Book Reviews

Learning Mathematics: A Survey of 12 Countries

International Study of Achievement in Mathematics. TORSTEN HUSÉN, Ed. Almqvist and Wiksell, Stockholm; Wiley, New York, 1967. 2 vols, 672 pp., illus. $19.95.

These two volumes report on a comparative study of the mathematics achievement of school children in Australia, Belgium, England, Finland, France, Israel, Japan, the Netherlands, Scotland, Sweden, the United States, and West Germany. The data were collected in the early part of 1964 by means of a series of standardized multiple-choice tests especially devised for international use. These were administered in each country to samples of 13-year-olds and of pupils finishing secondary school. Although "the tests were not devised primarily in order to make total score comparisons between countries possible and certainly not as yard sticks for an 'international contest' " (vol. 2, p. 26), such comparisons seem to have been encouraged by the publisher's press release, which declared, among other things, that "some nations do a far better job than others in teaching mathematics . . . the best overall job of mathematics instruction in public schools appears to be done in Japan. The United States is among the least effective." Thus, in a subsequent article entitled "Why Johnny Can't Add," *Newsweek* (13 March 1967) reported that "Japanese schools do the best overall job of math instruction and U.S. schools do just about the worst."

The principal scientific goal of the project was to use these comparative international data in order to identify social and educational practices that influence the student's achievement in mathematics. Presumably such information would be of great value in many aspects of educational planning, particularly in curriculum development. Since the comparability of data from the respective countries was thus vital to the success of the project, a great deal of care was devoted to the sampling procedures. The basic plan was first to sample schools within each country and then to sample pupils within schools. The adequate sampling of schools was considered important because of the possibility that there would be substantial differences among the schools of any one country. Except for two or possibly three of the countries, the stratified sampling of schools within countries appears to have been well carried out. It is much less clear, however, how representative the within-school sampling was; the authors give only sketchy information on this point: "some countries attached importance to testing whole classes while others did not" and "the subsampling process . . . appears to have given difficulty in some cases" (vol. 1, pp. 161–62). Since the local school officials were usually given the responsibility for selecting the pupils who were to take the tests at each school, it is unfortunate that no independent checks on the representativeness of these subsamples were made. Ideally, each participating pupil should have been asked to report data (for example, marks in recent courses) which could have been checked against comparable statistics based on all pupils in the school. The pupils tested did give their parents' occupations and educational levels; these data, which are summarized by country in the report, could have been checked against comparable data from national censuses as a rough measure of representativeness. In the absence of such independent checks, the representativeness of the samples of pupils tested, and hence the validity of the international comparisons, are open to question.

Most of the second volume is devoted to formulating and testing hypotheses about how differences in the organization of schools, in curricula, in instructional methods, and in social factors affect mathematics achievement. Unfortunately, the results of most of these analyses are ambiguous, because of the cross-sectional design that was used. For example, while it was frequently found that the average achievement in one type of school differed from the average achievement in another type, the data offer no way of determining whether these differences are due to the schools' different effects on achievement or to one school's being more attractive to able students than the other. A longitudinal study, in which relative changes over time in the achievement of pupils in different schools were compared, would have helped greatly to resolve many such ambiguities.

Although the authors recognize that "the design of this study does not allow for the confirmation of imputed causal relationships" (vol. 1, p. 31), the discussion and conclusions frequently betray an uncritical acceptance of certain causal interpretations even when equally plausible alternative explanations suggest themselves. For example, positive associations between the child's mathematics achievement and his parents' educational level, which were found for 13-year-olds in nearly every country, are consistently interpreted as attributable to environmental influences in the home: "It appears that parents with high socio-economic characteristics do a better job of preparing their children for school" (vol. 2, p. 254). The possibility that these relationships may be mediated, even in part, by genetic factors is given virtually no consideration. A similar attitude is revealed in the authors' recurrent use of the term "social bias" in reference to the marked tendency, in every country, for those students who finish secondary school to come from better-educated families than those who do not finish.

Potential consumers of the findings based on between-country correlations should be cautioned that some of these correlations may be seriously in error because of mistakes in the calculations for one country, Finland—mistakes that the editor discovered when the volumes were in galley form. While some approximations to the correct data for Finland are provided in an appendix to the second volume, the possible effects of the miscalculations are not discussed. A comparison between these last-minute approximations and the data used for Finland in the text, however, shows a startling shift in that country's relative ranking on mean total scores. In one of the populations of 13-year-olds (there were two samples of this age group in each country), Finland drops from third place to last place among the 12 countries in total mathe-

matics achievement; in the other 13-year-old population, the drop is from fourth to 11th place. Apparently, the entire text was written without knowledge of this error.

That these errors in Finland's data have markedly distorted many of the between-country correlations can be seen if one recomputes some of the coefficients, using the data in Table 2.2 and the corrected achievement mean for Finland. For example, in calculating by hand just a few trial coefficients, I found striking changes in magnitude and even reversals in sign. The Spearman rank-difference correlation between total mathematics achievement and "number of subjects taken in grade 12" changed from +.32 to −.10. Conversely, the correlation between total mathematics achievement and "student opportunity to learn all items" jumped from a modest .62 to a highly suggestive .95. As another consequence of the change in Finland's relative position, the tentative conclusion (vol. 2, p. 68) that late entry into school (age 7) may have a detrimental effect on mathematics achievement at age 13 is greatly strengthened. The 13-year-olds of the only two countries uniformly employing this later age of school entry (Finland and Sweden) show the lowest mean mathematics achievement.

In the interests of correcting any false impressions that might be created by the discussions based on these erroneous between-country correlations, the authors might want to consider recomputing the coefficients and distributing a revised presentation and interpretation of these findings.

The magnitude of the observed differences in performance between students in different countries is not trivial: for instance, the average 13-year-old student in Japan answered correctly nearly twice as many items (about 31) in the 70-item test as did the average 13-year-old in either the United States or Sweden (about 16 items each). If one assumes that the within-school sampling was random in each country and that these mean scores are therefore truly representative of national achievement levels in mathematics, what evidence does the study provide concerning the possible reasons for these differences? The simplest and perhaps most plausible explanation would seem to be that the countries varied considerably in the extent to which their curricula provided opportunities to learn the types of material covered in the test. We have already noted that the students' opportunity to learn the test material (as judged by their teachers) correlated very highly (.95) with between-country differences in mean achievement. The *within*-country correlations between these variables were consistently positive (median $r = .19$), although the range among countries was from nonsignificance to more than .50. It seems likely that the nonsignificant correlations in certain countries can be explained by the relative homogeneity of the mathematics curriculum within these countries. If the causal relations implied here can be accepted, an important task for curriculum experts in this and other low-scoring countries is to determine whether this apparent lack of fit between the test questions and the student's opportunity to learn this type of material is a fault of the curriculum or of the test. If the test questions can be accepted as an appropriate and reasonably representative sample of tasks in modern mathematics, then this may be one situation where "teaching for the test" makes good sense.

Considering that more than 130,000 students, 13,000 teachers, and 5000 schools from the 12 countries participated, this project stands as a major technical and, perhaps, diplomatic achievement. The educational and scientific value of the findings is much less clear, however, because the method used was not entirely adequate to the job. It can only be hoped that the planners of the next large-scale international study, which is scheduled to begin in 1968, will most seriously consider the advantages of using a longitudinal design.

ALEXANDER W. ASTIN
American Council on Education, Washington, D.C.

Social Change in a Philippine People

Mountain Arbiters. The Changing Life of a Philippine Hill People. EDWARD P. DOZIER. University of Arizona Press, Tucson, 1966. 319 pp., illus. $10.

Dozier's book is a significant and timely statement concerning an ethnographic area and a set of theoretical problems that have become important in anthropology since the early 1950's. Kalinga society of northern Luzon was last described in a major publication by Roy Barton, whose posthumous work *The Kalinga: Their Institutions and Custom Law* (University of Chicago Press, 1949) is now a classic. Dozier's book places Barton's treatment in broader ethnological perspective by comparing the social organization of the Kalinga with those of other well-known Mountain Province groups and of lowland societies that have been studied only in the two decades since World War II. The author presents an outline of social change using historical and economic controls within the period the Kalinga have been known as an entity to outside observers. Although some Philippinists may cavil at such a comprehensive treatment as this based on a relatively brief period of field research, the problems Dozier attacks and the framework within which he attacks them should provide grounds for much worthwhile research and discussion.

Qualified, first-rate anthropologists have been slow to publish books (although not articles) on their findings about Philippine societies. Almost since the beginning of the American period, at the start of the century, a small group of anthropologists—headed by H. Otley Beyer and Fay-Cooper Cole (who died in 1966 and 1961 respectively) in the early days, by Fred Eggan and the late Felix Keesing from the '30's on, and by Robert Fox and Harold Conklin since World War II—has been working away in relative obscurity on Philippine materials. Until recently, the area has been overshadowed by the more "interesting" regions of Africa, the rest of Asia and Oceania, and Amerindia. Dozier's book marks a new period in this respect. Its publication comes at a time when much work has been accomplished and is about to be put into print (not only on so-called "pagans" like the Kalinga, but on the other 80 to 90 percent of the population as well), and when general interest in the Philippines has been rekindled by the events of 1966. Even though Dozier is a brash newcomer to the field of Philippine studies (he is an outstanding figure in the field of Amerind studies) and will suffer the displeasure of the old-timers, his attractive volume is doubly important because he has beaten the others to the draw and with his outsider's point of view has been able to see things and ask questions sure to jolt the more deeply involved. Moreover, he treats the Kalinga not as "primitives" but as a