Mastery Learning: Does It Work?

Robert B. Burns

Research evidence shows that mastery learning is much more effective than conventional methods. We do not yet know whether it works equally well for all kinds of learning and for all kinds of students.

The past decade has seen the rise of two distinct types of "mastery learning" instructional strategies: Bloom's (1968) Learning for Mastery (LFM) and Keller's Personalized System of Instruction (PSI). Although the LFM and PSI instructional strategies have evolved from different scientific traditions, affect classroom practice in different ways, and are typically used at different levels of education,[1] they share the common assumption that quality learning is possible for virtually all students. In fact, proponents of mastery learning often claim that as many as 90-95 percent of students can learn as well under a mastery strategy as the top 15-20 percent of students learn under nonmastery approaches. This is certainly a bold assertion, and it should be checked empirically before practitioners decide whether or not to use mastery learning. The purpose of this paper is to summarize the research data addressed to the question of "Do mastery learning instructional strategies work?"

The research evidence comparing mastery strategies with conventional methods of instruction contains an array of anecdotal impressions, research reports, experimental studies, field trials in a single classroom, and large-scale programs across a number of classrooms and content areas. One way to handle the diversity of research is to first establish a set of ground rules about choosing and weighting particular studies and then to determine, for each study, whether statistically significant results favor the mastery approach or the conventional approach. After tallying each study, one can simply tabulate the number favoring the mastery or nonmastery approach and report the final "box score." This is a fairly typical way of summarizing research in the social sciences.

Figure 1 reports the results of two recent box score reviews on the comparable effectiveness of mastery strategies, one by Block and Burns (1977), reviewing both LFM and PSI studies, and one by Kulik, Kulik, and Cohen (1979), reviewing only PSI studies. The box scores have been tabulated for three traditional measures of instructional effectiveness: cognitive achievement, retention of cognitive learning, and affective achievement. Although the form of these three types of learning measures varied considerably

[1] For a good discussion of the differences and similarities between LFM and PSI approaches to mastery learning, see Block (1974) and Block and Burns (1977).

across studies, in general the cognitive achievement measures consisted of end-of-unit and end-of-course final examinations; the retention measures consisted of the same measures given a second time with the interval between administrations ranging from several weeks to several months; and the affective achievement measures consisted of a number of rating scales ranging from interest in and attitude toward the course and teaching method to academic self-concept.

The Block and Burns review reported that for 97 comparisons of average achievement scores between mastery and nonmastery groups, 59 comparisons indicated statistically significant results in favor of mastery-taught students while only three comparisons favored the nonmastery-taught students. The remaining 34 comparisons reported no statistically significant results in favor of either method of instruction, although mastery-taught students scored higher than nonmastery-taught students in 28 of the 34 comparisons. The Kulik and others review reported that for 61 comparisons of average achievement scores, 48 comparisons indicated statistically significant results in favor of mastery-taught students while no study reported statistically significant results favoring nonmastery-taught students. The remaining 13 comparisons reported no statistically significant differences between the two methods of instruction, although mastery-taught students scored higher than nonmastery-taught students in nine of those 13 comparisons. Figure 1 also shows less extensive data for retention and affective learning measures in favor of the mastery strategies.

While this method of summarizing research is informative and yields an overall impression of the effects of mastery learning instructional strategies, it suffers from one major deficiency. The box score method provides no means of determining the *magnitude* of differences between mastery and nonmastery-taught students. Thus, while the results tabulated in Figure 1 suggest that mastery methods of instruction consistently outscore conventional methods of instruction, the box scores do not indicate whether the mastery method wins each time by a single run in close competition or wins by ten runs in a runaway. As educational psychologists have pointed out for years, the mean score differences between two groups may be statistically significant, yet at the same time the difference may be so small as to be educationally insignificant.

How Well Does it Work?

What is needed, then, is a method of reviewing literature that at least begins to answer, "*How well* do mastery strategies work?" Such a method has

recently been suggested by Glass (1978), who labeled it "meta-analysis." The major objective is to determine the average magnitude of differences between two or more instructional strategies (or any two or more variables being compared). Using this method, one can start to determine if the differences between two competing methods of instruction are educationally as well as statistically significant.

One procedure suggested by Glass is to calculate an "effect size" for each study included in a review and then to simply average these effect sizes over all studies. The effect size for each study is determined by subtracting the mean score of the nonmastery group from the mean score of the mastery group and dividing this difference by the standard deviation of the nonmastery group. This procedure is a kind of standardization, where the magnitude of mean score difference between two groups is specified in standard deviation units of the nonmastery group. The larger the effect size, the greater the difference between the mastery and nonmastery groups.

Both the Block and Burns and the Kulik and others reviews reported meta-analyses on the research

Robert B. Burns is Assistant Professor of Education, University of South Carolina, Columbia.

comparing mastery and nonmastery approaches to instruction. The average effect sizes they calculated are given in Figure 2. Block and Burns reported an average effect size for LFM cognitive achievement studies of .83 of a standard deviation while Kulik and others reported an average effect size of .49 of a standard deviation for PSI cognitive achievement studies. Block and Burns also reported an average effect size for both LFM and PSI retention studies of .67, and Kulik and others reported an average effect size of .65 for PSI affective achievement studies.

Now, what do these numbers mean in everyday language? If there were no differences between mastery and nonmastery mean scores on the learning measures, then the average effect size would be zero. Positive effect sizes indicate that, on the average, mastery students outscore the nonmastery students, and the larger the effect size, the greater the difference between mastery and nonmastery students. In general, an effect size of .20 indicates a weak effect, an effect size of .50 indicates a moderate effect, and effect sizes greater than .80 indicate strong effects. Thus, the data reported in Figure 2 indicate moderate to strong effects in favor of mastery strategies.

Another way of indicating the magnitude of the effect sizes reported in Figure 2 is to assume that effect sizes are distributed normally. The effect size of .83 reported by Block and Burns for cognitive achievement measures means that on the average, one would expect 50 percent of mastery-taught students to achieve as well as or better than the top 20 percent of the nonmastery students. Again, on the average, mastery teaching would be expected to move the typical student from the fiftieth to the eightieth percentile of the nonmastery group. By any standards, these effect sizes are impressive. One has to conclude, given the research evidence to date, that mastery methods not only work, but work very well.

For What and for Whom?

The results presented so far indicate that mastery strategies do indeed have moderate to strong effects on student learning when compared to conventional methods of instruction. It is important, however, that proponents of a given instructional method ask at least two additional questions. First, given that an instructional method works, does it work equally well for different kinds of learning outcomes? And, second, does the instructional method work equally well for different types of students? Unfortunately, the mastery learning research to date is not extensive enough to answer either of these questions conclusively. This problem is not peculiar to mastery learning research, however. All instructional research has historically been weak in distinguishing different kinds of learning outcomes and different types of students. Yet these are the types of questions that need to be answered before practitioners can make informed

Figure 1. Box Score for Mastery Learning Instructional Strategies

Learning Outcome		Number of Statistically Significant Results Favoring Mastery Approaches	No. of Results Reporting No Statistically Significant Differences Between Mastery and Conventional Approaches	Number of Statistically Significant Results Favoring Conventional Approaches
Cognitive Achievement	Block & Burns (LFM only)	59	34	3
	Kulik and others (PSI only)	48	13	0
Retention	Block & Burns (LFM and PSI)	17	10	0
	Kulik and others (PSI only)	—*	—*	—*
Affective Achievement	Block & Burrns (LFM and PSI)	16	14	0
	Kulik and others (PSI only)	19	16	—**

* No box score data reported.
** The exact number for this cell cannot be determined from the information given in the paper. The maximum possible number is 4, but the impression given by Kulik and others is that the number is 0.

decisions about the type of instructional methods required for their particular educational goals and student clientele.

Consider the first question, "Do mastery methods work equally well for different kinds of learning outcomes?" We would like to know if they work as well for higher-order learning—for example, the development of efficient cognitive strategies and problem-solving skills—as they work for lower-level learning such as the acquisition and comprehension of subject-matter content. The few studies examining this question suggest that mastery strategies do, in fact, work for both kinds of learning, but better and more systematic research is needed.

My hunch is that when many kinds of learning are considered simultaneously—both cognitive and affective—we will find that no one instructional method is best for developing them all. Rather, we will find some methods better for some kinds of learning and others better for other kinds of learning.

Now consider the second question, "Do mastery methods work equally well for different types of students?" In other words, do mastery strategies interact with the entry characteristics students bring to the learning environment? The review by Kulik and others suggests that mastery methods work equally well for both low- and high-aptitude students, but the studies are few. There is currently considerable research under the rubric of "aptitude-treatment interaction" that suggests that no one method is best for all students. Rather, some types of students learn better with one method while others learn better with a different method. My hunch is that this will also be true for mastery learning. Until more research is completed, however, I prefer to withhold judgment.

Some Answers, Some More Questions

This paper began by attempting to answer the question, "Do mastery strategies work?" I summarized two major reviews of research which found that they do in fact work. I then turned to a relatively new method of reviewing research called meta-analysis which reveals not only that mastery strategies work, but that they work very well. However, I concluded by suggesting two additional issues that need to be addressed before practitioners can make informed choices of instructional methods. The issues are summarized in the question, "Do mastery strategies work equally well for different kinds of learning and for different types of students?" Research on these points is inconclusive. This is not the fault of mastery learning per se, but merely reflects the current state of instructional research. Perhaps if practitioners begin to ask such questions, it will spur researchers to search for the answers.

Figure 2. Average Effect Sizes for Mastery Learning Instructional Strategies

Learning Outcome		Average Effect Size
Cognitive Achievement	Block & Burns (LFM only)	.83
	Kulik and others (PSI only)	.49
Retention	Block & Burns (LFM and PSI)	.67
	Kulik and others (PSI only)	—*
Affective Achievement	Block & Burns (LFM and PSI)	—*
	Kulik and others (PSI only)	.65**

* Effect sizes were not reported.

** This effect size is an average of four different effect sizes reported by Kulik and others.

References

Block, J. "A Description and Comparison of Bloom's Learning for Mastery Strategy and Keller's Personalized System of Instruction." In: J. Block, editor. *Schools, Society and Mastery Learning.* New York: Holt, Rinehart and Winston, 1974.

Block, J., and R. Burns. "Mastery Learning." In: L. Shulman, editor. *Review of Research in Education,* Volume 4. Itasca, Illinois: F. E. Peacock, 1977.

Bloom, B. "Learning for Mastery." *UCLA-CSEIP Evaluation Comment* 1: 1968.

Glass, G. "Integrating Findings: The Meta-analysis of Research." In: L. Shulman, editor. *Review of Research in Education,* Volume 5. Itasca, Illinois: F. E. Peacock, 1978.

Keller, F. "Goodbye, Teacher . . ." *Journal of Applied Behavior Analysis* 1: 79-89; 1968.

Kulik, J., C. Kulik, and P. Cohen. "A Meta-analysis of Outcome Studies of Keller's Personalized System of Instruction." *American Psychologist* 34: 307-18; 1979.