking" by Americans and "streaming" by the British), ability groupings for selected subjects, the Joplin plan (regardless of their grade level, students meet in a homogeneous class for reading, based on their achievement levels), non-graded plans, special classes for high achievers (the gifted), special classes for low achievers (slow learners), and within class ability groupings can be cited (Slavin, 1987a).

Two opposing views have emerged regarding ability grouping as its merits began to be questioned with its utilization over time and over setting. These views can be summarized by what Kerckhoff (1986) calls the ***traditional*** and the ***divergence hypotheses.*** The traditional hypothesis states that ability grouping yields positive results for all students, while the divergence hypothesis states that there are positive gains for students in high ability groups, but students in low ability groups actually lose in performance.

Slavin (1990) lists the advantages and disadvantages of ability grouping in secondary schools in his ***best-evidence synthesis.*** Slavin states the advantages as, "It permits pupils to make progress commensurate with their abilities: it makes possible an adaptation of the technique of instruction to the needs of the group; it reduces failures; it helps maintain interest and incentive, since bright students are not bored by the participation of the dull; slower pupils participate more when not eclipsed by those much brighter; it makes teaching easier; it makes possible individual instruction to small groups", (1990, p. 473).

One of the points made here, i.e. that bright students would be bored by the participation of the dull, seem to be a construed reality which may not be all that congruent with the experiences of students. First, what Slavin calls the "dull students" do not participate often. Tautologically, because these students do not participate, they do not learn as much; and since they don't know, they cannot participate. If this state of affairs defines a "dull student", with no reference to unalterable characteristics, then certainly environmental remediation is possible. Secondly, if these students do in fact participate, why should any other group of students be bored with it? Au contraire, peer group interaction and remediation seems to work as much for the "bright students" as for others as decades of Mastery Learning (Hackenberg, 1993; Yıldıran and Hackenberg, 1996) as well as other research have shown (De Weerdt, 1996a, 1996b, 1998).

The empirical support for the advantages of ability grouping comes from Newfield and McElyea (1983) in their "High School and Beyond" study, where the researchers went through data on 58,000 high school seniors and sophomores. Their conclusion was that ability grouping leads to improved achievement and attitude toward

subject areas, both in regular and remedial classes. Kerckhoff (1986) in his review of British schools states that teachers in Britain favor the use of ability grouping despite the lack of empirical support for it, or they object to streaming on philosophical grounds but actually practice it.

Slavin (1990, p. 473) counts the disadvantages of ability grouping as slow students needing the presence of abler ones for stimulation and encouragement; stigmatization of slow groups, de-motivating the students in these groups; teachers not having the skills or the time to differentially assign work for various levels of ability; and teachers objecting to working with slower groups. Many of the above stated disadvantages have been validated by other research. The point that students in high ability groups show gains while those in low ability groups lose in academic performance is substantiated by research (Acland, 1973; Kelly, 1974; Hallinan and Sørensen 1983; Kerckhoff, 1986). Acland (1973) showed that while junior high school students in the upper stream gained 0.71 test points between the ages of 8-11, students in the lower stream lost an average of 0.49 points. Hallinan and Sørensen, (1983) found that students placed in high ability groups benefited more from practice. Kerckhoff (1986) followed a sample of 4,797 boys and 4,602 girls through their education until before graduation from secondary school. The cohorts attended one of four types of schools: grammar schools for high ability students preparing them for university education, secondary modern schools, comprehensive schools, and private secondary schools. In the schools that streamed their students into high, middle, and low ability groups (secondary modern and comprehensive schools), the divergence pattern of increased gains for high ability groups and decreased gains for low ability groups was found consistently in terms of math and English achievement levels.

One needs to be reminded by the work of Rosenthall and Jacobson (1966) in terms of how teachers are affected by their perceptions of students and end up structuring differential environments for them, which lead to differential learning outcomes. Students, as Jackson (1968) states, are put into very different environments in the same classroom, with very differential advantages and disadvantages. Reuman (1989) cites studies where assignment to low-ability classrooms or groups created low expectations for students, which became self-fulfilling prophecies. Students then developed negative affect for schools due to low academic self concept. The research of Yıldıran and Nwabueze (1991) illustrates how teachers interact most with students who need them the least, i.e. with high and average achievers, and interact minimally with low achievers. Kifer (1973) has shown in his quasi-longitudinal study that while in second grade, high and low achievers are similar to each other in terms of self-concept; by the time they are in fourth grade, they are very different; and when they are in sixth grade, they are two different populations, the differences enlarging even more in the eighth grade. Hallinan and Sørensen (1983) report the negative effects on self-esteem as well as the disturbance of parents when students are moved to a lower ability group. Students who fall behind their group tended to lose motivation and learned less than the case would have been if they were not grouped. Kelly (1974) states much more strongly that ability grouping leads to delinquent behavior for students in low ability groups.

Aside from the ethical difficulties related to the research on ability grouping, Acland (1973), and Hallinan and Sørensen a decade later (1983), complained about the methodological difficulties in research on ability grouping. According to Acland, ability

grouping was not usually done on acceptable criteria for separation but on the basis of either the social background of the student or teacher ratings of ability. Hallinan and Sørensen stated that students were often assigned to three stable groups of high, medium, and low in equal numbers rather than on the basis of actual homogeneity.

If research results have shown one of these disadvantages to be the case (and almost all of them have been validated by findings), how is it possible that we are still interested in ability grouping, unless fundamentally we are interested in and intend to teach some well and others not? If an intervention in medicine had equal probability of being as hazardous as the above stated situation, what would be its chances of ever being practiced? And what would happen to those who practice it in the face of evidence?

Unless groupings are formed holding all of the above stated motivational variables as well as expectations constant, this intervention will always have dubious outcomes. If groupings are formed with the aim of making all students reach the criterion level of learning, then it will not be important what kind of grouping is formed, provided that teachers are equally motivated to enable all students attain desirable levels of learning and expect this to happen to all of them. In that case only, will all students learn equally well, regardless of whether students in these classes are homogeneously or heterogeneously grouped. The comment made is similar to Slavin's conclusion in his ***best-evidence synthesis*** in 1990, that in secondary schools none of the between- class ability grouping plans have had a significant effect on achievement. His analysis leads him to the conclusion that ability grouping has no positive or negative effects on students of high, average, or low ability. This conclusion is not only due to methodological errors in ability grouping strategies and research, but also to the effects of uncontrolled intervening variables under ability grouping procedures that influence achievement or learning gains. Thus, it is not grouping itself, but what is done within those groups and what is expected from them that affect achievement levels or learning gains of students.

Within the above stated conceptual framework and concerns, the following hypotheses were developed:

I. The effects of Mastery Learning and grouping on achievement will be significant.

II. The achievement level of all three classes under Mastery Learning will be significantly higher than their conventionally taught counterparts. Thus, the Mastery Learning class with heterogeneous grouping will score significantly higher than the control class with the same type of grouping; the Mastery Learning class with homogeneous grouping will score higher than its control counterpart; finally the Mastery Learning class without grouping will score higher than the control condition without grouping.

III. While there will not be a significant difference in achievement levels between heterogeneously and homogeneously grouped Mastery Learning classes, both of these classes will achieve significantly higher than the Mastery Learning class without grouping.

IV. The class taught under conventional learning with heterogeneous grouping will achieve significantly higher than the class under conventional learning with homogeneous grouping. Both of these two classes will attain significantly higher achievement levels than the class under conventional learning without grouping.

V. Grouping will differentially affect low, average, and high achieving students under
different instructional methods.

Methodology

Subjects of the Study

The study was carried out at Kuleli Military High School in Istanbul, Turkey. The students are highly selective due to admittance procedures. Students are selected according to the results on an examination in junior high school math, science, and Turkish language and literature. In addition, there is a physical education test, in addition to requirements of height and weight and lack of any type of illness and impairment. In 1995, the top 1.6 % was admitted from a sample of 52,000 applicants.

After admittance into this military academy after junior high school, students study English intensively for a preparatory year before lycée one. The students in this study were in lycée one, after having completed their preparatory year. Thus, they were the same age as their 10th grade American cohort. Kuleli students come from middle and lower-middle SES backgrounds and are in the high ability group, being homogeneous in terms of the characteristics they share with other students in the same school.

There were 14 sections of lycée one comprised of 378 students, one English teacher instructing every two sections. Six sections with a total of 143 students were randomly selected from among 14 sections. In the final sample, there were 136 students in the six sections, seven students having dropped out of the study due to illness. Table 1 shows the student numbers of the classes and their attrition rates.

Table 1. Distribution of subjects into six classes in the study

Classes	Section No.	Beginning Number	Attrition	Final Number
CL+HT	2	25	2	23
CL+HM	4	23	3	20
CL+NG	6	24	1	23
ML+HT	7	23	1	22
ML+HM	8	24	-	24
ML+NG	13	24	-	24
Total		143	7	136

All sections were taught by different teachers except for the Mastery Learning with heterogeneous grouping (ML+HT) and Mastery Learning with homogeneous grouping (ML+HM), which were taught by the same teacher. All five teachers took part in the study on a voluntary basis.

Subject Area

The subject area was lycée one English, using the textbook *Destinations***,** which is the third book of the *American Streamline Series* (Hartley and Viney, 1985). There were 80 units in the book. The first two books review grammar. There were four learning tasks in the study. The first learning task included units 57 and 58, which were on reduced relative clauses, and clauses of reason (introduced by 'because'), as well as clauses of contrast (introduced by 'although', 'though', 'even though'). Learning task 2 was comprised of units 59, 60, and 61 on sentence connectors ('not only...but also'), clauses of reason (introduced by 'because of' and 'due to'), clauses of contrast (introduced by 'in spite of' and 'despite'), and sentence connectors which express contrast ('however' and 'nevertheless'). Units 62 and 63 comprised the third learning task on noun clauses beginning with 'what' in the subject and object positions, and the usage of '-ever'. Finally, learning task four included units 65, 66, and 67 dealing with 'would rather', time expressions 'It's time', and 'as if', 'as though'. The first and second learning tasks included 15 objectives each, the third learning task had 13, while the fourth had 10 objectives on lower as well as higher mental processes, altogether totaling to 53 objectives. Instruction time for each learning task was eight 40 minute class hours a week, the study lasting for four weeks. Table 2 shows the distribution of units into learning tasks, their content, number of objectives included in each learning task, and the number of class hours for each learning task.

The Design of the Study

The study was a field experiment with a three by two design comprised of six conditions, three under Mastery Learning and three under conventional instruction. There was one Mastery Learning class with heterogeneous student grouping (ML+HT), another with homogeneous student grouping (ML+HM), while in the third Mastery Learning class, the students were not grouped at all (ML+NG). Each of these three Mastery Learning classes had a counterpart with the same type of grouping under conventional instruction, (the CL+HT, CL+HM and CL+NG classes). Table 3 shows the design of the study.

Table 2. Number and content of the units, number of objectives and instruction time in each task of the study

Task No.	Units	Content of the Units	Number of Objectives	Hours of instruction
1	57, 58	Reduced relative clauses; because, although, though, even though.	15	8' of 40"
2	59, 60, 61	Not only ... but also; because of, due to; in spite of, despite; however, nevertheless.	15	8' of 40"
3	62, 63	Noun clauses with "what" as subject and/or object; "-ever" words.	13	8' of 40"
4	65, 66, 67	Would rather; it's (about) time; as if/as though.	10	8' of 40"
Total	10	-	53	32' of 40"

Table 3. The design of the study

	Mastery Learning (ML)	**Conventional Learning (CL)**
Heterogeneous Grouping (HT)	ML+HT	CL+HT
Homogeneous Grouping (HM)	ML+HM	CL+HM
	ML+NG	CL+NG

Subjects were assigned to the six classes randomly, the assignment of the treatments to the classes being also random. The assignment of groupings in each of the four classes (ML+HT, ML+HM, CL+HT, and CL+HM) was done by matching the students on the basis of their English Comprehension Level (ECL) test scores. The ECL test is developed by the American Defense Institute to evaluate the comprehension level in English for officers from allied countries who are nominees for various courses in the

U.S.A. It is a standardized test of 120 items with 40 alternate forms, 20 of which are renewed annually. The ECL questions are aimed to test listening and reading comprehension abilities, vocabulary power, and knowledge of basic grammar. The test is administered in groups in 45 minutes. Kuleli Military Academy students take this test five times through their education; once in entering the school, and once at the end of each academic year in order to follow their improvement in English. The scores on the ECL were obtained before this study, after the students had completed their preparatory year at the end of the 1993-1994 academic year. The ECL scores served as the criterion for forming homogeneous and heterogeneous groups.

Four to six students in each class with the highest scores on the ECL comprised the high achieving group. Similarly, four to six students in each section with the lowest scores comprised the low-achieving group. Students whose scores ranged between these two groups were identified as the average achieving students.

In the classes with heterogeneous student groupings (ML+HT and CL+HT), six groups were formed with four students on the average in each group. In each of these four groups, there was one high, one low, and two average achieving students. In the two classes with homogeneous groupings (ML+HM and CL+HM), four students with the highest scores on the ECL formed the high achieving group, while four students with the lowest ECL scores formed the low achieving group. The remaining closest scoring students were placed in groups of four, comprising the average achieving students.Table 4 shows the distribution of high, average and low achieving students in each class.

Table 4. Number of low, average, and high achieving students in ML+HT, ML+HM, CL+HT and CL+HM classes

Classes	Number of	No of Low	No of Average	No of High
ML+HT	6	5	11	6
ML+HM	6	4	16	4
CL+ HT	6	6	12	5
CL+HM	6	3	13	4
Total	24	18	52	19

As can be observed from the table, all classes are normally distributed in terms of the number of students in the low, average and high achieving clusters due to the procedures followed in determining the low and high achieving students. In the two Mastery Learning classes (ML+HT and ML+HM), six homogeneously or heterogeneously formed groups (four students in each group) came together for one class period for corrective learning after completing each learning task and taking the formative test related to that learning task. The aim of group work was correction of mistakes on the formative test for the learning task proper. The teacher monitored group work without participating in it, unless a question was directly asked to the teacher. In that condition, the teacher made the explanation to the whole class. After this group

work, students in the ML+HT and ML+HM classes, who could not reach the 85 % criterion level, took the parallel form of the formative test for that particular learning task. No student in these two Mastery Learning classes needed another corrective loop to reach the 85 % criterion level of learning on a particular learning task. Consequently all students in these two Mastery Learning classes went on to the next learning task after the parallel form of the formative test. In the control classes with homogeneous or heterogeneous groupings, students in the respective groups came together at the end of each learning task, again for one class period, for any purpose their teachers considered appropriate. After group work, students went on to the next learning task without taking any formative tests. In the two other classes (ML+NG and CL+NG) no groupings were formed. In the Mastery Learning class without grouping, students took the formative test after the completion of a learning task. Correctives were given to students on common errors for each learning task. The parallel form of the formative test was given to students who had not reached the 85 % criterion level. No student required another corrective loop for the same learning task in this Mastery Learning class. Consequently, the class went on to the next learning task. In the control class without grouping, after the completion of a learning task, the students went directly on to the next learning task.

Type of grouping (heterogeneous, homogenous, or no grouping), type of instruction (Mastery Learning, or conventional) and ECL performance levels of students (low, average, and high) constituted the independent variables of the study. The dependent variable of the study was the summative test scores of students based on the objectives for all four learning tasks.

Training of the Teachers

The teachers of the Mastery Learning classes were instructed for six hours on two consecutive days on the theory and implementation of Mastery Learning. On the first day, a one hour lecture was given to the teachers of the Mastery Learning classes on the theory. This was followed by a session on how to write learning objectives for a learning task. A one hour demonstration of how to develop questions from these objectives ended the sessions on the first day of instruction. The activities of the second day involved going over the objectives prepared for the teachers (not the objectives the teachers themselves prepared) on the learning tasks included in the study. During the last hour on the second day, the teachers of the Mastery Learning classes practiced in preparing corrective materials.

Another kind of training was given to the teachers of the Mastery Learning and conventional classes in which homogeneous and heterogeneous groupings were to be formed. The Mastery Learning teachers who would be using different types of groupings were instructed on when and how group work would be done, and were given the names of students in each group (of four students in a group), in their classes. This session was followed by a separate session for the teachers of conventional instruction who would be using groupings in their classes.These teachers were also given information on groupings. To avoid bias, the teachers were not given information about which groups were made of low, average, or high achieving students. No teacher in the study saw the formative or the summative instruments before they were administered to students.

Data Collection Procedures

Initial Measures

There were two initial measures obtained for all students. The first measure was the students' English grades of the previous year used as an index of cognitive entry behaviors related to the subject area of the present study. The second measure was the ECL score of each student, administered at the end of the previous academic year. The ECL scores were used to match students on their English performance levels in assigning the homogeneous or heterogeneous groups.

Process Measures

A formative test was given to all students in the three Mastery Learning classes (ML+HT, ML+HM and ML+NG) after the completion of each of the four learning tasks. For students who had not reached the criterion level of learning of 85 % in these classes, correctives were assigned. A parallel form of the formative test for a particular learning task was given to students after correctives. As no student needed more help after the parallel forms of the formative tests in the Mastery Learning classes, a second repetition was not necessary. The students in the control classes did not receive the formative tests.

Final Measures

On the final day of the study, a summative test was administered to all students at the same time. The summative questions were derived from objectives for each learning task. Ten questions tapping five of the 15 objectives of the first learning task, 14 questions tapping seven of the 15 objectives of the second learning task, eight questions tapping four of the 13 objectives of the third learning task, and eight questions tapping four of the 10 objectives of the fourth learning task comprised the summative test of 40 items.

Hypotheses and Results

Prior to the study, all six classes were compared with each other in terms of their previous year English grades from the 1993-1994 academic year, and their ECL scores, to find out if there were any significant differences among these classes prior to the implementation of treatments. Tables 5 and 6 show the one-way analyses of variance done on each of these measures.

Table 5. One-way analysis of variance on the previous year's (1993-1994) English grades in the six classes

Source	df	Sum of Squares	Mean Squares	F	Significance Level
Between Groups	5	16.5284	3.3057	2.0975	.0697
Within Groups	130	204.8761	1.5760		N.S.
Total	135	221.4044			

Table 6. One-way analysis of variance on the previous year's (1993-1994) ECL scores in the six classes

Source	df	Sum of Squares	Mean Squares	F	Significance Level
Between Groups	5	173.2564	34.6513	.6125	.6904
Within Groups	130	7354.1259	56.5702		N.S.
Total	135	7527.3824			

Both tables indicate that the classes are not different from one another either in terms of their previous year English grades or their ECL scores prior to the implementation of the treatments.

Analysis Done on Each Hypothesis

The first hypothesis of the study states that the effects of Mastery Learning and grouping will be significant on achievement levels of students as measured by the summative test. Table 7 shows the descriptive statistics of this test.

All the trends expected in the four hypotheses of the study are illustrated in Table 7. Students under Mastery Learning methodology scored higher than students under conventional instructional methods; students who were in heterogeneous and homogeneous groups scored higher than students who were not put into groups; there is a small difference between students in heterogeneous groups in comparison to homogeneous groups favoring the former; and the alignment of the four classes in terms of summative achievement is in the expected direction.

Table 7. The descriptive statistics of the summative achievement test

Groups	N	Possible	Mean	Standard	Min.	Max.
ML_{Total}	70	40	33.98	3.80	25	40
CL_{Total}	66	40	24.39	5.44	12	35
$Grouped_{Total}$	89	40	30.44	6.12	14	39
NG_{Total}	47	40	27.23	7.20	12	40
ML Grouped	46	40	34.74	3.33	25	39
$_{Total}$	43	40	25.84	4.97	14	35
CL Grouped $_{Total}$						
HT Groups $_{Total}$	45	40	30.76	6.20	17	39
HM Groups $_{Total}$	44	40	30.11	6.02	14	39
ML+HT	22	40	35.36	3.08	28	39
ML+HM	24	40	34.17	3.44	25	39
ML+NG	24	40	32.54	4.22	25	40
CL+HT	23	40	26.35	5.14	17	35
CL+HM	20	40	25.25	4.71	14	35
CL+NG	23	40	21.69	5.23	12	30
Total	133	40	29.33	6.72	12	40

Figure 1 shows the percentages of students reaching the criterion level of learning of 85 % on the summative test in each of the six classes.

The figure indicates that in all Mastery Learning classes, the percentage of students reaching the 85 % level of learning far exceeds the conventional classes. The effect of grouping in enabling students to reach the criterion level of learning seems to be stronger under Mastery Learning conditions in comparison to control conditions, since the numbers reaching the criterion level in the three Mastery Learning classes show wider differences than conventional learning classes. In the heterogeneously grouped Mastery Learning class, 17/22 (77.3 %) students reached the criterion level, 14/24 (58.3 %) reached the same level in the homogeneously grouped, and 9/24 (37.5 %) in the non- grouped Mastery Learning classes. Only 2/23 (8.7 %) students in the control class with heterogeneous grouping reached the criterion level of learning, 1/20 (5 %) in the control class with homogeneous grouping, and no student reached this level in the conventional class with no grouping. Although the effect of grouping is minimal under control conditions, still the trend is aligned with the divergence hypothesis of Kerckhoff.

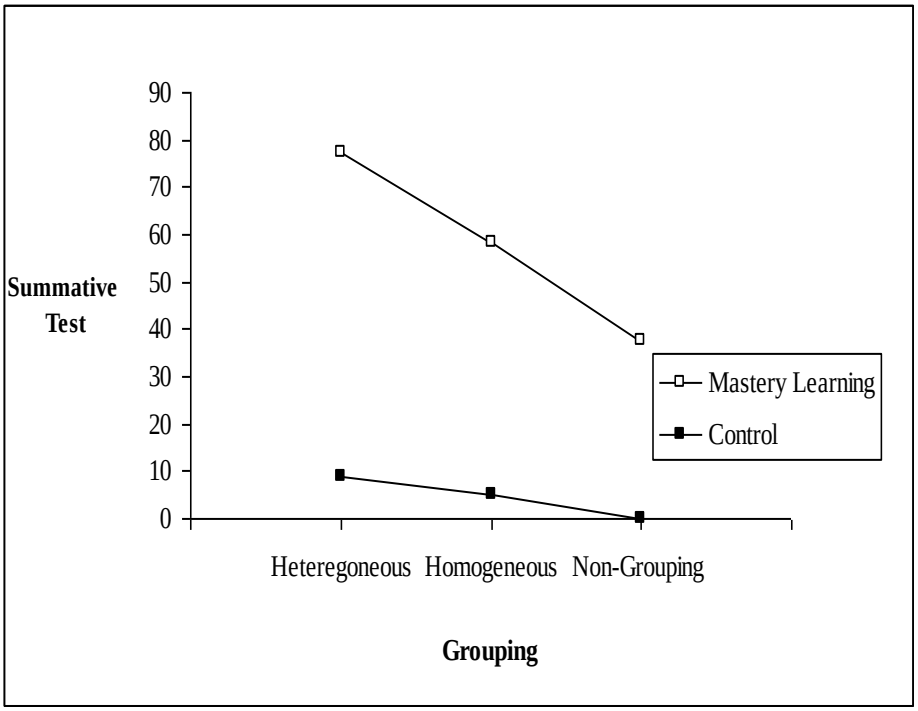

Figure 1. Comparison of the percentages of students reaching mastery on the summative test in six classes

The effects of instruction type and type of grouping were analyzed by a two-analysis of variance. Table 8 indicates that both type of instruction and type of grouping have significant effects on achievement (F=155.81, p=.000; and F=8.290, p=.005 respectively). The effect of instruction, however, is 18.795 times greater (MS_{INS}/MS_{GR}=18.795) than grouping. The interaction between the two interventions is not significant (F=0.666).

Table 8. Two-way analysis of variance on the summative test scores showing the effect of type of instruction and type of grouping

Source of variation	Sum of Squares	df	Mean Square	F	Significance Level
Type of Instruction	3117.1690	1	3117.1690	155.81	0.000
Type of Grouping	331.7002	2	165.8501	8.290	0.005
Interaction	26.6440	2	13.3220	0.666	N.S.
Error	6076.2332	130	20.0060		

Figure 2 plots the achievement levels of the six classes on the summative test and gives a visual description to the lack of interaction between the two treatments. The figure shows that regardless of grouping type, Mastery Learning classes performed better on the summative test in comparison to classes under conventional instruction. The means of the classes out of 40 possible points in order of magnitude are 35.36 for the Mastery Learning with heterogeneous grouping, 34.17 for the ML class with homogeneous grouping, and 32.54 for the non-grouped Mastery Learning class; while for the conventional class with heterogeneous grouping the mean is 26.35, for the homogeneously grouped control class 25.25, and 21.69 for the non-grouped conventional class. The trend is aligned with the divergence hypothesis of Kerckhoff.

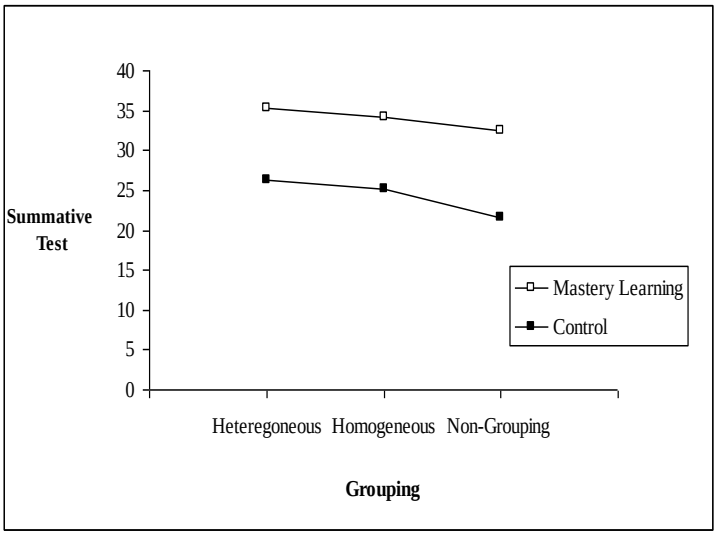

Figure 2. The interaction between type of instruction and type of grouping on summative achievement

Conditions were further compared with each other using the Newman-Keuls formula. Table 9 shows the results of this analysis.

Table 9 shows that Mastery Learning classes did better than the classes under conventional learning at $p<.01$ ($q=17.67$). Although there are no significant differences between heterogeneously or homogeneously grouped classes in terms of their summative achievement levels, there are significant differences between classes grouped either way and those which are not grouped at $p<.01$ level of significance, favoring the grouped classes ($q=5.31$ and 4.33 respectively).

Table 9. Comparison of the summative test scores of the groups by using the Newman-Keuls formula

Class Comparison	df	MS Error	Calculated q	Table Value	Significance Level
ML_{Total}+CL_{Total}	134	20.006	17.67	3.70	0.01
HT_{Total}+HM_{Total}	133	20.006	0.98	3.70	N.S.
HT_{Total}+NG_{Total}	133	20.006	5.31	4.20	0.01
HM_{Total}+NG_{Total}	133	20.006	4.33	3.70	0.01

The last analysis for the first hypothesis involves the E correlation ratios and the amount of variance accounted by each intervention. Table 10 shows this analysis.

Table 10. E correlation ratios and the amount of variance accounted by type of instruction and type of grouping on summative achievement

	E Correlation Ratio	Amount of Variance Accounted for (%)
Mastery Learning and Achievement	0.7162	51.30
Grouping and Achievement	0.2336	5.46
Multiple E Correlation Ratio	0.7534	56.76

Table 10 indicates that Mastery Learning accounts for 51.30 %, while type of grouping accounts for 5.46 % of the variance in summative achievement. Since the two-way analysis of variance results indicated a non-significant interaction (Table 8), the effects of the two treatments are additive. Thus, the amount of variance the two interventions explain in terms of summative achievement is 56.76 %.

In light of the evidence, the first hypothesis of the study is confirmed. Type of instruction and type of grouping both have significant effects on student achievement levels. Students under Mastery Learning scored significantly higher than students under conventional instruction. So long as students were grouped one way or another, they scored significantly higher than students who were not grouped. Whether students were grouped heterogeneously or homogeneously did not significantly affect performance, although both the means and percentages of students reaching the criterion level of learning show a trend aligned with the divergence hypothesis of Kerckhoff, with heterogeneous groups doing better than homogeneous groups.

The second hypothesis of the study states that all three classes under Mastery Learning method of instruction will score significantly higher than their conventional counterparts on the summative test. Table 11 shows the comparison of the classes with each other using the Newman-Keuls formula.

Table 11. Comparison of the summative test scores of the groups by using the Newman-Keuls formula

Class Comparison	df	MS	Calculated	Table	Significance
ML_{Total} and CL_{Total}	134	20.006	17.67	3.70	0.01
ML+HT and CL+HT	130	20.006	9.57	4.50	0.01
ML+HM and CL+HM	130	20.006	9.48	4.50	0.01
ML+NG and CL+NG	130	20.006	11.53	4.50	0.01
ML+HT and CL+HM	130	20.006	10.74	4.71	0.01
ML+HTand CL+NG	130	20.006	14.52	4.87	0.01
ML+HM and CL+HT	130	20.006	8.31	4.20	0.01
ML+HM and CL+NG	130	20.006	13.26	4.71	0.01
ML+NG and CL+HT	130	20.006	6.58	3.70	0.01
ML+NG and CL+HM	130	20.006	7.75	4.20	0.01

Table 11 shows that all Mastery Learning classes scored significantly higher than classes under conventional learning strategies at $p<.01$ level. When each Mastery Learning class is compared with its counterpart conventional class, the greatest difference is observed between the Mastery Learning and the control classes without grouping (q=l1.53). This is the bare effect of Mastery Learning over control conditions. The other comparisons of Mastery Learning classes with their counterparts also yielded very high q values, (q=9.57 for the heterogeneously grouped Mastery Learning and control classes, and q=9.48 for the homogeneously grouped counterparts), all comparisons significantly favoring the Mastery Learning classes.

The highest value in the cross comparisons of Mastery Learning and control classes is observed between the heterogeneously grouped Mastery Learning class and the non-grouped control class (q=14.52), followed by the difference between the homogeneously grouped Mastery Learning class and the non-grouped control class (q=13.26). This finding is totally in line with our expectations that the effect of instruction under Mastery Learning, which is already strong, is facilitated by the input of students within groups (especially heterogeneous), in comparison to non-grouped control conditions. Along the same line of thinking, the smallest difference is observed between the non-grouped Mastery Learning class and the heterogeneously grouped control conditions (q=6.58), indicating that the heterogeneously grouped conventional class comes closest to the mere effect of Mastery Learning. Thus, students seem to profit from peer interaction, and more so if this interaction is varied as is the case under heterogeneous groupings.

Table 12 shows the effect size analyses of the differences between Mastery Learning and conventional classes.

Table 12. Comparison of the differences on summative test achievement between mastery learning and conventional learning classes by means of effect size analyses

Class Comparison	Mean Difference	Standard Deviation	Effect Size
ML_{Total} and CL_{Total}	9.59	5.43	1.7644
ML+HT and CL+HT	9.01	5.14	1.7529
ML+HM and CL+HM	8.92	4.71	1.8938
ML+NG and CL+NG	10.85	5.23	2.0745
ML+HT and CL+HM	10.11	4.71	2.1464
ML+HT and CL+NG	13.67	5.23	2.6137
ML+HM and CL+HT	7.82	5.14	1.5214
ML+HM and CL+NG	12.48	5.23	2.3862
ML+NG and CL+HT	6.19	5.14	1.2042
ML+NG and CL+HM	7.29	4.71	1.5477

Table 12 shows that there is a difference of 1.76 standard deviations between all Mastery Learning classes combined and the control classes combined, which is aligned with other research both in Turkey and other countries. The biggest difference between Mastery Learning classes and their counterpart control classes is observed in the comparison of non-grouped ones (2.07 standard deviations). The difference between the homogeneously grouped Mastery Learning and control classes is slightly more than the heterogeneously grouped ones (1.89 and 1.75 standard deviations respectively), again supporting the divergence hypothesis.

The 2.14 standard deviations of difference between the heterogeneously grouped Mastery Learning class and the homogeneously grouped control class is larger than the difference between the homogeneously grouped Mastery Learning class and heterogeneously grouped control class (1.52 standard deviations), indicating again the superiority of heterogeneous grouping. This finding is further supported by the smallest difference between the non-grouped Mastery Learning class and the heterogeneously grouped control class which shows an effect size of 1.20 standard deviations. This is smaller than the difference between the non-grouped Mastery Learning class and the homogeneously grouped control class (1.54 standard deviations)

The greatest difference observed in the study is between the heterogeneously grouped Mastery Learning class and the non-grouped control class (2.61 standard deviations), followed by the difference between the homogeneously grouped Master Learning class and the non-grouped control class (2.38 standard deviations). This is confirmatory evidence to Bloom's two-sigma hypothesis that when Mastery Learning is used in combination with another intervention, their combined effects increase achievement levels of students to over two standard deviations in comparison to traditional instructional methods.

In light of the evidence, the second hypothesis of the study is strongly confirmed. Mastery Learning method of instruction produces significantly higher levels of performance for students in all comparisons with conventional instruction. In addition, when another intervention like grouping is used with Mastery Learning, a two-sigma effect is clearly produced in comparison to mere conventional methods of instruction. The trend that heterogeneous grouping is more beneficial under both Mastery Learning and conventional instruction is also observed through the analyses done for this hypothesis.

The third hypothesis states that while there will not be a difference between the homogeneously and heterogeneously grouped Mastery Learning classes in terms of their achievement levels on the summative test; these two classes will achieve significantly higher than the Mastery Learning class without grouping. The logic behind this hypothesis is that since all students are helped to reach the criterion level of learning under Mastery Learning conditions, and since all group activities are geared to enable all students to share with each other their various resources in learning as well as to provide missing links that have occurred, it should not make any difference whether they are grouped homogeneously or heterogeneously. Groupings are not made under Mastery Learning to enlarge the variation among students but rather to make it vanish (Bloom, 1971). Thus different expectations held for low and high achieving groups are obliterated by the theory. Since all students are expected to reach the criterion level of learning, they are expected to cooperate with each other towards that end within learning environments designed for this very aim. Consequently type of grouping is stripped of all expectations except being a means of enabling all students to reach high levels of learning.

The reason that homogeneously or heterogeneously grouped Mastery Learning classes were expected to enhance learning more than no grouping was due to peer guidance and sharing in group activities rather than just the guidance and direction of the teacher. In the Mastery Learning class with no grouping, the teacher went over common errors made by students after the formative tests. However, it was thought that when students are put into groups, corrective help would become more focused to the individual errors of each student by the help of peers. Secondly in groups, students would not only have the teacher as a source of help, but be exposed to another source, the contribution of their peers in completing a particular learning task. This point is not usually given due attention by teachers who need the contribution of every resource available in producing effective learning outcomes. Consequently, it was thought that grouping, whether it was homogeneous or heterogeneous under Mastery Learning conditions would be more enhancive of learning than no grouping.

Table 13 and 14 give the descriptive statistics of the three Mastery Learning classes on the formative and summative tests.

Table 13 shows that in all cases, the heterogeneously and homogeneously grouped Mastery Learning classes scored higher than the non-grouped Mastery Learning class.

Table 13. Comparison of the mean and standard deviation values of the mastery learning (ML+HT, ML+HM and ML+NG) classes in terms of formative and summative tests

		ML+HT		ML+HM		ML+NG	
Type of Test	**Poss. Pts.**	**Mean**	**Stand. Dev.**	**Mean**	**Stand. Dev.**	**Mean**	**Stand. Dev.**
Formative 1	27	25.18	2.04	25.50	1.29	24.88	1.92
Formative 2	28	25.86	1.36	25.54	2.53	24.71	2.46
Formative 3	23	22.55	0.94	22.08	1 22	20.75	1.74
Formative 4	20	19.73	0.75	19.83	0.47	19.13	1.74
Summative	40	35.36	3.08	34.17	3.44	32.54	4 22

Table 14 shows the number and percentages of students reaching the 85 % criterion level of learning in these classes.

The table shows that in all cases except for the first formative test, homogeneously and heterogeneously grouped Mastery Learning classes had a larger proportion of students reaching the criterion level of learning than the Mastery Learning class without grouping; albeit in all three classes, more than 85 % of the students reached this level on the formative tests. This finding shows the strength of Mastery Learning under all grouping conditions in raising achievement levels of students. The effect of grouping is more evident at the end of a sequence of learning tasks on the summative test, where 77.27 % of the students in the heterogeneously grouped Mastery Learning class reached the 85 % criterion level of learning, while 58.33 % of students reached the same level in the homogeneously grouped Mastery Learning class. Approximately one student in three was able to reach the criterion level in the Mastery Learning class without grouping (32.54 %). This finding, despite the strength of Mastery Learning in reducing learning differences, is in line with Kerckhoffs divergence hypothesis. Even under the facilitative effect of Mastery Learning, heterogeneous grouping enabled more students to reach higher levels of learning than homogeneous grouping, although there are no significant differences between the means in performance under these two conditions (Table 16). However, even under the homogenizing effect of Mastery Learning in enabling students to reach high levels of performance, peer interaction seems to be a resource in learning in addition to the contribution of the teacher. Students in groupings with peer interaction did seem to profit from this.

Table 14. The number and percentage of students reaching the 85% criterion level in the mastery learning (ML+HT, ML+HM and ML+NG) classes on formative and summative tests

	(85%)	ML+HT (n=22)		ML+HM (n=24)		ML+NG (n=24)	
Type of Test	**Criterion**	**Number**	**Percentage**	**Number**	**Percentage**	**Number**	**Percentage**
Formative	23	19	86.36	23	95.83	22	91.66
Formative	24	21	95.45	21	87.50	21	87.50
Formative	20	21	95.45	23	95.83	21	87.50
Formative	17	22	100.00	24	100.00	23	95.83
Summative	34	17	77.27	14	58.33	9	37.50

Table 15 shows the one-way analysis of variance results of the effect of grouping type in these three Mastery Learning classes.

Table 15. One-way analysis of variance on the summative test scores in the mastery learning (ML+HT, ML+HM and ML+NG) classes

Source	**df**	**Sum of Squares**	**Mean Squares**	**F**	**Significance Level**
Between Groups	2	92.6031	46.3016	3.3706	.0403
Within Groups	67	920.3826	13.7371		
Total	69	1012.9857			

Table 15 statistically validates the above stated arguments that grouping does have a significant effect on student achievement at $p<.05$ level. Table 15 shows the comparison of the achievement levels of the three ML classes using the Newman-Keuls formula.

Table 16. Comparison of the summative test scores of the mastery learning (ML+HT, ML+HM and ML+NG) classes by using the Newman-Keuls formula

Class Comparison	**df**	**MS Error**	**Calculated q**	**Table Value***	**Significance Level**
ML+HT and ML+HM	67	13.7371	1.55	2.83	N.S.
ML+HT and ML+NG	67	13.7371	3.67	3.40	0.05
ML+HM and ML+NG	67	13.7371	2.12	2.83	N.S.

*Table values at p=.05 level

Table 16 shows that there is no significant difference between the homogeneously and heterogeneously grouped Mastery Learning classes in terms of their achievement levels. There is a significant difference at $p<.05$ level ($q=3.67$) between the heterogeneously grouped Mastery Learning class and the one in which grouping was not used. There is no significant difference between the homogeneously grouped Mastery Learning class and the non-grouped one in terms of achievement levels; although the difference approaches the .05 level of significance (calculated q value is 2.12 against the table value of 2.83). It seems that even under Mastery Learning, heterogeneous grouping is more beneficial to students in terms of their mean performances as well as in enabling more than 3/4 of students to reach the criterion level of learning on the summative test.

Table 17 shows the effect size analyses, comparing the achievement levels of students as measured by the summative test.

Table 17. Comparison of the differences between ML+HT, ML+HM and ML+NG classes on summative test achievement by means of effect size analyses

Class Comparison	Mean Difference	Standard Deviation	Effect Size
ML+HT and ML+HM	1.19	3.44	0.3459
ML+HT and ML+NG	2.82	4.22	0.6682
ML+HM and ML+NG	1.63	4 22	0.3862

The largest difference is between the heterogeneously grouped and the non-grouped Mastery Learning classes with more than 2/3 of a standard deviation between them. There is a difference of .38 standard deviations between the homogeneously grouped and the non-grouped Mastery Learning classes. The smallest difference is between the heterogeneously and homogeneously grouped Mastery Learning classes (.34 standard deviations).

In light of the evidence, the first premise of the third hypothesis that there would be no significant differences between the heterogeneously and homogeneously grouped Mastery Learning classes is confirmed. The trend, however, is in favor of the heterogeneously grouped Mastery Learning class, especially in enabling a larger proportion of students to reach the criterion level of learning.

The second premise of the hypothesis that Mastery Learning classes either with homogeneous or heterogeneous groupings would both do significantly better than the non-grouped Mastery Learning class is only partially confirmed. Only the Mastery Learning class with heterogeneous grouping performed significantly higher than the non- grouped Master Learning class, the homogeneously grouped one not being significantly different than either the non-grouped or the heterogeneously grouped Mastery Learning classes. Thus it seems that when groupings are used under Mastery Learning conditions, heterogeneous clusters seem definitely better than no peer exchange opportunities in learning environments, and show a trend in being better than homogeneous clusters, albeit not significantly so. Homogeneous clusters on the other hand, do not produce significantly better results in comparison to no groupings, but

show a trend in that direction. All of the evidence points in the direction of the divergence hypothesis, even under Mastery Learning conditions.

The fourth hypothesis of the study states that the class under conventional instruction with heterogeneous grouping will achieve significantly higher than the conventional class under homogenous grouping, and both of these classes will attain achievement levels significantly higher than the conventional class without grouping. Table 18 shows the means and standard deviations of the three conventional instruction classes on the summative test only, since these classes did not take the formative tests.

Table 18. Comparison of the mean and standard deviation values of the conventional instruction (CL+HT, CL+HM and CL+NG) classes in terms of the summative test

		CL+HT		CL+HM		CL+NG	
Type of Test	**Poss. Pts.**	**Mean**	**Standard Dev.**	**Mean**	**Standard Dev.**	**Mean**	**Standard Dev.**
Summative	40	26.35	5.14	25.25	4.71	21.69	5.23

Table 18 shows that the highest achieving conventional class is the one with heterogeneous grouping, while the lowest achieving class is the one without grouping.

Table 19 shows the numbers and percentages of students reaching the 85 % criterion level in the three control classes. The table shows that two students (8.69 %) reached the criterion level under the heterogeneously grouped conventional class and only one student (5 %) reached the same level in the homogeneously grouped conventional class. There are no students who were able to reach the 85 % level of learning in the conventional class with no grouping.

Table 19. Numbers and percentages of students reaching the 85% criterion level on the summative test in the conventional instruction CL+HT, CL+HM and CL+NG classes

				Mastery Attainment	
Classes	**Number**	**Poss. Pts.**	**85% Criterion Level**	**Number**	**Percentage**
CL+HT	23	40	34	2	8.69
CL+HM	20	40	34	1	5.00
CL+NG	23	40	34	0	0.00
Total	66	40	34	3	4.55

Table 20 shows the results of the one-way analysis of variance done on the summative test scores of the three classes under conventional instruction with different types of grouping.

Table 20. One-way analysis of variance on the summative test scores in the conventional instruction (CL+HT, CL+HM and CL+NG) classes

Source	df	Sum of Squares	Mean Squares	F	Significance Level
Between Groups	2	269.9206	134.9603	5.0615	.0092
Within Groups	3	1679.837	26.6641		
Total	65	1949.7576	-		

Table 20 indicates that there are significant differences among the three conventional classes with different types of grouping on the summative test.

Table 21 shows the comparison of these classes using the Newman-Keuls formula.

Table 21. Comparison of the summative test scores of the conventional instruction (CL+HT, CL+HM and CL+NG) classes by using the Newman-Keuls formula

Class Comparison	df	MS Error	Calculated q	Table Value	Significance Level
CL+HT and CL+HM	63	26.6641	1.00	2.83	N.S.
CL+HT and CL+NG	63	26.6641	4.22	3.40	0.05
CL+HM and CL+NG	63	26.6641	3.23	2.83	0.05

Table 21 shows that while there are no significant differences in terms of summative achievement between the heterogeneously and homogeneously grouped conventional classes; both of these classes achieved significantly higher than the conventional class without grouping at $p<.05$ level. This finding shows that peer group interaction is more beneficial than using the teacher as the only resource of learning under conventional instruction as well. There are no significant differences between the heterogeneously and homogeneously grouped conventional classes, although the trend is aligned with the divergence hypothesis of Kerckhoff.

Table 22 shows the effect size differences among the three conventional classes with different groupings. The findings indicate that there is approximately 9/10 (.89) standard deviations of a difference between the heterogeneously grouped and the non-grouped conventional classes. There is a difference of about 7/10 (.68) standard deviations between the homogeneously grouped and the non-grouped conventional classes. A difference of about 1/4 standard deviations (.23) exists between the heterogeneously and homogeneously grouped conventional classes, the trend again favoring the divergence hypothesis.

Table 22. Comparison of the differences among the conventional instruction (CL+HT, CL+HM and CL+NG) classes on summative test achievement by means of effect size analyses

Class Comparison	Mean Difference	Standard Deviation	Effect Size
CL+HT and CL+HM	1.10	4.71	0.2335
CL+HT and CL+NG	4.66	5.23	0.8910
CL+HM and CL+NG	3.56	5.23	0.6806

In light of the evidence, the fourth hypothesis of the study is partially confirmed. While no significant differences exist between the heterogeneously and homogeneously grouped conventional classes, the trend is in line with the divergence hypothesis of Kerckhoff. However, there is a significant difference in achievement between the heterogeneously grouped conventional class and the non-grouped one, as well as between the homogeneously grouped class and the one without grouping. It seems that peer group interaction significantly facilitates learning, especially if the grouping is heterogeneous[3].

The fifth hypothesis states that grouping will differentially affect low, average, and high achieving students under different instructional methods. Table 23 shows the three-way analysis of variance of the effects of type of instruction (Mastery Learning or conventional), type of grouping[4] (homogeneous or heterogeneous), and level of achievement (high, average, or low based on the ECL[5] test) on the summative test scores of students.

Table 23 shows that type of instruction has the strongest effect ($F=103.344$, $p<.000$). Level of performance on the ECL also has significant effects at $p<.020$ level ($F=4.141$); however type of grouping does not significantly affect summative achievement.

[3] Since all students in this school are highly selective and the environment is very structured, it is unlikely that teachers would have different expectations from different groups of students. Thus, the operating factor in groupings was peer input. The more heterogeneous it was, the more facilitative to learning it seemed.

[4] It must be noted that the Mastery Learning and conventional classes with no groupings were not included in the analyses for the fifth hypothesis since the aim was to investigate the differential effects of heterogeneous and homogeneous groupings on low, average, and high achieving students.

[5] This test was given prior to the implementation of treatments in order to determine the comprehension level of students in English. It was used as a basis for forming the homogeneous and heterogeneous groups in different classes.

Table 23. Three-way analysis of variance of the effects of instruction, grouping and ECL level of performance on the summative test scores in ML+HT, ML+HM, CL+HT and CL+HM classes

Source of Variation	Sum of	df	Mean	F	Significance
Main Effects	1931.936	4	482.984	28.277	.000
Type of Grouping	23.480	1	23.480	1.375	N.S.
Level of Achievement	141.472	2	70.736	4.141	.020
Type of Instruction	1765.174	1	1765.174		.000
2-way Interaction	27.792	5	5.558	.325	N.S.
Grouping and Level	.958	2	.479	.028	N.S.
Grouping and	.005	1	.005	.000	N.S.
Level and Instruction	26.521	2	13.261	.776	N.S.
3-way Interaction	58.972	2	29.486	1.726	N.S.
Error	1315.209	77	17.081	-	-
Total	3333.910	88	37.885	-	-

Table 24 gives the combined distributions of low, average, and high achieving students in the heterogeneously (ML+HT and CL+HT) and homogeneously (ML+HM and CL+HM) grouped classes as well as their mean scores on the summative test.

Table 24. Mean achievement summative scores and numbers of low, average, and high achieving groups (based on ECL scores) in ML+HT and CL+HT combined and ML+HM and CL+HM combined classes

Type of Grouping	Low		Average		High	
	Number	Mean	Number	Mean	Number	Mean
Heterogeneous	11	29.00	23	30.26	11	33.55
Homogeneous	7	29.43	29	29.72	7	32.13

The table shows that low, average, and high achieving students under both types of grouping show a normal distribution with 11 low, 23 average, and 11 high achieving students under heterogeneous, and 7 low, 29 average, and 7 high achieving students under homogeneous groupings, due to the procedures followed in forming the groups for this study. Further, the average and high achieving students score slightly higher on the summative test when grouped heterogeneously. For low achieving students, type of grouping does not seem to make a difference when students under both Mastery Learning and conventional instruction are combined. What matters is the effectiveness of instruction (Tables 7, 11, 23).

Table 25 shows the same analysis for the homogeneously and heterogeneously grouped Mastery Learning classes.

Table 25 shows that under Mastery Learning conditions, heterogeneous grouping seems to produce higher scores for all students in comparison to homogeneous grouping. The low achieving students seem to have profited most from heterogeneous grouping (with a difference of 3.85 points on the summative test), while the differences between heterogeneously and homogeneously grouped average and high achieving students are rather small (with differences of .42 and .33 summative points respectively). It seems that under Mastery Learning conditions, type of grouping only makes a difference for low achieving students, and does not matter much for average or high achieving ones, since all students are required to reach the criterion level of learning.

Table 25. Mean achievement summative scores and numbers of low, average and high achieving groups (based on ECL scores) in mastery learning (ML+HT and ML+HM) classes

Type of Grouping	Low		Average		High	
	Number	Mean	Number	Mean	Number	Mean
Heterogeneous	5	35.60	11	34.73	6	36.33
Homogeneous	4	31.75	16	34.31	4	36.00

Table 26 shows the same analysis in the heterogeneously and homogeneously grouped conventional instruction classes.

Table 26. Mean achievement summative scores and numbers of low, average and high achieving groups (based on ECL scores) in the conventional instruction (CL+HT and CL+HM) classes

Type of Grouping	Low		Average		High	
	Number	Mean	Number	Mean	Number	Mean
Heterogeneous	6	23.50	12	26.17	5	30.20
Homogeneous	3	26.33	13	24.08	4	28.25

There is again a normal distribution in both heterogeneously and homogeneously grouped conventional instruction classes due to grouping procedures used in this study. Heterogeneous grouping seems beneficial for average and high achieving students in that order, under conventional instruction. The differences between heterogeneously and homogeneously grouped students are 2.09 points for the average, and 1.95 points for the high achieving ones respectively on the summative test in classes under conventional instruction, favoring heterogeneous grouping.

Homogeneous grouping seems to be more beneficial for low achieving students under conventional instruction with 2.83 points of difference in comparison to heterogeneous grouping. In fact, homogeneously grouped low achieving students scored higher than homogeneously grouped average students under conventional instruction. Contrary to intuition, homogeneous grouping seems to be most detrimental for average students followed by the high achieving ones under conventional instruction.

Table 27 shows the comparisons between heterogeneous and homogeneous groupings, using the Newman-Keuls formula. Table 27 shows that the differences in all comparisons are not significant. The highest q value is obtained for the difference between heterogeneously and homogeneously grouped low achieving students under Mastery Learning (q=1.96), favoring heterogeneous grouping. The second highest q value (1.79) is obtained for average achieving students under conventional instruction in favor of heterogeneous grouping.

Table 27. Comparison of the summative test scores of the low, average, and high achieving students by using the Newman-Keuls formula

Class	Comparison	df	MS Error	Calculated q	Table Value	Significance Level
TOTAL	**HM_{Low} & HT_{Low}**	16	17.081	0.22	3.00	N.S.
TOTAL	**$HT_{Average}$ & $HM_{Average}$**	50	17.081	0.66	2.86	N.S.
TOTAL	**HT_{High} & HM_{High}**	17	17.081	1.05	2.98	N.S.
ML	**HT_{Low} & HM_{Low}**	7	17.081	1.96	3.34	N.S.
ML	**$HT_{Average}$ & $HM_{Average}$**	25	17.081	0.37	2.92	N.S.
ML	**HT_{High} & HM_{High}**	8	17.081	0.17	3.26	N.S.
CL	**HM_{Low} & HT_{Low}**	7	17.081	1.37	3.34	N.S.
CL	**$HT_{Average}$ & $HM_{Average}$**	23	17.081	1.79	2.95	N.S.
CL	**HT_{High} & HM_{High}**	7	17.081	0.99	3.34	N.S.

* Table values at p=.05 level

In conclusion, type of instruction and level of prior knowledge of students in English (obtained from the ECL test before the interventions) significantly affect achievement levels, whereas type of grouping does not have a significant effect but only shows trends. In six comparisons under Mastery Learning and conventional instruction, five comparisons show that heterogeneous grouping enables students to reach higher levels of learning while only one shows that homogeneous grouping is more beneficial, albeit not significantly so. Homogeneous grouping seems more beneficial for low achieving students under conventional instruction. However, under Mastery Learning the opposite is true. On the other hand, homogeneous grouping is most detrimental for average achieving students under conventional instruction, since these students scored even lower than low achieving ones when grouped homogeneously. In all comparisons heterogeneous grouping seems to be slightly more beneficial in terms of learning outcomes in comparison to homogeneous grouping for average and high achieving

students. In light of the evidence, the fifth hypothesis is generally supported in terms of trends rather than significance levels.

Limitations of the Study and Suggestions for Further Research

The study was carried out in a military lycée in Istanbul. The population of students is very homogeneous in terms of aptitude, ability, and performance, since these students are selected through extremely competitive examinations. The students in this lycée are among the most competent in their age groups, the school being known for its superior academic excellence. Even within such a homogeneous environment, the impact of Mastery Learning on student achievement levels was very striking, and was 24.95 times stronger than the effect of prior knowledge of English (Table 22; $MS_{INS}/MS_{LEV}=24.95$).

Although this was a well controlled and carried out research, facilitated by the discipline of a military setting which was extremely cooperative[6], the effect of type of grouping did not lead to significant results when only homogeneous and heterogeneous groupings were compared, but only showed general trends favoring heterogeneous groupings. One would tend to think that under these conditions, the lack of support for grouping could not have been due to implementation errors. It is not clear why the effect of grouping was not significant but only showed a trend, unless Slavin's (1990) conclusion that ability grouping does not have positive or negative effects on low, average, and high ability students is a valid assertion. As I have remarked earlier, what makes the difference seems to be other intervening variables such as the purpose behind grouping, and implementations aligned with this purpose. What in the end has an impact is whether the goal of grouping is the equality of learning outcomes for all or the expected variations in achievement aligned with the variables used for grouping.

However, grouping showed significant effects when compared to conditions where there were no groupings, at $p<.01$ level of significance, as presented in Table 8. Table 20 shows that under conventional instruction, students in both heterogeneous and homogeneous groupings achieved significantly higher than students who were not grouped. Furthermore, there were no significant differences between homogeneously or

[6] A large number of the officers in Kuleli Military Lycée were educated at Boğaziçi University, both as teachers of English as well as of the sciences. As I was teaching three of the nine requirements (Educational Psychology, Measurement and Evaluation, and Curriculum and Instructional Programs) for the pedagogical formation courses of the teacher training program during those years at the undergraduate level, and the Research Design and Methodology course at the graduate level, all of the officers who were graduates of Boğaziçi University and were in teaching positions at this academy had been my students. In addition, the Kuleli Military High School, like Robert Academy, Şişli Terakki Lisesi, and Darüşşafaka Lisesi (the school for the orphans) has always been interested in educational intervention and research. Although these schools need the least help since they have high quality educational programs already, they are most open to ideas and educational interventions. This may indeed be the reason for their academic excellence. Consequently, when İlker Tuğal carried out the research at Kuleli, he was already a teacher officer (now he is a retired Colonel, having worked as the head of the Department of Foreign Languages at Işıklar Air Force High School in Bursa, Turkey), and I had been involved in teacher training as well as in other educational interactions through seminars and conferences at the Kuleli Military Lycée.

heterogeneously grouped students under conventional instruction as is seen in Table 20. Under Mastery Learning, the effect of grouping is only significant in one of the three comparisons, between heterogeneously grouped ML students and the non-grouped ones, favoring the heterogeneous grouping (Table 15). It seems that as Slavin (1990) points, kind of grouping is not important in affecting achievement levels of students. However, peer interaction seems to produce positive significant effects on achievement, especially under conventional instruction.

Aside from these major findings, there are also curious trends obtained through the results of this study. One interesting outcome of the study is the advantages of two opposing trends in grouping for low achieving students. Heterogeneous grouping as a trend was most beneficial for low achieving students under Mastery Learning with a difference of 3.85 summative test points in comparison to homogeneous grouping, while homogeneous grouping was most beneficial for low achievers under conventional instruction with a difference of 2.83 summative test points in comparison to heterogeneous grouping. It is not clear through this research why heterogeneous grouping worked, albeit as a trend, for all students under both instructional methods except for low achieving students under conventional instruction, (Table 25). However, this research shows that as a trend homogeneous grouping is most detrimental for average achieving students, and does not lead to better results high achieving ones. The argument that heterogeneous grouping may slow down the high achievement students did not hold for this research as well as another one (Yıldıran and Hackenberg, 1996), carried out in two very divergent settings, one in an elementary school in Nürnberg, Germany, and the other in a military academy at the senior high school level in Istanbul, Turkey. Au contraire, average and high achieving students showed a trend of profiting more from heterogeneous group interactions, whether grouping was based on ability (elementary school students in Nürnberg, Germany) or on achievement (military high school students in Istanbul, Turkey).

To see the effects of grouping more clearly in future research, increasing the number of low and high achievers may be useful. Since the distribution of low, average, and high achieving students in the total sample as well as under individual instructional methods was normal, the number of students in the low and high categories under separate instructional methods may have been too small to obtain more reliable results. Another possibility is integrating group work more extensively into instruction rather than using it only at the end of each learning task for corrective or remedial help. Perhaps by making the grouping intervention more part of instruction rather than an auxiliary to it, its effects can be more differentiated.

Conclusions and Implications

Bloom's contention that when another intervention is added to Mastery Learning, the effect of both would be even stronger than the single effect of Mastery Learning, approaching a two-sigma difference in comparison to the implementation of just conventional instruction, was the guiding proposition of this research. In fact, a difference of 2.6 standard deviations was obtained between the heterogeneously grouped Mastery Learning class and the non-grouped conventional learning class as seen in Table 11; the two interventions having an additive effect as observed in Table 8.

It is thus evident that heterogeneously grouped students who are under an instructional methodology like Mastery Learning, which enables almost all students to reach desired levels of learning (85 % criterion level), attain levels of achievement which are more than 2.5 standard deviations higher in comparison to plain conventional teaching/learning strategies.

This research was further carried out because grouping students is a contemporary trend in education and because the results from the research on grouping are inconclusive. Consequently, this research was carried out with a more complex design, including not only heterogeneous and homogeneous groupings, but also non-grouped class comparisons. The results indicated that grouping had a significant effect when heterogeneously, homogeneously, and non-grouped Mastery Learning and conventional classes were compared with each other, the grouped classes attaining significantly higher achievement levels than non-grouped classes (Tables 8, 9, 15, 16, 20 and 21). However, there were no significant differences between the homogeneously and heterogeneously grouped classes (Tables 9, 16, and 21). Thus, grouping type did not seem to affect academic performance significantly in this setting either, except in comparison to a lack of peer interaction (no grouping). The trend, however, was toward the divergence hypothesis of Kerckhoff in five out of six comparisons.

After two research projects, I tend to agree with Slavin that type of grouping does not make much of a difference in student learning. However, grouping allows peer interaction, which in this research has been shown to have significant effects on learning outcomes. It seems that students do profit from peer input and resources, no matter what kind of grouping is formed. Any type of grouping in the end enables students to share with one another. What challenges the boundaries of ethics is using groupings as a means to define our own expectations rather than what students can accomplish and can be helped to accomplish. Heterogeneous groupings have the advantage of being more stigma proof and have within their boundaries more varied resources. I believe that it is not grouping itself but what is expected from groups and what is done in groups that define the learning environment of students and determine their learning outcomes.

Instruction in the world now not only relies on the teacher but also on other fancy technologies which are always very expensive, but are only sometimes used effectively. I have no argument against making learning easier or more enjoyable, so long as these are the aims in using educational technology. However, it seems to me that we are missing the forest for the tree when technology becomes an end in itself rather than a means for raising the competence levels of students so they become more self confident and enjoy learning to learn more. It is not what we use, but what aims we have for what we use that makes the difference. Buying expensive equipment without the intention of enabling all students to reach high levels of learning can be as functional as aimless toys at best, and a great wastage of resources at worst.

On the other hand, are we using available resources which are part of the natural learning environment at no cost? One of these great resources is the students themselves and their exchange with each other that no single teacher can provide. This research shows in all conditions that peer exchange significantly facilitates learning.

Another contribution of this research is the formation of groups, not on the basis of stable and unalterable characteristics like aptitude or I.Q., but on previous levels of learning (in this case English comprehension levels), which are totally

learnable/teachable and thus alterable. The present research shows the effect of Mastery Learning to be 24 times stronger than the effect of previous English comprehension levels of students (Table 23). Thus, this research projects indicates that it is what we do in instructional settings with what we have, much more than what the students bring into this setting that has an effect on learning outcomes. This is no longer a hypothesis but a reality that needs to be faced vis à vis our choices, values, and decisions.

When we decide to teach all those who sit in front of us the areas which are esteemed as desirable domains of functionality, we have the know-how and methods of doing so. The main point is confronting our intentions in the face of values which are parts of international documents, representing supposedly universally accepted ideas and goals for humanity.

References

Acland, H. (1973). Streaming in English primary schools. *British Journal of Educational Psychology, 43*, 151-160.

Binet, A., & Simon, T. (1905a). Méthodes nouvelles pour Ie diagnostic du niveau intellectuel des anormaux (New methods for the diagnosis of the intellectual levels of subnormals). In A. Binet et al., *L'Annee Psychologique, XI,* 191-244.

Binet, A., & Simon, T. (1905b). Application des méthodes nouvelles au diagnostic du niveau intellectual chez des enfants normaux et anormaux d'hospice et d'école primaire [Application of new methods for the diagnosis of the intellectual level of normal and abnormal children of hospices and elementary schools]. In A. Binet et al., *L'Annee Psychologique, XI*, 245-336.

Bloom, B. S. (1987). A response to Slavin's Mastery Learning reconsidered. *Review of Educational Research, 57* (4), 507-508.

De Weerdt, P. A. (1996a). Peer tutoring in Dutch secondary education. *Boğaziçi University Journal, XVI, 1994-1995*, 25-38.

De Weerdt, P. A. (1996b). The Netherlands. In J. Gibson-Cline (Ed.), *Adolescence, from crisis to coping: A thirteen nation study* (pp. 153-168). Oxford: Butterwoth-Heinemann.

De Weerdt, P. A. (1998, August 5). *Counseling as a profession: A mixed call*. Keynote speech at the yearly conference of the International Association for Counseling (Incorporating IRTAC)**,** UNESCO Headquarters, Paris, France.

Hackenberg, R. (1993). *The effects of instruction, aptitude, and grouping on mathematics achievement of fourth grade German students studying under Mastery Learning and normal classroom instruction.* Unpublished M.A. thesis, Boğaziçi University, Istanbul, Turkey.

Hallinan, M. T. (1990). The effects of ability grouping in secondary schools: A response to Slavin's best evidence synthesis. *Review of Educational Research, 60*, 501-504.

Hallinan, M. T. and Sørensen, A. B. (1983). The formation and stability of instructional groups. *American Sociological Review, 48*, 838-851.

Hartley, B., & Viney, P. (1985). *American Streamline*. Oxford: Oxford University Press.

Jackson, P. W. (1968). *Life in classrooms.* New York: Holt, Rinehart and Winston.

Kelly, D. H. (1974). Track position and delinquent involvement: A preliminary analysis. *Sociology and Social Research, 58*, 380-386.

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

Kifer, E. (1973). *The effects of school achievement on the affective traits of the learner*. Unpublished Ph.D. dissertation, University of Chicago.

Newfield, J., & McElyea, V. B. (1983). Achievement and attitudinal differences among students in regular, remedial, and advanced classes. *Journal of Experimental Education, 52*, 47-56.

Reuman, D. A. (1989). How social comparison mediates the relation between ability-grouping practices and students' achievement expectancies in mathematics. *Journal of Educational Psychology, 81* (2), 178-189.

Rosenthal, R., & Jacobson, L. (1966). Teacher' expectancies: Determinants of pupils' I.Q. gains. *Psychological Reports, 19*, 115-118.

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

Slavin, R. E. (1987b). Mastery Learning reconsidered. *Review of Educational Research, 57* (2), 175-213.

Slavin, R. E. (1990). Achievement effects of ability grouping in secondary schools: A best-evidence synthesis. *Review of Educational Research, 60*, 471-499.

Tuğal, İ. (1995). *The effects of grouping and Mastery Learning method of instruction on English achievement levels of lycée one students in Kuleli military high school.* Unpublished M.A. thesis, Boğaziçi University, Istanbul, Turkey.

Yıldıran, G., & Nwabueze, B. (1991). The effects of Mastery learning and improved teaching on mathematics achievement of seventh grade Turkish students. *Boğaziçi University Journal: Educational Sciences, XIV, 1990-1991*, 111-123.

Yıldıran, G., & Hackenberg, R. (1996). The effects of instruction, aptitude, and grouping on mathematics achievement of fourth grade German students under Mastery Learning and normal classroom instruction. *Boğaziçi University Journal: Educational Sciences, XVI, 1994-1995*, 83-94.

Gruplama ve Tam Öğrenme Modeli Öğretim Yönteminin Kuleli Askeri Lisesi Birinci Sınıf Öğrencilerinin İngilizce Başarı Seviyelerine Etkisi

Özet

Çalışmanın amacı, öğretim yöntemi (Tam Öğrenme ve geleneksel öğretim yöntemleri) ve gruplandırma türünün (düşük, orta ya da yüksek erişi düzeyindeki öğrencilerin heterojen, homojen gruplandırılmaları ya da hiç gruplandırılmamaları) İstanbul Kuleli Askeri Lisesi lise birinci sınıftaki öğrencilerin İngilizce dersi erişi düzeyleri üzerindeki etkilerini incelemektir. Bulgular, Tam Öğrenme yöntemi ve gruplandırma türünün erişi düzeyi üzerinde önemli düzeyde olumlu ve çoğalgan etkilerinin olduğunu ve Tam Öğrenme yönteminin etkisinin gruplandırma etkisinden 18.8 kat daha fazla olduğunu göstermektedir. Tam Öğrenme ve geleneksel öğretim yöntemlerinde heterojen gruplandırma hiç gruplandırmamadan önemli düzeyde daha etkili olmuş, geleneksel öğretim yöntemi altında düşük erişi düzeyindeki öğrenciler dışında her iki öğretim yöntemiyle erişi düzeyini daha olumlu yönde etkilemiştir.

Anahtar kelimeler: Tam öğrenme, gruplama, İngilizce başarı seviyesi