

## THE EFFECT OF MASTERY LEARNING ON AFFECTIVE CHARACTERISTICS OF STUDENTS A QUANTITATIVE RESEARCH SYNTHESIS

**Bahrul Hayat**

UIN Syarif Hidayatullah Jakarta

bahrul\_hayat@yahoo.co.id

### **Abstract**

*Many experimental researches have been conducted until recent years to see the effect of mastery learning approach on students' cognitive behavior and affective characteristics. But the question is how much evidence is there in the existing research results provides scientific conclusions by combining existing experimental results. By treating different experiments of mastery learning as research replications, the experimental results can be combined using a meta-analysis technique. This paper shows how a quantitative research synthesis can effectively be used to combine statistical evidences of researches conducted separately and independently. The effect of mastery learning on affective characteristics of students was selected for this research synthesis. The mastery learning approach to be investigated in this research synthesis is Bloom type of mastery learning strategy. Using 26 independent comparisons, the results of study show that: a) the effect sizes of mastery learning on affective characteristics of students are heterogeneous across studies, b) the source of study, either from dissertation or journal article, does not explain the variability among the effect sizes, c) mastery learning programs using a  $\geq 75\%$  mastery criterion seem to have positive affective impact on the students, while those using  $< 75\%$  mastery criterion have no impact on the affective characteristics of students, d) the mean effect size shows a decreasing trend as the level of education increases, e) the mean effect size is highly positive for mathematics class and low positive effect for science and social studies, and f) short treatment duration has a much larger positive effect size than the long term treatment duration.*

**Keywords:** *mastery learning, affective effect size, research synthesis, statistical meta-analysis*

### **Abstrak**

*Sampai saat ini telah banyak penelitian eksperimental dilakukan untuk melihat pengaruh pendekatan mastery learning pada perilaku kognitif dan karakteristik afektif siswa. Tapi pertanyaannya seberapa banyak bukti pada hasil penelitian yang ada yang memberikan kesimpulan ilmiah dengan mengkombinasikan hasil eksperimental yang telah ada. Dengan memberi perlakuan eksperimen mastery learning yang berbeda sebagai replikasi penelitian, hasil eksperimental dapat dikombinasikan menggunakan teknik meta-analisis. Penelitian ini menunjukkan bahwa sintesis penelitian kuantitatif dapat secara efektif digunakan untuk mengkombinasikan bukti statistik dari penelitian yang dilakukan secara terpisah dan independen. Pengaruh dari mastery learning pada karakteristik afektif siswa diseleksi dari sintesis penelitian ini. Pendekatan mastery learning yang diinvestigasi pada penelitian ini adalah Bloom type of mastery learning strategy. Menggunakan 26 perbandingan*



*independen, hasil penelitian menunjukkan bahwa: a) ukuran pengaruh mastery learning pada karakteristik afektif siswa heterogen di setiap penelitian, b) sumber penelitian, dari disertasi maupun artikel jurnal, tidak menjelaskan variabilitas antar ukuran pengaruh, c) mastery learning-program yang menggunakan kriteria penguasaan  $\geq 75\%$  tampak memiliki dampak afektif yang positif pada siswa, dan yang menggunakan kriteria penguasaan  $< 75\%$  tidak memiliki dampak pada karakteristik afektif siswa, d) rata-rata ukuran pengaruh menunjukkan kecenderungan menurun seiring meningkatnya tingkat edukasi, e) rata-rata ukuran pengaruh sangat positif terhadap kelas matematika dan pengaruh positifnya lebih rendah terhadap kelas ilmu pengetahuan alam dan social, dan f) durasi perlakuan yang pendek memiliki dampak positif yang lebih luas dibanding durasi perlakuan jangka panjang.*

**Kata Kunci:** *mastery learning, ukuran pengaruh afektif, sintesis penelitian, meta-analisis statistik*

Diterima: 16 April 2015

Direvisi: 12 Mei 2015

Disetujui: 21 Mei 2015

## INTRODUCTION

The mastery learning approach to instruction has proven to be one of the most effective procedures for promoting academic attainment. Since 1968 when the mastery learning strategy was developed, many researchers have conducted empirical studies that demonstrate the effectiveness of mastery programs in a wide variety of circumstances. Programs based on the mastery learning approach are used today at all levels of education from elementary school to graduate and professional schools covering diverse subject matters.

Mastery learning can be generally defined as —instruction organized to emphasize student mastery of specific learning objectives and to deliver corrective instruction as necessary in order to achieve that goal|| (Levin, 1985, ix). Although various contemporary approaches to mastery learning have been developed and practiced, there are two basic approaches to mastery learning, namely, Bloom's (1968) Mastery Learning (ML) strategy and Keller's (1968) Personalized System of Instruction (PSI). While both approaches share one common goal—to help all students to reach a very high level in their learning—ML is primarily designed for use in group-based instruction, whereas PSI, as its name indicates, is applicable in individual-based instruction.

Various topic were covered in mastery learning research, including, among other things, (a) aptitude and rate of learning; (b) ability to understand instruction; (c) quality of instruction; (d) the factor of time in mastery; and (e) the affective effect of school learning (Block, 1974). The learning criteria that are commonly used in investigating the effectiveness of mastery programs are the retention of the learned material, transfer of this material to novel situations, lower and high mental processes attained in the process of learning, and positive affect and interest in the learning task.

Of all the criteria mentioned, student cognitive achievement is the criterion most often used by schools and researchers. The full-scale reviews of research on mastery learning by Block and Burn (1976) and Guskey and Gates (1986) report extraordinary positive effects of this approach on student achievement. While results of research on the effect of ML on cognitive achievement seem to be very promising, research on the effect of ML on affective characteristics of students shows inconclusive results.

The purpose of this paper is to synthesize quantitatively research studies on the effect of mastery learning on student affective outcomes. Specifically, the study is supposed to answer several major questions, namely: Does the ML approach to instruction yield a larger affective effect size than the traditional approach? Does the affective effect size of the ML approach vary depending upon the mastery criterion, grade level, subject matter, and treatment duration? The mastery learning approach to be investigated in this research synthesis is the Bloom type of Mastery Learning strategy. To focus the study, the construct of affect is narrowed down to cover only interest in and attitude toward the subject. Other outcomes such as self-concept, self-efficacy, and mental health are excluded from this review.

## METHODS

The first step in this research synthesis was to identify and to collect studies that investigated the effect of the ML approach on student affective outcomes. As in primary research the quality of a research review depends, to a certain extent, on the completeness of the sample of studies included. The sample of this review was limited to studies investigating the impact of ML on students' interest in and attitude toward subject matter as compared to traditional instruction. Traditional instruction is understood to me. Teacher-centered whole-group instruction in a self-contained classroom.

Most research studies were located with the help of earlier research reviews by Block and Burn (1976) and Guskey and Gates (1986). A search was performed through the database of the Current Index to journals in Education, ERIC, covering the time span 1970-1990. Since the ML approach was originally developed by Bloom and his students at the University of Chicago, the writer also searched dissertations that examined the effects of ML programs. Since the search was limited to the studies available at the University of Chicago Library, the writer does not claim that the sample obtained for this study was complete. The search yielded more than twenty studies relevant for the purpose of the synthesis. For inclusion in this research synthesis the following selection criteria were established:

1. The study had to have some type of traditional instruction for comparison.
2. The study had to involve application of ML that was clearly group-based and teacher-paced.
3. The ML program had to have 'conventional' (not computerized) feedback and corrective procedures and not one-to-one tutorial procedures.
4. The necessary statistical information for computing the effect size had to be available.

After the twenty studies were read and scrutinized with regard to the mentioned selection criteria, thirteen of the studies were qualified for inclusion in this study. The final sample for the quantitative synthesis of affective outcomes consisted of seven dissertations and six journal articles. The thirteen studies reported 26 independent comparisons investigating the impact of ML on the students' interest in or attitude toward subject matter (See Appendix A).

### **Coding Scheme**

The coding scheme was basically developed to record information on two major variables. The first is related to methodological characteristics of the study covering study design, effect size characteristics, measuring instrument, and source of study. The second deals with the treatment characteristics which include the mastery criterion, education level, subject matter, and time length (program duration). Although the variables included in coding might be related to the affective effect sizes, the information provided by the study was not as detailed and complete as

expected. Consequently, the writer reconstructed the coding scheme to produce one which is somewhat simpler than the initial one.

### Data Analysis

The data were analyzed through the following statistical procedures. First, the effect size and its variance were estimated for each study. Second, detection of outliers was conducted using the homogeneity test statistic  $Q_i$  by deleting the observation under examination (Hedges & Olkin, 1985, 256). Third, the homogeneity of effect sizes for the complete sample was tested using the homogeneity test statistic  $Q_t$ . Finally, the interaction between the explanatory variables and the effect sizes ( $Q_b$ ). The statistical fit of the model indicates that the explanatory variable is a source of systematic variance which explains the variability among the effect sizes. The estimates of effect sizes for each category of the explanatory variable were also calculated. The equations used in the analysis are as follows.

1. Formula for calculating effect sizes:

where  $\bar{Y}^E$  and  $\bar{Y}^C$  are the respective experimental and control group means and  $s$  is the sample pooled standard deviation (Hedges & Olkin, 1985, 79),

and  $n^E$  and  $n^C$  are the experimental and control group sample sizes, respectively.

2. Formula for calculating estimated variance (Hedges & Olkin, 1985, 86):

$$s = \sqrt{\frac{(n^E - 1)(S^E)^2 + (n^C - 1)(S^C)^2}{n^E + n^C - 2}}$$

$$\hat{\sigma}^2(d) = \frac{n^E + n^C}{n^E n^C} + \frac{d^2}{2(n^E + n^C)}$$

3. Formula for computing homogeneity test statistics (Hedges & Olkin, 1985, 127):
  
4. Formula for calculating weighted mean effect sizes (Hedges & Olkin, 1985, 111):
  
5. Formula for calculating the sample estimate of the variance of  $d_+$  (Hedges & Olkin, 1985, 113):

$$Q = \sum \hat{\sigma}^2(d_+) = \frac{\sum_{i=1}^k \frac{d_i}{1 + \frac{1}{2(n^E + n^C)}} (d_i)^2}{\sum_{i=1}^k \frac{1}{\hat{\sigma}^2(d_i)^2} \frac{1}{(d_i)^2}}$$

## RESULTS

The descriptive statistics for the 26 comparisons indicate that affective effect sizes range from -.53 to 1.368 with a mean of .31. Table 1 presents the effect size estimates and the values of the homogeneity test statistic  $Q_i$  to locate outliers. Using this statistic study 3 is identified as a potential outlier, since the value of the homogeneity statistic with study 3 removed is the smallest of the  $Q_i$  values. Study 21 has the next smallest  $Q_i$  value. Although study 3 seems to be a potential outlier, it was not dropped from the analysis since the value of  $Q_i$  with study 3 excluded is still large enough to detect the heterogeneity of the sample effect sizes.

The overall homogeneity statistic  $Q$  for the affective effect size of the 26 comparisons is 104.06. As a chi-square with  $k-1$  degrees of freedom (where  $k$  is the number of comparisons), the homogeneity statistic  $Q$  is significant with  $p < .001$ . This shows that the effect sizes are heterogeneous across comparisons. The weighted common effect size for this complete sample is .393, with a standard error of .051.

Since the effect sizes are heterogeneous across comparisons (complete sample), the subsequent analyses were done to identify possible sources of variability of the effect sizes. Source of study, percentage of mastery criterion, education level, subject matter, and program duration are used as explanatory variables for this analysis. Analysis by other explanatory variables was not conducted due to the limited information available.

**Table 1**  
*Effect Size Estimates and Homogeneity Statistics to Locate Outliers*

	$n^E$	$n^C$		$\sigma^2(d)$	
1	50	50	-0.090	0.040	97.82
2	32	36	0.199	0.059	103.39
3	64	65	1.368	0.038	77.42
4	24	25	1.266	0.098	96.08
5	29	24	0.364	0.077	104.05
6	28	30	1.040	0.078	98.54
7	30	31	0.515	0.068	103.83
8	26	29	0.772	0.078	102.17
9	26	29	0.900	0.080	100.75
10	27	85	0.362	0.049	104.04
11	22	15	-0.170	0.113	101.17
12	27	14	0.565	0.112	103.79
13	21	15	0.218	0.115	103.79
14	26	41	-0.410	0.064	93.57
15	17	16	-0.200	0.122	101.11
16	26	16	-0.300	0.102	99.23
17	10	13	0.438	0.181	104.05
18	23	15	-0.530	0.114	96.40
19	13	14	-0.100	0.149	102.39
20	60	33	0.359	0.048	104.03
21	61	56	1.080	0.039	91.18



	$n^E$	$n^C$		$\bar{d}$	$\sigma^2(d)$	
22	26	23	-0.320	0.083	97.73	
23	26	24	-0.030	0.080	101.75	
24	26	28	-0.080	0.074	100.93	
25	28	33	0.336	0.067	104.01	
26	86	80	0.530	0.025	103.22	

The outcome of the attempt to fit a statistical model to affective effect size data with source of study as explanatory variable is presented in Table 2.

**Table 2**  
*Analysis of Effect Size Variability with Source of Study as Explanatory Variable*

Source of Variation	Test of Homogeneity			Effect Size	
	df	Q	p	Mean	SE
Between Classes	1	0.05	Not sig		
Within Classes	24	104.01	<0.001		
Dissertation	17	78.18	<0.001	0.402	0.065
Journal	7	25.83	<0.001	0.378	0.083
Total	25	104.06	<0.001		

The between-class homogeneity statistic  $Q_b$  with value .05 is not statistically significant as a chi-square statistic with one degree of freedom. This means that the variable source of study (dissertation and journal article) does not explain the variability among the effect sizes. As indicated the highly significant within-class homogeneity statistic of 104.01, the systematic variability among effect sizes is merely due to the large within-class variation of effect sizes. The positive mean effect size of .402 (SE=.065) for dissertation studies is not significantly different from the mean effect size estimate of .378 (SE=.083) for studies from journal articles.

Testing the association between the mastery criterion and the students' affective outcomes, a significant between-class homogeneity statistic  $Q_b$  of value 32.29 (df=1,  $p < .001$ ) is obtained as shown in Table 3.

**Table 3**  
*Analysis of Effect Size Variability with Percentage of Mastery Criterion as Explanatory Variable*

Source of Variation	Test of Homogeneity			Effect	Size
	df	Q	p	Mean	SE
Between Classes	1	32.29	<0.001		
Within Classes	19	51.33	<0.001		
< 75% mastery	8	10.53	Not sig	-0.135	0.108
≥ 75% mastery	11	40.80	<0.001	0.609	0.075
Total	20	83.62	<.0001		

The within-class homogeneity statistic  $Q_w$  is still highly significant ( $Q_w=51.33$ ,  $df=19$ ,  $p<.001$ ). This may be due to studies with a  $\geq 75\%$  mastery criterion. Studies using this mastery criterion have heterogeneous effect sizes with a common effect size of .609 and standard error of .075. Little variation in effect sizes is observed in studies using a  $< 75\%$  mastery criterion. The effect sizes of this category are homogeneous and its population effect size estimate is not significantly different from zero ( $M=.135$ ,  $SE=.108$ ). Using 95% confidence bands, the effect size estimates of the two criteria are significantly different from each other. Mastery learning programs using a  $\geq 75\%$  mastery criterion seem to have positive affective impact on the students, while those using a  $< 75\%$  mastery criterion have no impact on the affective characteristics of the students.

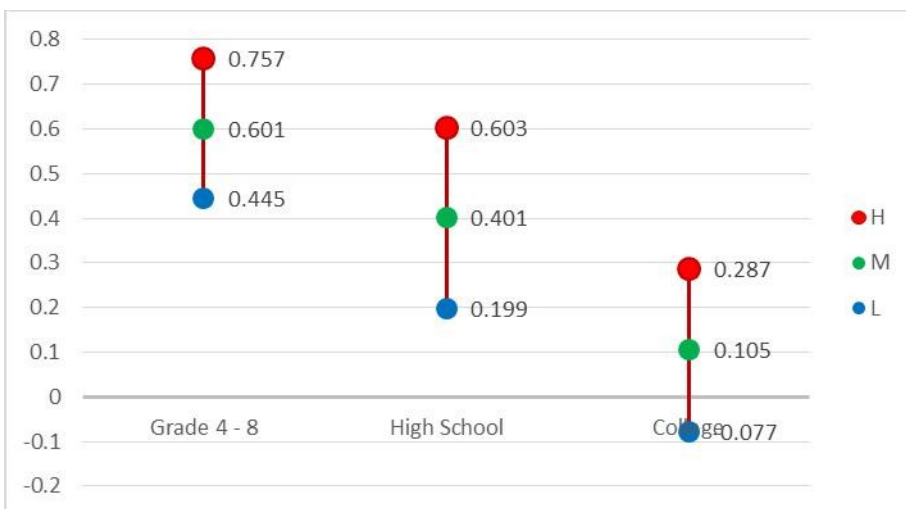
Three categories of education level are used to analyze the relationship between the variable level of education as explanatory variable and effect sizes. As the between-class homogeneity statistic  $Q_b=16.11$  ( $df=2$ ,  $p<.001$ ) indicates, level of education does explain some of the variation among the affective effect sizes as shown in Table 4.

**Table 4**  
*Analysis of Effect Size Variability with Education Level as Explanatory Variable*

Source of Variation	Test of Homogeneity			Effect	Size
	df	Q	p	Mean	SE
Between Classes	2	17.15	<0.001		
Within Classes	23	86.91	<0.001		
grade 4 – 8	9	46.46	<0.001	0.601	0.078
high school	5	19.61	<0.001	0.401	0.101
collage	9	20.84	<0.050	0.105	0.091
Total	25	104.06	<0.001		

However, it does not explain all of the systematic variation in effect sizes since a highly significant within-class homogeneity statistic  $Q_w=86.91$  is observed ( $df=2$ ,  $p < .001$ ). This large within-class homogeneity statistic is merely due to a large variation in effect size within each category of education level as indicated by significant homogeneity statistics for each category. Interestingly enough, the mean effect size computed from grades 4-8 ( $d = .601$ ,  $SE = .078$ ) is the highest, followed by high school and college. Figure 1 illustrates the decreasing trend in mean effect size as the level of education increases.

Table 5 presents the results of the analysis with subject matter as explanatory variable. For this purpose subject areas are classified into three categories, namely, science, mathematics, and social studies and language. The significant between-class homogeneity statistic  $Q_b=16.11$  ( $df=2$ ,  $p < .001$ ) allows the inference of an association between subject matter and the effect sizes.



**Figure 1**  
*Point estimate and 95% confidence band For the effect sizes classified by education level*

A significant within-class homogeneity statistic  $Q_w=87.95$  ( $df=23$ ,  $p < .001$ ) is also observed. This within-class variability is solely attributed to the mathematics and social studies/language categories.

**Table 5***Analysis of Effect Size Variability with Subject Matter as Explanatory Variable*

Source of Variation	Test of Homogeneity			Effect Size	
	df	Q	p	Mean	SE
Between Classes	2	16.11	<0.001		
Within Classes	23	87.95	<0.001		
Science	5	6.49	Not sig.	0.207	0.102
Mathematics	9	54.29	<0.001	0.637	0.079
Social/Language	9	27.17	<0.001	0.232	0.088
Total	25	104.06	<0.001		

Of particular interest is the relatively high positive mean effect size of .601 (SE= .79) for mathematics mastery programs which substantially higher than the mean effect size for science and social studies mastery programs. Science and social studies/language mastery programs both have a relatively low positive effect on students affective outcomes, with mean effect size estimates of .207 (SE= .102) and .232 (SE= 088), respectively.

The variable time length (treatment duration) addresses the question of whether the affective outcomes measured are long-term (less than eight weeks) or short-term effects (more than eight weeks). Table 6 displays results of the analysis with program duration as explanatory variable. The variable treatment duration is seen to be a significant explanatory variable with between-class homogeneity statistic  $Q_b=17.44$  ( $df=1$ ,  $p<.001$ ).

**Table 6***Analysis of Effect Size Variability with Time Length as Explanatory Variable*

Source of Variation	Test of Homogeneity			Effect	Size
	df	Q	p	Mean	SE
Between Classes	1	17.44	<0.001		
Within Classes	21	62.6	<0.001		
< 8 weeks	10	34.22	<0.001	0.584	0.071
≥ 8 weeks	11	28.38	<0.005	0.103	0.091
Total	22	80.04	<0.001		

However, this variable does not explain all of the systematic variation of effect size since the within-class homogeneity statistic  $Q_w=62.6$  ( $df=21$ ,  $p<.001$ ) is very highly significant. Both categories of treatment duration have large heterogeneous effect sizes. However, the two mean effect size estimates are significantly different from each other ( $M_1=.584, SE_1=.071$ ;  $M_2=.103, SE_2=.091$ ). the short treatment duration has a

much larger positive effect size than the long treatment duration. It seems that the longer the treatment duration, the lower the affective effect.

## DISCUSSION

Generally speaking, this quantitative synthesis of research on the effect of the Mastery Learning approach on student affect supports the findings of previous reviews that students who learned under the ML approach liked the subject they were studying. Like Block and Burn (1976) and more recently Guskey and Gates (1986), this writer finds that group-based applications of the ML strategy have positive effects on student affect. The synthesis reveals, however, that the magnitude of effect varies widely across studies. This variability of effect size suggests that conclusions from the study should be drawn cautiously.

Further analysis of the variability of effect magnitude indicates that many factors contribute to this variability. Mastery criterion, education level (age), subject matter, and treatment duration are found to be significant variables associated with affective effect size. Students who learned under a  $\geq 75\%$  mastery criterion benefitted from the ML approach. In contrast, students who learned under a  $< 75\%$  mastery criterion did not show higher affect than those learning under a traditional approach.

The affective effect size is also found to be negatively related to student age. The greater the age of students, the lower the effect size. The size effect tends to decrease sharply as the level of education to which ML is applied increases. The subject area to which ML is applied also explains the variability of effect size. Mathematics ML programs have a powerful effect on student affect, while science and social studies/language ML programs have very little positive impact on affect.

As predicted, a short treatment duration (less than eight weeks) for the ML approach yields a relatively high positive effect size, and a long treatment duration yields a very low effect size. Based on this result we may speculate that the effect seem to be a temporary condition that may result from the initial novelty of a new program.

## REFERENCES

- Ames, C. and Archer, J. (1988). Achievement goals in the classroom: students' learning strategies and motivational processes. Journal of Educational Psychology. 88 (3), 260-7.
- Anania, J. (1981). The effects of quality of instruction on cognitive and affective learning of student. Unpublished doctoral dissertation, University of Chicago.
- Anderson, L.W. (1973). Time and school learning. Unpublished doctoral dissertation, University of Chicago.
- Arlin, M.N. (1973). Learning rate and learning rate variance under mastery learning condition. Unpublished doctoral dissertation, University of Chicago.
- Block, J.H. and Burn, R.B. (1976). Mastery learning. In L Schulman (Ed) Review of Research in Education. 4, 3-49
- Cooper, H.M. (1989). Integrating Research. Newbury park, CA: Sage Publication, Inc.
- Covington, M.V. and Omelich, C.L. (1984). Task oriented versus competitive learning structures: motivational and performance consequences. Journal of Educational Psychology. 76 (6), 1038-50
- Dalton, D.W. and Hannapin, M.J. (1988). The effects of computer assisted and traditional mastery methods on computation accuracy and attitude. Journal of Educational Research. 82 (1) , 27-33
- Geeslin, D.H. (1984). A survey of pupil opinion concerning learning for mastery. Education. 105 (2), 147-50.
- Glass, G.N. (1976). Primary, secondary, and meta-analysis of research. Educational Researcher, 5, 3-8.
- Guskey, T. and Gates, S. (1986). Synthesis of research on the effect of mastery learning in elementary and secondary classroom. Educational Leadership, 43 (3), 73-80.
- Haver, W.E. (1987). Developing skills in college algebra-A mastery learning approach. Two-year College Mathematics Journal. 9 (5), 282-7.
- Hedges, L.V. and Olkin, I. (1985). Statistical methods for meta-analysis. San Diego: Academic Press, Inc.
- Jackson, G.B. (1980). Methods for integrative reviews. Review of Educational Research, 50, 438-460.
- Lee, M.M. and MacLean, J.E. (1978). A comparison of achievement and attitudes among three methods of teaching educational psychology. The Journal of Educational Research. 22 (2), 86-90.
- Levin, D.U. and Associates (Ed) (1985). Improving student achievement through mastery learning program. San Francisco: Jossey-bass Publishers.
- Long, J.C.; Okey, J.R. and Yeanny, R.H. (1981). The effects of a diagnostic-prescriptive teaching strategy on student achievement and

attitude in biology. Journal of Research in Science Teaching. 18 (6), 515-23.

Mevarech, Z.R. (1986). The role of a feedback-corrective procedure in developing mathematics achievement and self-concept in desegregated classroom. Educational Evaluation. 12, 197-203.

Nordin, A. (1979). The effects of different qualities of instruction on selected cognitive, affective and time variable. Unpublished doctoral dissertation, University of Chicago.

Slavin, R. (1987). Mastery learning reconsidered. Review of Educational Research, 57 (2), 175-213.

Tenenbaum, G. (1982). A method of group instruction which is as effective as one-to-one tutorial instruction. Unpublished doctoral dissertation, University of Chicago.

Whiting, B. and Render, G.F. (1987). Cognitive and affective outcomes of mastery learning. Clearing House. 60 (6), 276-80.

Yildiran, G. (1977). The effects of level of cognitive achievement on selected learning criteria under mastery learning and normal classroom instruction. Unpublished doctoral dissertation, University of Chicago.