Compensatory Education And the Theory of Intelligence

by Arthur R. Jensen

Compensatory education has made its least impressive impact on just those variables that it was originally intended to improve the most: I.Q. and scholastic achievement, claims Mr. Jensen. The plain truth, he says, is that compensatory programs have not resulted in any appreciable, durable gains in these areas.

THE PAST 20 years have been a period of unparalleled affluence for public education and educational research in the U.S. When the history of this era is written, two features will stand out prominently: racial desegregation of the schools and large-scale experimentation with compensatory education.

The nation focused its educational resources during this period primarily on extending the benefits of education to every segment of the population — especially to those groups that historically have derived the least benefit from the traditional system of schooling. During the past 20 years more young people have gone to school for more years and have obtained more diplomas, per capita, in the U.S. than in any other nation. Fifty percent of U.S. high school graduates in the 1970s went on to college.

These proud facts are one side of the picture. The other side is much less complimentary and should shake any complacency we Americans might feel. The past 20 years, which have brought the most energetic large-scale innovations in the history of U.S. education, have also brought an accelerating decline in Scholastic Aptitude Test scores. And there are other signs of malaise as well. On objective measures of the average level of educational achievement, the U.S. falls below all other industrialized nations, according to the International Association for the Evaluation of Educational Achievement.[1] In fact, average levels of educational achievement lower than that of the U.S. are found only in the industrially underdeveloped nations of the Third World.

Illiteracy in the U.S. has been grossly underestimated. Until recently, the U.S. Census Bureau routinely estimated the rate of illiteracy as the percentage of Americans with fewer than six years of schooling. The 1980 Census found that only two-tenths of 1% (0.2%) of the U.S. population between the ages of 14 and 24 met this definition of illiteracy — a rate that was the same for both black and white Americans.

Simple tests of actual reading ability reveal a much less rosy picture, however. According to lawyer and psychologist Barbara Lerner, evidence collected by the National Assessment of Educational Progress shows that "the overall rate of illiteracy for cohorts reaching their 18th birthday in the 1970s can safely be estimated to have been at least 20%. . . . [Moreover, the] black-white gap was still dramatic: 41.6% of all black 17-year-olds still enrolled in school in 1975 were functionally illiterate."[2] Lerner goes on to emphasize the broad implications of this finding:

> On this basis, it would have seemed reasonable to predict serious shortages of literate workers throughout the 1980s and perhaps beyond, along with high levels of structural unemployment, particularly among younger black workers, and increasing difficulty in meeting economic competition from foreign countries with more literate work forces.[3]

CLEARLY, THOSE conditions that originally gave rise to the aims and aspirations of compensatory education are as relevant today as they were 20 years ago. Of the many lessons that can be learned from assessments and meta-analyses of the results of 20 years of compensatory education, I intend to dwell in this article on what seems to me to be one of the most important. Because the lesson on which I will dwell is one of the clearest and seemingly least-debatable findings of studies of compensatory education programs of all kinds and because this lesson has important implications for both theory and practice, it is peculiar that this lesson has been soft-pedaled in most published summaries of compensatory education outcomes.

ARTHUR R. JENSEN is a professor of educational psychology in the School of Education, University of California, Berkeley.

The lesson to which I refer is this: compensatory education has made its least impressive impact on just those variables that it was originally intended (and expected) to improve the most: namely, I.Q. and scholastic achievement. The plain truth is that compensatory programs have not resulted in any appreciable, durable gains in I.Q. or scholastic achievement for those youngsters who have taken part in them. This is an important discovery, and the fact that we do not like this outcome or that it is not what we expected neither diminishes its importance nor justifies downplaying it. Rather, we are challenged to try to understand its theoretical implications for the study of intelligence and its practical implications for the practice of education.

Let us not be distracted from trying to understand the discrepancy between the expected and the actual outcomes of compensatory education programs by the too-easy response of retroactively revising our original expectations. We should gain more from our 20 years of experience than just a list of excuses for the disappointing discrepancy between our expectations and the actual results.

To be sure, Head Start and other compensatory education programs have produced some positive gains. The fact that the bona fide benefits of compensatory education have not been primarily cognitive in nature and are not strongly reflected in academic achievement per se should not detract from the social importance of these gains. The positive outcomes of Head Start and similar programs include such things as the improvement of participants' nutrition and of their medical and dental care. The list of positive outcomes also includes greater involvement of parents in their children's schooling, noticeable improvement in the children's attitudes toward school and in their self-esteem, fewer behavioral problems among participants, fewer retentions in grade, and a smaller percentage of special education placements.[4]

The socially desirable outcomes have not been accompanied by marked or lasting improvement in either I.Q. or academic performance.

These socially desirable outcomes have not been accompanied by marked or lasting improvement in either I.Q. or academic performance, however. Even the smaller percentage of special education placements may be attributable to teachers' and administrators' knowledge that certain children have taken part in Head Start or other compensatory education programs, because such children are less apt than nonparticipating peers to be labeled as candidates for special education. Gene Glass and Mary Ellwein offer an insightful observation on this point in their review of *As the Twig Is Bent*, a book on 11 compensatory education programs and their outcomes, as assessed by the Consortium for Longitudinal Studies. According to Glass and Ellwein:

> [T]hose whose ideas are represented in *As the Twig Is Bent* see themselves as developmental psychologists molding the inner, lasting core of the individual — one can almost visualize the cortical wiring they imagine being rearranged by ever-earlier intervention. And yet the true lasting effects of a child's preschool experiences may be etched only in the attitudes of the professionals and in the records of the institutions that will husband his or her life after preschool.[5]

Even studies of those compensatory programs that involve the most intensive and prolonged educational experience show

Illustration by John McGuffin

the effects of such programs on I.Q. to be relatively modest and subject to "fade-out" within one to three years. The highly publicized "Miracle in Milwaukee" Study by Rick Heber and Howard Garber appears to be a case in point. In that study, the researchers gave intensive training designed to enhance cognitive development to children who were deemed at risk for mental retardation because of their family backgrounds. The training lasted from birth until the participants entered school. Unfortunately, no detailed account of the conduct of the Milwaukee Study or of its long-term outcomes has yet appeared in any refereed scientific journal. Because the data are not available for full and proper critical review, I cannot legitimately cite this study with regard to the effects of early intervention on subsequent intelligence and scholastic achievement.

Fortunately, a similar study — the Abecedarian Project,[6] currently under way in North Carolina — is being properly reported in the appropriate journals, and the researchers conducting this study promise the kind of evaluation that Heber and Garber have failed to deliver. From infancy to school age, children in the Abecedarian Project spend six or more hours daily, five days a week, 50 weeks a year, in a cognitive training program. Their I.Q. gains, measured against a matched control group at age 3, look encouraging. However, the possibility exists that the program has merely increased participants' I.Q. scores and not the underlying *g* factor of intelligence that the I.Q. test is intended to measure and upon which its predictive and construct validity depend.[7]

Probably the most scholarly, thorough, and up-to-date examination of the variety of experimental attempts to improve intelligence and other human abilities is *How and How Much Can Intelligence Be Increased*, edited by Douglas Detterman and Robert Sternberg.[8] In a review of this book, I said:

> What this book may bring as something of a surprise to many psychologists who received their education in the 1950s and '60s, in the heyday of what has been termed "naive environmentalism" in American educational psychology, is the evident great difficulty in effecting practically substantial and durable gains in individuals' intelligence. In terms of some conceptions of human intelligence as predominantly a product of cultural learning, this fact should seem surprising. . . . The sum total of the wide-ranging information provided in this book would scarcely contradict the conclusion that, as yet, investigators have not come up with dependable and replicable evidence that they have discovered a psychological method by which they can increase "intelligence" in the sense of Spearman's *g*.[9]

THUS CURRENT claims regarding the plasticity of human intelligence are notably more subdued than were the promises of only 20 years ago. Edward Zigler, one of the founders of and leaders in compensatory education, and his colleague, Winnie Berman, have recently warned that workers in the field "must be on guard never again to make the errors of overpromising and overselling the positive effects of early childhood intervention."[10]

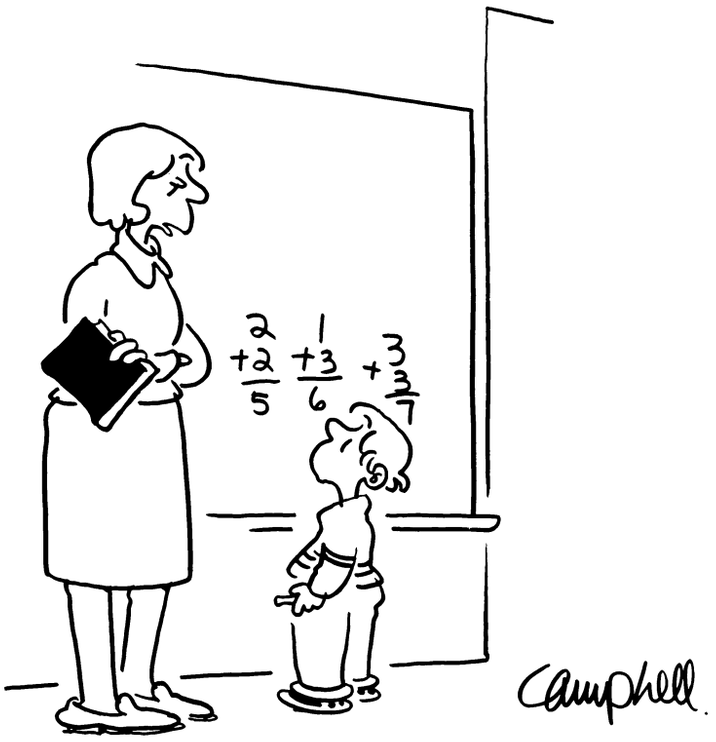
Current claims regarding the plasticity of human intelligence are notably more subdued than were the promises of only 20 years ago.

Despite their personal enthusiasm for compensatory education, Zigler and Berman have surveyed the history and developments in this field with critical objectivity. Of the beginning of preschool intervention in the 1960s, they say:

> It was widely believed that a program of early environmental enrichment would give lower SES [socioeconomic status] children the boost they needed to perform on a par with their middle SES peers. Intervention was supposed to impart immediate benefits so that class differences would be eliminated by the time of school entry. Furthermore, many expected that the brief preschool experience would be so potent a counteraction to the deficits in poor children's lives that it could prevent further attenuation in age-appropriate performance and a recurrence of the gap between social classes in later grades. . . . What we witnessed in the 1960s was the belief that intelligence quotients can be dramatically increased with minimal effort. . . . Unfortunately, "knowing more" was easily translated into "becoming smarter."[11]

Elsewhere, Zigler describes the thinking in the early days of Head Start, a program that he helped to initiate:

> . . . J. McV. Hunt, Benjamin Bloom, and others constructed for us a theoretical view that conceptualized the young child as possessing an almost unlimited degree of plasticity. Joe Hunt continued to assert that the norm of reaction for intelligence was 70 I.Q. points . . . and that relatively short-term intervention efforts could result in I.Q. gains of 49 or 63 points. With such environmental sugarplums dancing in our heads, we actually thought we could compensate for the effects of several years of impoverishment as well as inoculate the child against the future rav-



"You need to try for mainstream answers."

ages of such impoverishment, all by providing a six- or eight-week summer Head Start experience.[12]

This theoretical view of human intelligence — a view that governed the design and expectations of compensatory education programs in the 1960s — has been put to the test during the past 20 years. And the outcome seems remarkably clear. It turns out that the prevailing views of most psychologists and educators in the 1960s were largely wrong with regard to such questions as, What is the nature of intelligence? What is it that our I.Q. tests measure primarily? Why is the I.Q. so highly predictive of scholastic performance?

The error lay in believing that the disadvantage with which many poor or culturally different children entered school — and the disadvantage that compensatory education was intended to remedy — was mainly a deficiency in *knowledge*. Implicit in this belief was a view of intelligence as consisting of a general learning ability of almost unlimited plasticity plus the "knowledge contents" of memory, particularly those kinds of knowledge that serve to improve scholastic performance. Holders of this view saw the information content of I.Q. tests as an arbitrary sample of the specific items of knowledge and skill normally acquired by members of the white middle and upper classes.

In this highly behavioristic conception of intelligence, which I have elsewhere termed the *specificity doctrine*,[13] intelligence is erroneously identified with the content of the test items that psychologists have devised for assessing intelligence. These test items cover such things as general information, vocabulary, arithmetic, and the ability to copy certain geometric figures, to make block designs, and to work puzzles. To acquire the knowledge and skills to do these things — or to learn other, similar things that would have positive transfer to performance on I.Q. tests or in coursework — is to become more intelligent, according to this deceptive view of intelligence. As Zigler and Berman have put it, "knowing more" is erroneously translated into "becoming smarter."

STRIKING FINDINGS from two recent lines of research — that on test bias and that on mental chronometry — clearly contradict the view of individual and group differences in intelligence as differences primarily in knowledge.

The research on test bias has shown that the level of difficulty of I.Q. and achievement test items is consistent across all American-born, English-speaking ethnic and social-class groups. Moreover, I.Q. and achievement tests do not differ in their predictive validity for these groups. These findings are highly inconsistent with the hypothesis that cultural differences exist in the knowledge base that these tests sample. Available evidence from studies of test bias makes it extremely implausible that racial and social-class differences can be explained by cultural differences in the knowledge base or by differential opportunity for acquiring the knowledge that existing tests sample.[14] For every American-born social class and racial group, highly diverse test items maintain the same relative standing on indices of item difficulty, regardless of the culture loadings of the items. This phenomenon requires that we find some explanation for group differences on I.Q. and achievement tests other than cultural differences in exposure to the various kinds of knowledge sampled by the tests.

We must seek the explanation, I believe, at the most basic level of information processing. In recent years, both the theory and the technology of research on cognitive processes have afforded powerful means for analyzing individual and group differences in abilities. Within the framework of cognitive processes research, the kinds of questions that we can investigate are quite different and more basic than those we can study through traditional psychometric tests and factor analysis. Mental chronometry, or measurement of the time required for various mental events in the course of information processing, permits us to investigate individual differences at the level of elementary cognitive processes — those processes through which individuals attain the complex learning, knowledge, and problem-solving skills that I.Q. tests sample.

Researchers devise the tasks used to measure individual differences in various elementary cognitive processes in such a way as to rule out or greatly minimize individual differences in knowledge. These tasks are so simple, and the error rates on them are so close to zero, that individual differences can be studied only by chronometric techniques. For example, the cognitive tasks that we use in our laboratory are so easy that they typically require less than one second to perform.[15] Yet these very brief response latencies, derived from a number of elementary processing tasks, together can account for some 70% of the variance in scores on untimed standard psychometric tests of intelligence. Very little of the true score variance on such tests can be attributed to the knowledge covered by the tests' content per se.

It is important to understand that the items of standardized psychometric tests are mainly vehicles for reflecting the past and present efficiency of mental processes. That these items usually include some knowledge content is only an incidental and nonessential feature. The fact is that individual differences on these content-laden tests correlate with response latencies on elementary cognitive-processing tasks that have minimal intellectual content. This means that our standard I.Q. tests — and the scholastic achievement tests with which these I.Q. tests are highly correlated — reflect individual differences in the speed and efficiency of basic cognitive processes more than they reflect differences in the information content to which test-takers have been exposed. In fact, we can account for a substantial portion of the variance in I.Q. scores by measuring the evoked electrical potentials of the brain, using an electrode attached to the scalp — a measure that is not only free of any knowledge content but that is not even dependent on any voluntary or overt behavior by the subject.[16]

The design of compensatory education and the assessment of its effects should be informed by the recent studies on information processing.

THUS I SUGGEST that the design of compensatory education and the assessment of its effects should be informed by the recent studies on information processing. The variables that have been measured by researchers in this field to date have correlated not only with I.Q., but with scholastic achievement as well.[17] An important question for future research is, What proportions of the variance in I.Q. and in scholastic achievement are associated with elementary cognitive processes and with meta-processes respectively? A second but equally important question is, What possible effects can various types of compensatory training have on these two levels of cognitive processes?

Elementary cognitive processes include such variables as perceptual speed, stimulus scanning, stimulus encoding, mental rotation or transformation of visual stimuli, short-term memory capacity, efficien-

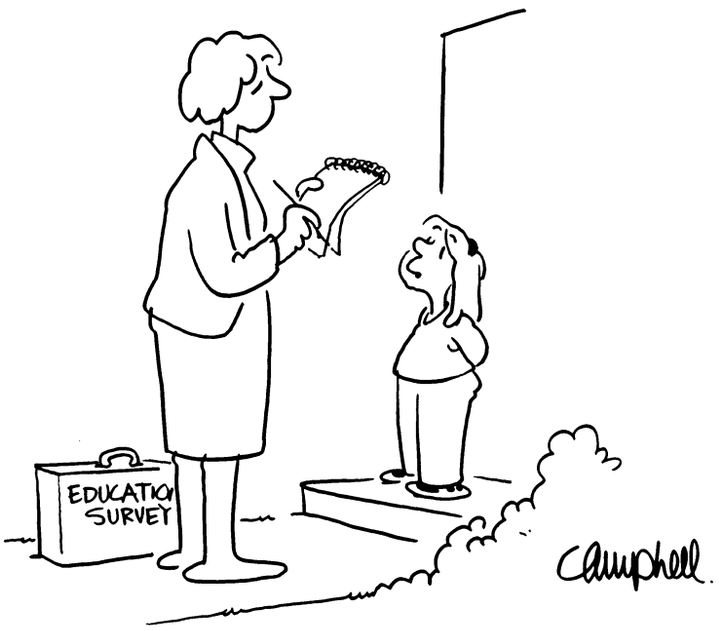
As yet, virtually nothing is known about the effects of compensatory education on the various levels of cognitive processing.

cy of information retrieval from long-term memory, generalization, discrimination, comparison, transfer, and response execution. *Meta-processes* include those planning and executive functions that select and coordinate the deployment of the elementary cognitive processes to handle specific situations, e.g., strategies for problem recognition, for selecting and combining lower-order cognitive processes, for organizing information, for allocating time and resources, for monitoring one's own performance, and the like.

Meta-processes are thought to be more amenable than elementary processes to improvement through training, but no solid evidence currently exists on this question. And, though much is already known about social-class and racial-group differences in I.Q. and scholastic achievement, psychologists have scarcely begun to try to understand the nature and locus of these differences in terms of the cognitive processes and meta-processes involved.[18] As yet, virtually nothing is known about the effects of compensatory education on the various levels of cognitive processing or about the extent to which the levels of cognitive processing can be influenced by training especially designed for that purpose.

I suspect that a substantial part of the individual variance in I.Q. and scholastic achievement – probably somewhere between 50% and 70%, according to the best evidence on the heritability of I.Q. – is not subject to manipulation by any strictly psychological or educational treatment. The reason for this, I assume, is that the main locus of control of that unyielding source of variance is more biological than psychological or behavioral.

At an even more fundamental level, we might ask why variance in intelligence should be so surprisingly resistant to experimental manipulation. As I have suggested elsewhere,[19] this apparent resistance to manipulation seems less surprising if we view human intelligence as an outcome of biological evolution. Genetic variation is the one absolutely essential ingredient to enable evolution to occur. If intelligence has evolved as a fitness characteristic in the Darwinian sense – that is, as an instrumentality for the survival of humankind – it is conceivable that the biological basis of intelligence has a built-in stabilizing mechanism, rather like a gyroscope, that safeguards the individual's behavioral capacity for coping with the exigencies of survival. If that were the case, mental development would not be wholly at the mercy of often erratic environmental happenstance. A too-malleable fitness trait would afford an organism too little protection against the vagaries of its environment. Thus, as humanity evolved, processes may also have evolved to buffer human intelligence from being pushed too far in one direction or another, whether by adventitiously harmful or by intentionally benevolent environmental forces.



"My mom's an Ed.D., my dad's an M.D./Ph.D., and I'm a CBS, ABC, NBC."

1. Barbara Lerner, "Test Scores as Measures of Human Capital," in Raymond B. Cattell, ed., *Intelligence and National Achievement* (Washington, D.C.: Cliveden Press, 1983).
2. Ibid., p. 73.
3. Ibid., p. 74.
4. Consortium for Longitudinal Studies, *As the Twig Is Bent . . . Lasting Effects of Preschool Programs* (Hillsdale, N.J.: Erlbaum, 1983); and Edward Zigler and Jeanette Valentine, *Project Head Start* (New York: Free Press, 1979).
5. Gene V Glass and Mary C. Ellwein, review of *As the Twig Is Bent . . .*, by the Consortium for Longitudinal Studies, in *Science*, 20 January 1984, p. 274.
6. Craig T. Ramey et al., "The Carolina Abecedarian Project: A Longitudinal and Multidisciplinary Approach to the Prevention of Developmental Retardation," in Theodore D. Tjossem, ed., *Intervention Strategies for High Risk Infants and Young Children* (Baltimore: University Park Press, 1976).
7. Craig T. Ramey and Ron Haskins, "The Modification of Intelligence Through Early Experience," *Intelligence*, January/March 1981, pp. 5-19; and Arthur R. Jensen, "Raising the I.Q.: The Ramey and Haskins Study," *Intelligence*, January/March 1981, pp. 29-40.
8. Douglas K. Detterman and Robert J. Sternberg, eds., *How and How Much Can Intelligence Be Increased* (Norwood, N.J.: Ablex, 1982).
9. Arthur R. Jensen, "Again, How Much Can We Boost I.Q.?," review of *How and How Much Can Intelligence Be Increased*, edited by Douglas K. Detterman and Robert J. Sternberg, in *Contemporary Psychology*, October 1983, p. 757.
10. Edward Zigler and Winnie Berman, "Discerning the Future of Early Childhood Intervention," *American Psychologist*, August 1983, p. 897.
11. Ibid., pp. 895-96.
12. Quoted in Peter Skerry, "The Charmed Life of Head Start," *Public Interest*, Fall 1983, pp. 18-39.
13. Arthur R. Jensen, "Test Validity: *g* Versus the Specificity Doctrine," *Journal of Social and Biological Structures*, vol. 7, 1984, pp. 93-118.
14. Arthur R. Jensen, *Bias in Mental Testing* (New York: Free Press, 1980); and Cecil R. Reynolds and Robert T. Brown, *Perspectives on Bias in Mental Testing* (New York: Plenum, 1984).
15. Arthur R. Jensen, "Chronometric Analysis of Intelligence," *Journal of Social and Biological Structures*, April 1980, pp. 103-22; idem, "The Chronometry of Intelligence," in Robert J. Sternberg, ed., *Advances in the Psychology of Human Intelligence* (Hillsdale, N.J.: Erlbaum, 1982); and idem, "Reaction Time and Psychometric *g*," in Hans J. Eysenck, ed., *A Model for Intelligence* (Heidelberg: Springer-Verlag, 1982).
16. Donna E. Hendrickson and Alan E. Hendrickson, "The Biological Basis of Individual Differences in Intelligence," *Personality and Individual Differences*, January 1980, pp. 3-34.
17. Jerry S. Carlson and C. Mark Jensen, "Reaction Time, Movement Time, and Intelligence: A Replication and Extension," *Intelligence*, July/September 1982, pp. 265-74.
18. John G. Borkowski and Audrey Krause, "Racial Differences in Intelligence: The Importance of the Executive System," *Intelligence*, October/December 1983, pp. 379-95; Arthur R. Jensen, "Race Differences and Type II Errors: A Comment on Borkowski and Krause," *Intelligence*, in press; and Philip A. Vernon and Arthur R. Jensen, "Individual and Group Differences in Intelligence and Speed of Information Processing," *Personality and Individual Differences*, in press.
19. Jensen, "Again, How Much Can We Boost I.Q.?," p. 758. □