

UC Irvine

UC Irvine Previously Published Works

Title

Mathematics Achievement of Chinese, Japanese, and American Children: Ten Years Later

Permalink

<https://escholarship.org/uc/item/8rh630g0>

Journal

Science, 259(5091)

ISSN

0036-8075

Authors

Stevenson, Harold W
Chen, Chuansheng
Lee, Shin-Ying

Publication Date

1993

DOI

10.1126/science.8418494

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

- Zool. inst. Akad. nauk SSSR 131, 104 (1985).
57. N. N. Dikov, *Arkhelogicheskie pamyatniki Kamchatki, Chukotki i Verkhnei Kalymy* (Nauka, Moscow, 1977).
 58. A. V. Lozhkin and V. P. Parii, *Opyt radiouglerodnogo datirovaniya verkhnechetvertichnykh otlozhenii* (Akademiya nauk SSSR, Magadan, 1985).
 59. M. A. Kir'yak, in *Drevnie pamyatniki Severa Dal'nego Vostoka* (Akademiya nauk SSSR, Magadan, 1990), pp. 35-52.
 60. J. A. Graham and R. F. Heizer, *Quaternaria* 9, 225 (1968); R. G. Klein, *Quat. Res.* 5, 391 (1975); A. J. Jelinek, *Am. Antiq.* 57, 345 (1992).
 61. T. F. Lynch, *Am. Antiq.* 55, 12 (1990).
 62. T. D. Dillehay, *Sci. Am.* 251, 106 (April, 1984).
 63. C. V. Haynes, *Rev. Archaeol.* 12, 8 (1991).
 64. Thousands of field surveys have been conducted in mid-latitude North America during the past few decades, and many of them have sampled geomorphic contexts that would be suitable for remains dating between 20,000 and 12,000 yr B.P. (such as cave deposits, loess, or alluvium in major river valleys). The results present a stark contrast to findings in northern Eurasia, which has yielded abundant evidence of human settlement in this time range from comparable settings (60).
 65. E. J. Dixon, R. M. Thorson, D. C. Plaskett, *Natl. Geogr. Soc. Res. Rep.* 20, 129 (1985).
 66. R. E. Ackerman, "The Archeology of the Central Kuskokwim Region" (National Geographic Society, Washington, DC, 1982); J. F. Hoffecker and C. F. Waythomas, *Curr. Res. Pleistocene* 8, 105 (1991).
 67. J. F. Hoffecker, *J. Archaeol. Sci.* 15, 683 (1988).
 68. C. V. Haynes, *Can. J. Anthropol.* 1, 115 (1980).
 69. R. Bonnicksen and K. L. Turnmire, Eds., *Clovis: Origins and Adaptations* (Center for the Study of the First Americans, Corvallis, OR, 1992).
 70. C. V. Haynes, in (20), pp. 383-398.
 71. ———, *Quat. Res.* 35, 438 (1991).
 72. ———, paper presented at the First Soviet-American Archaeological Symposium on Upper Paleolithic-Paleoindian Adaptations, Leningrad, U.S.S.R., 10 to 13 July, 1989; G. C. Frison, in (26), pp. 118-122.
 73. T. H. Dillehay, G. A. Calderon, G. Politis, M. Beltrao, *J. World Prehistory* 6, 145 (1992).
 74. C. V. Haynes, *Sci. Am.* 214, 104 (June, 1966); R. L. Humphrey, *Curr. Anthropol.* 7, 586 (1966).
 75. Many researchers believe that the northern fluted points are of Holocene age and reflect either independent local development or influence from the North American Plains (9, 52).
 76. T. Goebel, W. R. Powers, N. H. Bigelow, in (69), pp. 49-79.
 77. Examples of small bifacial points in Clovis assemblages that exhibit thinned, concave bases but lack classic flutes can be found in: J. J. Hester, E. L. Lundelius, R. Fryxell, *Blackwater Locality No. 1* (Publ. 8, Fort Burgwin Research Center, Southern Methodist University, Ranchos de Taos, TX, 1972), Fig. 89B; R. M. Gramly, *Bull. Buffalo Soc. Nat. Sci.* 30 (1982), Plate 9H.
 78. Beside fluting, the most significant difference between the Clovis and Nenana projectile points is the greater size (especially length) of the former. This size difference could reflect the increased abundance of high-quality chert available to the inhabitants of the North American Plains.
 79. During the Last Glacial Maximum, humans occupied the largely treeless Dnepr-Desna Basin of the Russian Plain, using mammal bone for fuel. They may have been at least partly dependent on large natural concentrations of frozen green bone [J. F. Hoffecker, thesis, University of Chicago (1986)]. Although such concentrations also accumulated in major basins of Beringia [such as the Kolyma (34) and Tanana (27)] during the Late Pleistocene, these areas would have been effectively isolated from human populations by the broad uplands of southwestern Beringia.
 80. C. M. Aikens and T. Higuchi, *Prehistory of Japan* (Academic Press, New York, 1982); Gai Pei, in *Palaeoanthropology and Palaeolithic Archaeology in the People's Republic of China*, Wu Rukang and J. Olsen, Eds. (Academic Press, London, 1985), pp. 225-241; C. T. Keally, *Curr. Res. Pleistocene* 7, 143 (1990).
 81. Some of the contrasts in the Beringian archaeological record between the occupations associated with the 12,000 to 11,000 yr B.P. interstadial and those associated with the 11,000 to 9,500 yr B.P. stadial may reflect climatic differences. Nenana Complex assemblages contain *pièces esquillées* and planes, which may be related to woodworking, whereas Denali assemblages yield numerous burins, which appear to have been used on bone, antler, and ivory. Traces of former dwelling structures are reported from both layers 7 and 6 at Ushki I, but only those from layer 6 exhibit evidence of entrance tunnels (57).
 82. J. H. Greenberg, C. G. Turner II, S. L. Zegura, *Curr. Anthropol.* 27, 477 (1986).
 83. Support for our research in the northern foothills of the Alaska Range was provided by the National Geographic Society, the National Science Foundation, Alaska Division of Parks, the University of Alaska, and the National Park Service. Additional support for manuscript preparation was provided by the U.S. Department of Energy under contract W-31-109-Eng-38. An earlier draft of this paper was reviewed by T. A. Ager, J. P. Cook, M. E. Edwards, D. M. Hopkins, K. E. LaGory, D. J. Meltzer, C. F. Waythomas, and F. H. West.

Mathematics Achievement of Chinese, Japanese, and American Children: Ten Years Later

Harold W. Stevenson, Chuansheng Chen, Shin-Ying Lee

A decade of heightened emphasis in the United States on mathematics and science education has had little influence on academic achievement or parental attitudes. American elementary school children in 1990 lagged behind their Chinese and Japanese peers to as great a degree as they did in 1980. Comparison of the performance of elementary and secondary school students between 1980 and 1990 reveals a decline from first to eleventh grade in the relative position of American students in mathematics. Parental satisfaction with American students' achievement and education remains high and standards remain low. Innate ability continues to be emphasized by Americans as a basis for achievement. American eleventh graders report more indications of stress than do their Chinese and Japanese counterparts.

The American educational system received greater attention and scrutiny in the 1980s than in any decade since the 1950s. President Bush and governors proposed an educational agenda for the nation, commissions were appointed, and boards of education and school systems throughout the country attempted to initiate reforms aimed at improving the academic achievement of American students. We are now at a point, 10 years after the reform movement began, where it is useful to ask whether these activities have resulted in any improvement in the performance of the students.

In 1980, we initiated a comparative study of American, Japanese, and Chinese elementary school students in Minneapolis, Sendai (Japan), and Taipei (Taiwan). The results showed that Chinese and Japanese first and fifth graders greatly surpassed their American counterparts in mathematics and that Chinese children were more capable readers than the Americans (1).

The low levels of achievement found in Minneapolis are especially worrisome be-

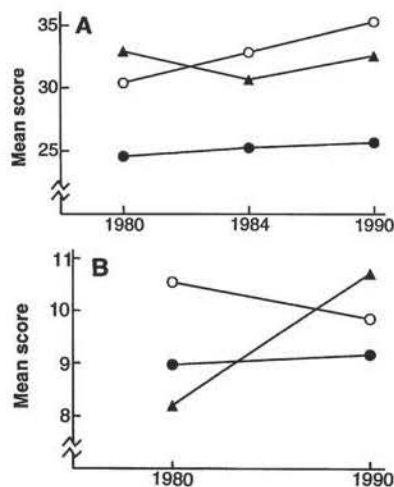
cause Minnesota students rank high among the states in mathematics achievement, and Minnesota has the highest percentage in the nation of students graduating from high school (2, 3). When problems are found in Minnesota, more severe ones might be expected to occur in many other states.

Four years after our original study, we returned to the same schools and followed up the first graders who were now fifth graders. No significant improvement occurred in the mathematics achievement of Minneapolis fifth graders during the 4 years, and cross-cultural differences were as great in 1984 as they had been in 1980 (4).

We began a new study in 1990. Once again we visited the schools included in the original study and tested a third sample of fifth graders. We also attempted to locate the first graders we had tested in 1980 (who were now eleventh graders) to trace their subsequent levels of academic achievement. To supplement the longitudinal sample, we tested over 1000 eleventh graders in each city. Data from the samples of students tested in 1980, 1984, and 1990 and 1991 form the basis of this article. These data allow us to assess possible changes in the performance of elementary school students

H. W. Stevenson is a professor of psychology and S.-Y. Lee is an assistant research scientist at the University of Michigan, Ann Arbor, MI 48109. C. Chen is an assistant professor of social ecology at the University of California, Irvine, CA 92717.

Fig. 1. (A) Fifth graders' performance on mathematics test given at three testing periods for students from Japan (▲), Taiwan (○), and the United States (●). Standard deviations for 1980: 6.45, 4.52, and 4.70, respectively. The comparable values for 1984 were 5.82, 5.68, and 5.07 and for 1990 were 5.34, 5.41, and 5.21. Sample sizes: 239, 240, and 244 (1980); 165, 164, and 118 (1984); and 273, 242, and 241 (1990). In 1980, the Japanese-American difference was 8.36 ± 0.52 ; the Chinese-American difference, 5.83 ± 0.42 ; and the Japanese-Chinese difference, 2.53 ± 0.51 . In 1984, the respective differences were 5.39 ± 0.65 , 7.59 ± 0.64 , and 2.19 ± 0.63 ; in 1990, 6.92 ± 0.47 , 9.65 ± 0.48 , and 2.73 ± 0.48 . **(B)** Fifth graders' performance on the reading vocabulary test given at two testing periods. Standard deviations for 1980: 3.39 (Japan), 1.97 (Taiwan), and 2.64 (United States); for 1990: 2.39, 1.70, and 2.58, respectively. Sample sizes: 241, 241, and 246 (1980); 273, 240, and 242 (1990). In 1980, the Japanese-American difference was 0.77 ± 0.28 ; the Chinese-American difference, 1.57 ± 0.21 ; and the Japanese-Chinese difference, 2.34 ± 0.25 . In 1990, the respective differences were 1.55 ± 0.22 , 0.68 ± 0.20 , and 0.87 ± 0.18 .



over the past decade and to evaluate the status of American students as they progress from elementary school to high school.

Outline of the Study

Our sample in 1980 included approximately 240 first graders and 240 fifth graders in each city. Nearly half of the students who were first graders in 1980 attended fifth grade at the same schools in 1984, the time of our first follow-up.

In 1990 and 1991, we attempted to follow up the first graders in our original study (now in eleventh grade). We located and obtained the cooperation of 212 students in Minneapolis, 169 in Taipei, and 93 in Sendai. Japanese eleventh graders were reluctant to participate in the study after school hours, primarily because they were preparing for the college entrance examinations they would take the following year. To place the eleventh-grade follow-up students in a broader context, we also studied cross-sectional samples totaling nearly 4000 students in Minneapolis, Taipei, and Sendai. These students attended 9, 18, and 8 different high schools, respectively. A larger number of schools was necessary to obtain a representative sample in Taipei because there are separate schools for boys and girls, and there are both regular and vocational high schools. New samples of approximately 240 fifth graders in each city were also studied in 1990 (5).

We interviewed over 85% of the mothers of the elementary school students and of the longitudinal sample of eleventh graders in each city. Over 73% of the fathers of the eleventh graders in the longitudinal sample filled out a questionnaire. We also interviewed the mathematics teachers of the eleventh graders.

We constructed all tests, interviews, and questionnaires used in the study. The tests, based on detailed analyses of the textbooks used in each city, were highly reliable (6). By including only concepts and operations common to all three locations, we reduced the potential problem of differential exposure, as discussed recently by Westbury (7). The same mathematics and reading tests were given to the fifth graders in all three testing periods.

We also included a test of general information. This test provided a measure of knowledge that children did not usually learn in school but had acquired through their everyday experiences. The test for elementary school students included items such as "What are two things a plant needs in order to grow?" and "Why can't people live under water?" The eleventh-grade test included more difficult items, such as "Why has it become possible to make smaller computers in recent years?" and "What do we mean by inflation when we talk about a country's economy?" (8).

The interviews and questionnaires covered parental attitudes and beliefs and many aspects of children's lives at home and after school. We included some common questions in the interviews and questionnaires constructed for each grade level and for each aspect of the study.

Academic Achievement

Fifth graders. The achievement gap in mathematics persisted throughout the decade (Fig. 1A). Japanese children were consistently superior to the American children, and the difference between the performance of the Chinese and American children was greater in 1990 than in 1980. The gap is large: for example, only 4.1% of the Chinese children and 10.3% of the Japanese

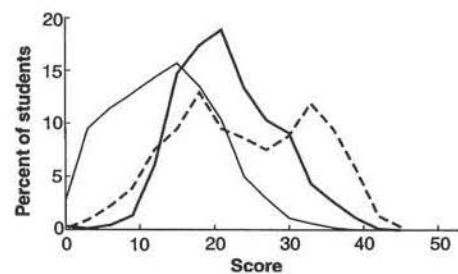


Fig. 2. Eleventh grader's scores on the mathematics test. Japan, heavy line; Taiwan, dashed line; United States, light line. Mean \pm 1 standard deviation: 21.72 ± 6.59 , 24.10 ± 9.47 , and 13.39 ± 7.06 , respectively. Sample sizes: 1120, 1475, and 1197.

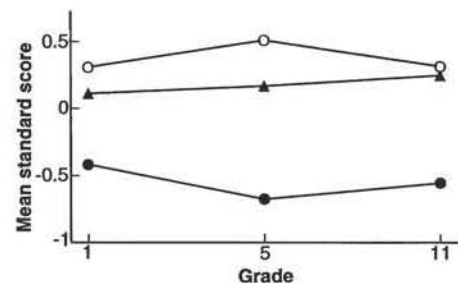


Fig. 3. Mean standard scores on the mathematics tests at grades 1, 5, and 11 for the longitudinal sample. Standardization procedures were conducted separately for each grade level. Sample sizes for Japan (▲), Taiwan (○), and United States (●), respectively: 240, 241, and 237 (1980); 165, 164, and 118 (1984); and 82, 156, and 198 (1990). The Japanese-American difference at first grade was 0.53 ± 0.09 , and the Chinese-American difference was 0.73 ± 0.09 . At fifth grade the respective differences were 0.84 ± 0.10 and 1.19 ± 0.10 ; at eleventh grade, 0.80 ± 0.10 and 0.87 ± 0.11 .

children in 1990 had scores as low as those of the average American child.

Reading ability also differed among students in the three cities. In 1980, Chinese children obtained the highest average score on reading vocabulary, and Japanese children received the lowest scores. In 1990, Japanese fifth graders had become the top performers, and American students received the lowest scores (Fig. 1B).

Eleventh graders. Of the three groups, American students continued to receive the lowest scores in mathematics at eleventh grade. Distributions of the scores of 3792 eleventh graders in the cross-sectional samples (Fig. 2) reveal a wide disparity between the average performance of the American and of the Chinese and Japanese students. Only 14.5% of the Chinese and 8.0% of the Japanese students received scores below the average score of the American students.

The distribution of scores in Taipei was bimodal. This was related to the type of school the students attended. The average score of students enrolled in vocational

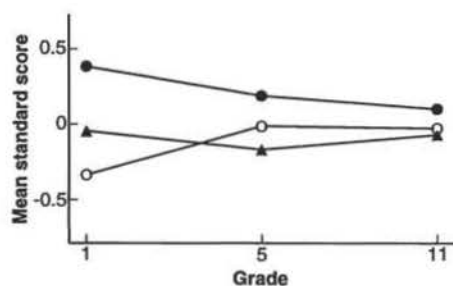


Fig. 4. Mean standard scores for the test of general information at grades 1, 5, and 11 for the longitudinal sample. Sample sizes for Japan (▲), Taiwan (○), and United States (●), respectively: 240, 241, and 237 (1980); 161, 161, and 119 (1984); and 82, 156, and 200 (1990).

high schools was 18.3; for students enrolled in the regular academic high schools it was 30.7. Even the vocational high school students obtained a higher average score than the Minneapolis students' average of 13.4 points.

The relative status of students at eleventh grade can also be compared with what existed at first and fifth grades. For this purpose, we computed standard scores at each grade level for students in the longitudinal sample (9). The data (Fig. 3) offer no evidence of improvement in the status of the American students as they moved from first through eleventh grade (10).

If we analyze the data from all students tested at each grade level rather than from only those in the smaller longitudinal sample, we find that the achievement gap in mathematics increased between first and eleventh grades. At first grade, the average Chinese and Japanese scores differed from the American scores by 0.72 ± 0.09 and 0.53 ± 0.09 standard deviation (SD) computed in the manner indicated above. By eleventh grade the corresponding differences were 1.18 ± 0.04 and 0.92 ± 0.03 SD. The increased divergence of scores, 0.45 ± 0.10 for the Chinese-American comparison and 0.39 ± 0.09 for the Japanese-American comparison, was statistically significant.

Another question sometimes raised is whether average scores offer the most appropriate measure for comparing children's levels of achievement in different countries. For example, might not the scores of high achievers, say the top 10% of students in each location, be comparable? We sought to answer this question by comparing the scores of the top 10% of American fifth and eleventh graders with those of the top 10% of Chinese and Japanese students.

The performance of the top Minneapolis students was more similar to that of the average Taipei and Sendai students than it was to that of the top students in those

cities. For example, at fifth grade the mean score of the American top achievers was 1.00 SD below the mean for the Asian top achievers but only 0.04 SD higher than the mean of all of the Asian students. The same effect occurred at eleventh grade: the top American students were 1.13 SD below the mean of the top Chinese and Japanese students and only 0.37 SD above the mean of all Asian students.

Generally, boys and girls were equally capable in mathematics during elementary school. This was true in all three cultures. At eleventh grade, however, scores of the boys in all three locations were significantly higher than those of the girls, and the gender differences were greater among Chinese and Japanese than among American students (11).

General information. The general information test offers an interesting contrast to the test of mathematics, where performance is highly dependent on academic instruction. Asian superiority was not evident in the general information scores (Fig. 4). Rather than diverge as grade level increased, as happened with the mathematics scores, the average scores of the American, Chinese, and Japanese students on the general information test became increasingly similar. At first grade, the average scores of the American students exceeded those of their Chinese peers by 0.72 ± 0.09 SD and of their Japanese peers by 0.43 ± 0.08 SD. By eleventh grade, the scores differed by only 0.13 ± 0.11 and 0.17 ± 0.13 SD, respectively. Early American superiority was also found in the performance of nearly 900 kindergarten children in the three cities (12). Scores of the American kindergarten children exceeded those of the Chinese children by 1.33 ± 0.07 SD and of the Japanese children by 0.69 ± 0.06 SD. Thus, throughout their schooling American students proved to be as capable as or even more capable than the Asian students when they were tested with items not based on the school curriculum.

We attribute the early superiority of the American children to the greater cognitive stimulation provided by their parents, who indicated that they read more frequently to their young children, took them on more excursions, and accompanied them to more cultural events than did the Chinese or Japanese parents (13). As American children grow older, parents appear to be less likely to provide the kinds of enriched out-of-school experiences that they did before the children entered first grade.

Parents' Satisfaction

One of the most dismayed findings in our 1980 study was the high level of satisfaction expressed by American mothers when they

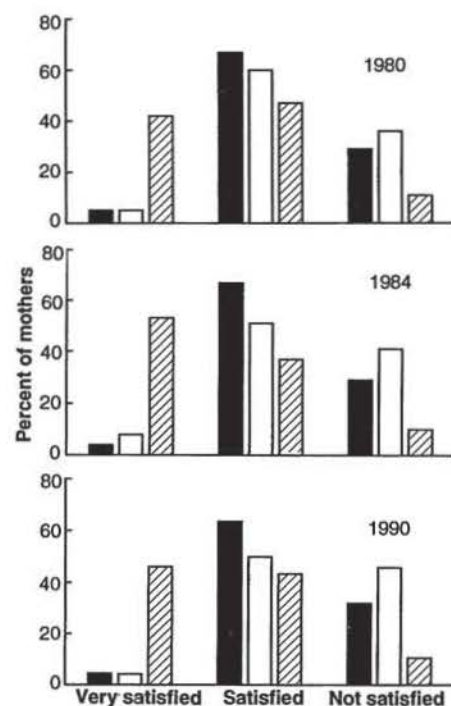


Fig. 5. Attitudes of Japanese (solid), Chinese (open), and American (hatched) mothers toward children's academic performance as reflected in ratings on a three-point rating scale. Sample sizes for Japan, Taiwan, and United States, respectively: 233, 197, and 214 (1980); 157, 104, and 117 (1984); and 258, 238, and 220 (1990).

were asked about their children's academic achievement. Few Chinese and Japanese mothers, but over 40% of the American mothers, expressed high degrees of satisfaction with their children's academic performance. One reason for their uncritical attitude may have been their lack of information about the relative status of American children compared to their peers in other industrialized countries. Since 1980, however, the media has disseminated many reports of the weaknesses of American educational systems and of American students in international comparisons. We can ask, therefore, whether the publication of these critical reports may have led to diminished satisfaction by American mothers. It is evident (Fig. 5) that it did not. If anything, somewhat more American mothers said they were "very satisfied" with their children's performance in 1990 compared with 1980.

The continued satisfaction of American mothers is surprising because they seemed to be aware of the country's low status in comparative studies. For example, we told the mothers of the eleventh graders, "Several recent studies have compared students' school achievement in different countries. One recent study compared the performance in math of high school students from

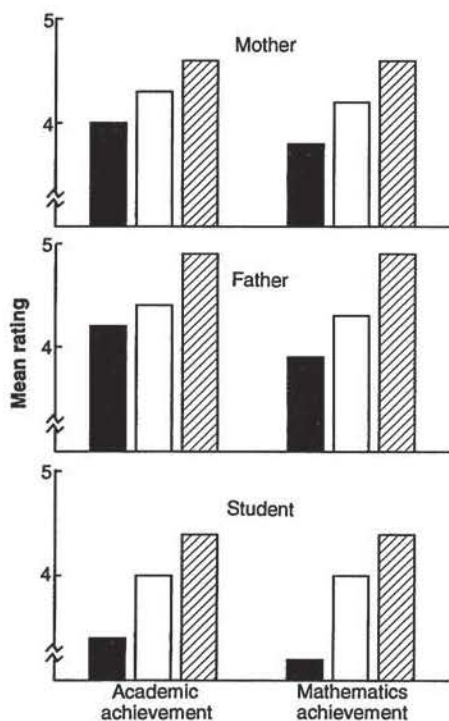


Fig. 6. Ratings made by parents and eleventh-grade students of the students' academic achievement and mathematics achievement on a seven-point rating scale ranging from "much below average" (1) to "much above average" (7). Sample sizes for Japan (solid), Taiwan (open), and United States (hatched), respectively: 157, 215, and 196 (mother); 131, 193, and 155 (father); and 149, 225, and 200 (student).

eight industrialized countries. Among the countries, where do you think (American, Taiwan, Japanese) high school students ranked in math?" American mothers estimated that American students' scores would fall between sixth and seventh place. Chinese and Japanese mothers were aware of their students' high status: they estimated that students in Taiwan and Japan would be between second and third place. In other words, Americans appeared to be aware that American education is in trouble but did not ascribe the phenomenon to their own children.

One reason American parents may express such satisfaction is that they seldom receive clear, explicit information about their children's standing in academic subjects. In the United States, most elementary school teachers convey their evaluations of children through general phrases, such as "satisfactory" or "needs improvement" or through drawings of smiling or frowning faces. As students progress through school, however, the grading system becomes more informative and parents should have a clearer idea of how their children are performing. Even so, parents of the American eleventh graders remained very positive

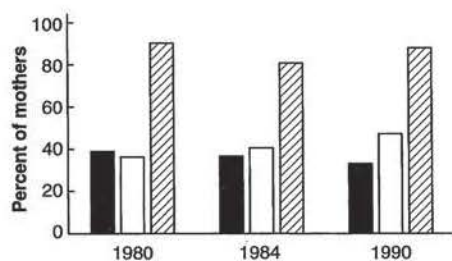


Fig. 7. Percent of Japanese (solid), Chinese (open), and American (hatched) mothers who believed their children's schools were doing a "good" or "excellent" job. Sample sizes for Japan, Taiwan, and United States, respectively: 234, 227, and 216 (1980); 159, 105, and 117 (1984); and 261, 237, and 220 (1990).

about their children's level of achievement. One-third of the American mothers said they were "very satisfied" with their child's academic achievement, a strong contrast with the 10% of the Chinese and 2% of the Japanese mothers who said they were "very satisfied." Fathers expressed similar attitudes: the percentages of fathers of eleventh graders who were "very satisfied" were 37% (Minneapolis), 8% (Taipei), and 6% (Sendai).

American students, as well as their parents, had a high regard for their own academic accomplishments. For example, we asked the eleventh graders and their mothers and fathers to rate the students' academic achievement and achievement in mathematics compared with corresponding achievements of other students of the same age. For all three sets of ratings (Fig. 6), the average ratings made by the American students and their parents were above those made by their Chinese and Japanese counterparts and were consistently higher than the score of "4," which was defined as that of average students.

We also asked the parents about how good a job they thought their children's schools were doing in educating their child. The American mothers of fifth graders were consistently more positive than the Chinese and Japanese mothers (Fig. 7). The positive attitude persisted to eleventh grade: 79% of the American mothers, but only 44% of the Chinese and 48% of the Japanese mothers, rated their children's schools as doing a "good" or "excellent" job. American fathers were also very positive: 77% said their eleventh grader's school was doing a "good" or "excellent" job, but Chinese and Japanese fathers were no more positive than the mothers: 51% of the Chinese and 40% of the Japanese fathers gave their children's schools the top ratings.

The children assimilated these positive attitudes. As early as fifth grade, American students expressed confidence that they

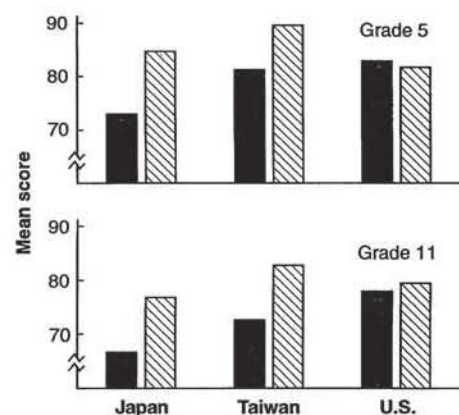


Fig. 8. Scores mothers would expect their children to obtain (solid) and scores with which the mothers would be satisfied (hatched) on a hypothetical mathematics test. Sample sizes for Japan, Taiwan, and United States, respectively: 256, 235, and 207 (grade 5) and 157, 206, and 191 (grade 11).

were doing as well in school as their parents and teachers wanted. Asian students were less sure (14).

Mothers' attitudes about the academic curricula were equally stable over the decade. Despite the publicity about the need for reform, for upgrading standards, and for assigning more homework, mothers of fifth graders were no more enthusiastic about instituting changes in 1990 than they had been in 1980 or 1984. Over 80% of the American mothers in all three periods thought the level of difficulty of the curriculum was "just right." In 1990, less than 10% thought it was either too hard or too easy. Very few Japanese or Chinese mothers ever thought the curriculum was too easy, but from 10 to 30%, depending on the year and location, thought it was too difficult.

American mothers expressed little enthusiasm about increasing the amount of homework. The great majority said the amount was "just right." A little more than a quarter of the American mothers thought their fifth graders had too little homework, both in 1980 (27%) and in 1990 (28%). Nearly a third of the Japanese mothers and 10% of the Chinese mothers expressed this attitude. These attitudes existed even though students in Taipei and Sendai spend much more time on homework than the students in Minneapolis (15).

In short, Minneapolis parents continued to be as satisfied in 1990 as they had been in 1980 with their children's academic achievement, the quality of education provided by their children's schools, the curricula, and the amount of homework assigned. In view of the persistence of such positive attitudes and of the continued poor performance of American students, one must wonder about the degree of popular

support that exists generally in the United States for extensive changes in elementary and secondary education.

Many American parents do not have high standards for their children's academic achievement. For example, we told the mothers, "Let's say there is a math test in which there are 100 points. The average score is 70. What score do you think your child would get?" We then asked, "What score would you be satisfied with?" Asian mothers were less easily satisfied than the American mothers. The tendency in all three locations, especially among mothers of fifth graders, was to expect that their child would receive an above-average score (Fig. 8). American mothers tended to give the highest estimates and Japanese mothers, the lowest. In responding to the second question, American mothers said they would be satisfied with the score they expected their child would receive (16). Chinese and Japanese mothers required a score higher than the expected score to be satisfied.

We found similar results when we asked the fathers and students these questions and when we asked about reading as well as mathematics. Americans consistently indicated that the score with which they would be satisfied was about the same as or lower than the score they expected their children or they would receive.

Ability versus Effort

In earlier reports, we presented evidence of cultural differences in the relative emphasis given to effort and ability in accounting for the academic achievement of elementary school children. Chinese and Japanese fifth graders and their mothers, following long-held tenets of Asian philosophy, stressed the importance of hard work as the route to success. American mothers, to a greater degree than Chinese and Japanese mothers, emphasized the importance of innate ability. We wondered whether this characteristic would remain applicable when students were in high school. Years of trying hard and not succeeding or of finding that high ability cannot compensate for failure to study might lead eleventh graders and their parents to modify their earlier beliefs. This did not prove to be the case. Eleventh graders in Japan and Taiwan remained strong adherents to the belief that hard work was of primary importance, and their American peers retained their strong belief in the importance of ability as a modifier of the effects of effort.

Several types of evidence support this conclusion. For example, we told the eleventh graders, "Here are some factors that may influence students' performance in mathematics: a good teacher, innate intel-

ligence, home environment, and studying hard. Which do you think is the most important factor?" More Chinese and Japanese than American students thought studying hard was the most important factor (59 and 72% versus 27%, respectively). In contrast, more than twice as many American as Chinese or Japanese students chose "a good teacher" (54% versus 18 and 14%, respectively).

We asked the teachers of the eleventh graders the same question. Among Japanese teachers, 93% selected "studying hard." Only 26% of the American teachers chose this alternative. In contrast, the first choice of 41% of the American teachers but of only 7% of the Japanese teachers was innate intelligence. Chinese teachers' choices fell between those of the American and Japanese teachers.

In another set of questions, we asked the students how strongly they believed in such statements as "Everyone in my class has about the same natural ability in math." American students disagreed to a significantly greater degree than did Chinese or Japanese students (17). When mothers were asked this type of question, American mothers strongly disagreed; Chinese and Japanese mothers were more neutral (18).

Psychological Adjustment

Critics of the academic success of Chinese and Japanese students often suggest that their high levels of performance come at great psychological cost. This criticism is based, we suspect, on informal reports and old data; we know of no recent comparative studies supporting this belief. We sought such information in our study by asking the eleventh graders to tell us about how frequently within the past month they had experienced feelings of stress, depression, aggression, and somatic complaints of possible psychological origin, such as being unable to sleep. We also asked about the frequency with which they felt nervous when they took tests or when the teacher handed tests back.

Japanese, not American students, reported the lowest frequencies of occurrence of all these characteristics. American students reported the most frequent feelings of stress, academic anxiety, and aggression. The most common source of stress was school. It was mentioned by 70% of the American students—more than three times as often as four other major sources (peers, family, sports, and jobs). Chinese students expressed somewhat more frequent feelings of depression and somatic problems (19).

These data do not support the Western stereotype of Asian students as tense young persons driven by relentless pressures for academic excellence. Had they exhibited

frequent indications of psychological disturbance, it would be reasonable to argue that high standards for academic achievement may occur at unacceptable costs. Rather, it was the American students who were more likely to express indications of distress. We believe this occurred because American students do not have a clear idea about the importance they should place on education. Chinese and Japanese students are expected to devote themselves primarily to their studies. American students, in contrast, are faced with many opposing demands. For example, more than three times as many American as Asian students have after-school jobs (74% versus 21%) and more than twice as many have "dates" (over 85% versus 37%). The motivation for economic independence and broad social experience, as well as the desire to engage in sports and to assist with family chores, make it difficult for American high school students to devote themselves wholeheartedly to their studies.

Implications

Unless radical reforms can be instituted in American educational systems, it seems unlikely that American students will lead the world in mathematics by the end of this century, one of the well-publicized goals adopted by the governors and President Bush in 1990. American schools may currently be fulfilling the roles expected of them by the American public, but its expectations prove to be insufficient when judged by international standards—American students received significantly lower scores on a curriculum-based mathematics test than their Chinese and Japanese peers, there was no significant improvement in the scores of fifth graders on the same test given three times over a period of 10 years, and the status of American students relative to their Chinese and Japanese peers declined between the first and eleventh grades.

These differences cannot be attributed to differential sampling. Enrollment in school is nearly universal among first and fifth graders in all three locations, and the percentage of adolescents enrolled in high school is similar if vocational high schools are included, as was the case in our study. Other factors were also carefully controlled, such as the time in the school year the data were collected, testing conditions, and the relevance of the test items.

We conclude that the achievement gap is real, that it is persistent, and that it is unlikely to diminish until, among other things, there are marked changes in the attitudes and beliefs of American parents and students about education. American parents appeared to be no more likely in

1990 and 1991 than they were in 1980 to believe that there is an urgent need for educational reform. They did not seem to be incensed by the low levels of performance by American students. Rather, they appeared to be pleased with their children's academic achievement, to be satisfied with the job their children's schools were doing, and to believe that children's innate abilities guide their course of progress through school. Attitudes and beliefs are difficult to change. But the likelihood of improving the nation's competitive position through better education depends, at least in part, on changing such optimistic but ultimately self-defeating views.

REFERENCES AND NOTES

1. H. W. Stevenson, S.-Y. Lee, J. W. Stigler, *Science* 231, 693 (1986).
2. *Kids Count Data Book* (Center for the Study of Social Policy, Washington, DC, 1989).
3. *The State of Mathematics Achievement* (National Center for Educational Statistics, U.S. Department of Education, Washington, DC, 1991).
4. H. W. Stevenson and J. W. Stigler, *The Learning Gap* (Summit, New York, 1992).
5. Consent was obtained from the American children's parents for their participation. In Japan and Taiwan such rights of consent are vested in school authorities and teachers. The students and their parents were told that if there was anything they felt uncomfortable about answering, they could go on to the next question. The examiners, all native speakers and residents of each city, explained that the session could be terminated if the respondent wished.
6. The mathematics test given to first and fifth graders had 54 items arranged in order of difficulty. Some items required only computation; others required application of mathematical principles to word problems. We tested all children individually. Any child who failed to answer four successive items correctly was stopped. The eleventh-grade mathematics test contained 46 items covering a broad range of problems. A 40-minute time limit was imposed. Questions on both tests were open-ended, and no multiple-choice items were included. Reliabilities of the mathematics tests were high: Cronbach alphas, a measure of test reliability defined by psychological testing theory, ranged from 0.86 to 0.95 for the elementary school test and from 0.91 to 0.95 for the high school test. The reading vocabulary test was highly reliable in Minneapolis and Sendai but lower in Taipei. Cronbach alphas were, respectively, 0.80, 0.88, and 0.61.
7. I. Westbury, *Educ. Res.* 22, 18 (1992).
8. The general information test for the elementary school students contained 26 items; the eleventh-grade test contained 12 items. At first and fifth grades, students proceeded through the test until they missed four successive items. Eleventh graders responded to as many of the questions as possible within a 12-minute time limit. The reliability of the test ranged from 0.79 to 0.91 for first and fifth graders; at eleventh grade it was 0.67 (Sendai), 0.80 (Taipei), and 0.82 (Minneapolis).
9. We determined standard scores separately for each grade by combining data for students in all three locations in a single distribution. We then computed the standardized scores at each grade level using the mean and standard deviation of the weighted sample to yield equal proportions of students from each of the three locations.
10. Longitudinal comparisons across the 10-year span assume that the eleventh-grade sample of students was not biased. This assumption proved to be valid for the Chinese and Japanese students. Comparisons of the first-grade scores in mathematics, reading, and a cognitive test composed of ten subtests revealed no significant differences between eleventh graders included in the longitudinal sample and those who were not. For the American students, the average first-grade scores of the 29 students we were unable to include in the follow-up sample were significantly lower on all three tests than were those of the 212 follow-up students. On the mathematics test, for example, the respective scores were 15.07 ± 1.04 and 17.30 ± 0.36 ; for the reading test, the respective scores were 5.72 ± 1.29 and 10.37 ± 0.67 . We conclude that the achievement scores of the American eleventh graders would have been even lower if the 29 additional American students had been included.
11. The mean difference between the scores of the eleventh-grade boys and girls in Minneapolis was 1.36 ± 0.45 ; in Taipei, 4.19 ± 0.49 ; and in Sendai, 5.43 ± 0.34 . The respective numbers of boys in the three cities were 478, 617, and 639; the numbers of girls were 526, 851, and 480.
12. H. W. Stevenson, S.-Y. Lee, T. Graham, in *Handbook of Research on Early Education*, B. Spodek, Ed. (Macmillan, New York, in press).
13. H. W. Stevenson *et al.*, *Monogr. Soc. Res. Child Dev. Ser. No. 221* (1990), vol. 55.
14. The mean ratings on a seven-point scale for responses referring to parents and teachers were, respectively, 3.92 ± 3.77 (Japanese), 4.20 and 4.26 (Chinese), and 5.27 and 5.00 (American), where 7 indicates "very true" and 4 indicates "somewhat true." Sample sizes for Japan, Taiwan, and the United States, respectively: 273, 242, and 241. The Japanese-American differences were 1.35 ± 0.13 (parents) and 1.23 ± 0.12 (teachers); the Chinese-American differences were 1.07 ± 0.14 (parents) and 0.74 ± 0.13 (teachers); and the Japanese-Chinese differences were 0.28 ± 0.14 (parents) and 0.49 ± 0.12 (teachers).
15. C. Chen and H. W. Stevenson, *Child Dev.* 60, 551 (1989).
16. The differences between the satisfactory and expected scores for the United States, Taiwan, and Japan were, respectively, -0.79 ± 0.69 , 8.51 ± 0.60 , and 11.54 ± 0.73 (grade 5); and 1.49 ± 1.00 , 10.47 ± 0.87 , and 10.08 ± 0.94 (grade 11).
17. On a seven-point scale, the mean ratings for the Japanese, Chinese, and American students were 3.80, 3.84, and 2.86. On this scale, 1 indicated "strongly disagree" and 4 indicated "neither agree nor disagree." Sample sizes for Japan, Taiwan, and the United States, respectively: 1060, 1472, and 1044. The Japanese-American difference was 0.95 ± 0.07 ; the Chinese-American difference was 0.98 ± 0.06 ; and the Japanese-Chinese difference was 0.04 ± 0.06 .
18. On a seven-point scale, the mean ratings for the Japanese, Chinese, and American mothers were 3.50, 2.97, and 1.74. Sample sizes for Japan, Taiwan, and the United States, respectively: 157, 215, and 194. The Japanese-American difference was 1.76 ± 0.13 ; the Chinese-American difference was 1.23 ± 0.14 ; and the Japanese-Chinese difference was 0.53 ± 0.16 .
19. These results are discussed in greater detail in D. Crystal *et al.*, in preparation.
20. Our study was undertaken with S. Kitamura and S. Kimura of Tohoku Fukushi University in Sendai, Japan, and C.-C. Hsu and H.-J. Ko of National Cheng Kung University, Taiwan. The research was supported by a grant from the National Science Foundation. Collection of the Taiwan data was supported by grants from the National Science Council.