Review of Educational Research
Summer 1990, Vol. 60, No. 2, pp. 300–302

Mastery Learning Re-Reconsidered

Robert E. Slavin
Johns Hopkins University

Kulik, Kulik, and Bangert-Drowns (1990) have written a technically proficient and comprehensive meta-analysis of research on mastery learning. In addition to the usual statistical distillation of effect sizes, they have discussed many of the key substantive and methodological issues that have been central to the debate about the effectiveness of mastery learning methods, and have even discussed specific critical studies. In these respects the paper is a significant advance over earlier meta-analyses, including their own earlier articles on the topic (Kulik, Kulik, & Bangert-Drowns, 1986; Kulik, Kulik, & Cohen, 1979).

In 1987, I published what I call a "best-evidence synthesis" on group-based mastery learning in elementary and secondary schools (Slavin, 1987a). In it, I concluded that in studies of at least 4 weeks' duration, mastery learning had essentially no effect on standardized achievement measures and a small effect on experimenter-made measures, which, I argued, were often biased in favor of the mastery learning treatment. The recent Kulik et al. review takes issue with these conclusions, claiming a substantial impact of group-based mastery learning.

There are several points of agreement as well as disagreement between our reviews of the mastery learning literature. Kulik et al. (1990) agree with what was to me the key finding of my review, that the effects of mastery learning on standardized measures are effectively zero. They make the valid point that both standardized and experimenter-made tests should be considered in the review. However, the fact that the great majority of mastery learning studies used experimenter-made tests means that the "overall" effect size estimates derived by Kulik et al. are heavily influenced by these tests. Had most studies used standardized measures, they would presumably have come to the opposite conclusion. This illustrates a problem I pointed out in critiques of meta-analysis (Slavin, 1984, 1986; also see Hedges & Olkin, 1985): When meta-analyses compute average effect sizes across categories of studies which differ from one another, the means are determined by the number of studies in each category, a criterion that has no substantive meaning.

In previous writings I have explained why I think experimenter-made tests may be biased in favor of mastery learning (see Slavin, 1987a, 1987b). First, in some studies there are clear indications that the tests used as the dependent measures were designed to cover the objectives taught in the mastery learning program, without regard to what was taught in the control group. One indication of this is

This article was written under funding from the Office of Educational Research and Improvement, U.S. Department of Education Grant No. OERI-R-117-R90002. However, any opinions expressed are my own and do not necessarily represent OERI positions or policies.

the fact that many of the largest effects of mastery learning on experimenter-made measures of achievement were in studies that did not hold curriculum constant in experimental and control classes.

Second, even when curriculum is held constant, it seems likely that the mastery learning procedures hold teachers more narrowly to the mastery objectives, whereas control teachers may be teaching material that is useful or important but is not assessed on the final test.

Third, the substantial difference in outcomes on some of the studies that used both standardized and curriculum-specific measures demands some explanation. What explanation can account for the finding by Anderson, Scott, and Hutlock (1976) of an effect size of +.64 on experimenter-made tests and +.04 on standardized tests? The insensitivity of standardized tests just does not wash as an explanation in a year-long study of elementary mathematics. A much more plausible explanation for this discrepancy is the likelihood that the experimenter-made test was keyed to the objectives taught in the mastery learning classes, and the description by the authors of how the test was constructed adds to this suspicion (see Slavin, 1987a).

The second major point of difference between the Kulik et al. review and my own is in our estimates of the effects of group-based mastery learning on experimenter-made measures. My review used medians rather than means, and arrived at a median across studies using experimenter-made tests of +.225 for equal-time studies and +.31 for extra-time studies, or +.27 overall. Kulik et al. estimate a mean effect size of +.51 for all LFM (group-based mastery learning) studies in elementary and secondary schools. If they had used medians, their studies would yield an estimate of +.43. Leaving aside a few minor quibbles, the largest difference between our different estimates was that I excluded studies of less than 4 weeks' duration (3 of the Kuliks' 17 LFM studies), studies lacking evidence that the experimental and control group were initially equivalent (1 study), and studies in which there was only one class per treatment, completely confounding treatment and teacher effects (1 study). In addition, I excluded a pilot study in Chile by Cabezon (1984) in which the experimental groups started off markedly higher in socioeconomic status and IQ than the control group, including instead another cohort in which experimental and control groups were equivalent. With these changes, the Kuliks's median is identical to mine. I doubt anyone would argue that studies of 2 weeks' duration that lacked evidence of initial equivalence should be given the same weight as longer, better-controlled experiments.

However, the main issue of the effectiveness of mastery learning boils down to a question of values. The findings of positive effects of mastery learning on experimenter-made measures can be interpreted as supporting the view that this technique can help focus teachers on a given set of objectives. Educators who value this have a good rationale for using group-based mastery learning. However, the claim that mastery learning can accelerate achievement in general in elementary and secondary schools is still awaiting convincing evidence.

References

Anderson, L., Scott, C., & Hutlock, N. (1976, April). *The effects of a mastery learning program on selected cognitive, affective, and ecological variables in grades 1 through 6*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Cabezon, E. (1984). *The effects of marked changes in student achievement pattern on the students, their teachers, and their parents: The Chilean case.* Unpublished doctoral dissertation, University of Chicago.

Hedges, L, & Olkin, I. (1985). *Statistical methods for meta-analysis.* New York: Academic Press.

Kulik, C., Kulik, J., & Bangert-Drowns, R. (1986, April). *Effects of testing for mastery on student learning.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Kulik, C-L., Kulik, J., & Bangert-Drowns, R. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational Research, 60,* 265–299.

Kulik, J., Kulik, C., & Cohen, P. (1979). A meta-analysis of outcome studies of Keller's Personalized System of Instruction. *American Psychologist, 34,* 307–318.

Slavin, R. (1984). Meta-analysis in education: How has it been used? *Educational Researcher. 13*(8), 6–15, 24–27.

Slavin, R. (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. *Educational Researcher, 15*(9), 5–11.

Slavin, R. (1987a). Mastery learning reconsidered. *Review of Educational Research, 57,* 175–213.

Slavin, R. (1987b). Taking the mystery out of mastery: A response to Guskey, Anderson, and Burns. *Review of Educational Research, 57,* 231–235.

Author

ROBERT E. SLAVIN is Co-Director, Elementary School Program, Center for Research on Effective Schooling for Disadvantaged Students, The Johns Hopkins University, 3505 N. Charles St., Baltimore, MD 21218. He specializes in cooperative learning, school and classroom organization, programs for disadvantaged students, and research review.