

Five Decades of Public Controversy Over Mental Testing

LEE J. CRONBACH *Stanford University*

As the United States mobilized for war in the spring of 1917, Professor Terman filled a briefcase with materials on the group intelligence test his student Arthur Otis had just designed, and went East to meet with other leading psychologists of the time. Within weeks they had organized the Army Alpha Examination for use in testing recruits. Their short mental test did indeed locate men who made satisfactory officers and noncoms. Delighted with this achievement, the psychologists then pressed for civilian testing.

"Teachers must learn to use tests," said Terman (1919). Otherwise, "the universal grading of children according to mental ability must remain a Utopian dream" (p. 291). Terman's individual test published in 1916 already was widely used, and immediately after the War the group tests were snatched up by school systems and colleges as a basis for pupil classification, guidance, and college admissions. Within 30 months of the first publication of a group test, some four million children had been tested. The test technology became an accepted, and increasingly influential, feature of American life. The momentum of the tests overrode all criticism.

Today, however, the critics are in the ascendant. The California Legislature, for example, has twice voted to prohibit group mental testing in schools, on the grounds that their effect is to limit the edu-

cation black children receive. (The Republican Governor vetoed the bill, which Republican legislators had opposed.) On many fronts, a procedure that came in as an impartial application of scientific findings about talent is now under bitter attack. The irony is all the greater, in that the attack comes largely from those who speak for the poor. Proponents of testing, from Thomas Jefferson onward, have wanted to open doors for the talented poor, in a system in which doors often are opened by parental wealth and status.

Public controversy deals in stereotypes, never in subtleties. The Luddites can smash up a device, but to improve a system requires calm study. So, following the advice of a professional study committee, the same California Legislature that voted to outlaw group testing by local districts instituted a new, carefully safeguarded, statewide test of mental ability for first graders to provide data needed as a baseline for evaluation of reading instruction that the Legislature desires. On advice of the same committee they canceled a statewide mental test in Grades 6 and 10 that had no proper function. Sound policy is not for tests or against tests; what matters is how tests are used. But all the public hears is endless angry clamor from extremists who see no good in any test, or no evil.

The topic of this article is not policies about testing. Rather, the concern is with the difficulties that arise when the scholar enters the arena of policy. The controversies about testing are an instructive example.

The able psychologists who wrote popular articles about individual differences set out to tell the public what they knew so that it could be weighed into policies. Human differences, however, are an explosive theme. The most dramatic fulminations were set off by Arthur Jensen's (1969) article, and his story can be our starting place.

The Jensen incident is often referred to as if his article were an isolated effort to promote an idiosyncratic theory. Few lay readers realize that

This article was prepared for an American Academy of Arts and Sciences study entitled "Social Science Controversies and Public Policy Decisions" which was financed by a grant from the Russell Sage Foundation. The article is preprinted by permission of the American Academy of Arts and Sciences. The work leading to this article was begun while I was a Fellow of the John Simon Guggenheim Memorial Foundation and a Visiting Scholar at the Center for Advanced Study in the Behavioral Sciences. Paul Chapman has been of great help in preparing this article. I thank many of the persons mentioned in the history for comments and information.

Requests for reprints should be sent to Lee J. Cronbach, School of Education, Stanford University, Stanford, California 94305.

controversy over these issues has waxed and waned since test reports began in 1905. Views that arouse heated rejoinders in one climate of opinion have been accepted matter-of-factly in another. For perspective, the Jensen affair must be seen alongside other controversies. Particularly important to the story is the role of the media in shaping controversies and, at times, in neglecting an opportunity for controversy. Nor are the media to be seen as independent agents; what catches fire depends, at any time, on the way public opinion is blowing.

Testing in the Public Eye: 1967-1974

THE POLITICAL AND SCIENTIFIC SETTING

To set the stage for Jensen, we must go back two decades. World War II dramatized the value of specialized training. After the War, talent was cherished as a resource. Specialized and higher education expanded rapidly, "merit scholarships" became available, and schools emphasized career guidance as never before. Even as college education became more widely available, entrance to prestige colleges became more competitive. Draft exemption, too, came to depend on test scores. The more the tests came to determine life chances, the more they became foci of anxiety. Objections to elitism and meritocracy followed.

Around 1960, traditional concepts were challenged in the newly vigorous field of child psychology. The new ideas, coming from studies of child learning and colored by Piaget's views on the role of experience in intellectual development, were strenuously "environmentalist." The importance of the genes and the biological substrate was not denied, but the emphasis was on the use the child made of experience.

If logical powers are constructed out of experience, it becomes natural to decry "the old belief in fixed intelligence" and to search for methods of stimulating intellectual growth. This trend within psychology crested just as the condition of blacks became a national concern. The Head Start program for the children of the poor was adopted on the warrant of demonstrations that such activities on a small scale had benefited children. Evidence that the hasty large-scale programs had not produced the promised miracles was just beginning to surface in 1969.

Racial differences in ability had rarely been examined by scholars in the preceding two decades.

A few psychologists had assembled the studies showing that blacks average lower on tests than whites, and these psychologists had tried to interpret the finding biologically rather than environmentally, but their work was considered irrelevant if not disreputable. Criticisms of the early work on the problem had convinced nearly all social scientists that no research design can disentangle genetic and environmental components of group differences, and the view prevailing after 1940 was close to these statements of Otto Klineberg (1952):

The available data offer no support for the view that racial or national origins set different limits to the potentialities of a child. The teacher has the right to assume that under similar conditions both the range of capacities and the average capacity of various groups will almost certainly be about the same. He has the duty, therefore, to treat each child as an individual. . . . (p. 953)

The middle sentence, however, slides from the view that the question of innate group differences is unanswerable to the "assumption" that no differences exist. What was an assumption in the 1940s had crystallized into a combative assertion in the 1960s.

A few more words of preamble, and we can turn to Jensen. Much of the dispute centers around the contention that a child's standing in ability depends largely on what genes he inherits. The evidence is that differences among American or British whites in the past generation have been due in part to genetic differences. The precise proportion, but not the principle, of hereditary influence can be debated. Yet the statistic is an index of sociocultural conditions, not a biological inevitability. Change the distribution of nutrition, home experience, and schooling in the next generation, and the heritability index will change. Findings on heritability within white populations tell us nothing whatsoever about how white and black groups would compare if their environments had been equalized. Note also that a high degree of heritability does not imply that improved environment can have no effect. Even if the heritability index is as high as .80, two children with the same genotype may differ by as much as 25 IQ points if one is reared in a superior environment and the other in an unstimulating one.

JENSEN: SCHOLARSHIP, CONFUSION, INFERRED ADVOCACY

Arthur Jensen, a well-recognized educational psychologist, and the geneticist Ernst Caspari were

invited to present companion lectures to the 1967 meeting of the American Educational Research Association. Caspari dealt with biological questions, while Jensen crossed over into social class and educational policy. The Jensen (1968) and Caspari (1968) articles agreed wherever they touched the same subject, and caused no controversy. Although Jensen dealt with heritability, he emphasized his own research on the relations among memory, IQ, and school performance in different social groups. At the end, he recommended ways to teach pupils having adequate memory but mediocre IQs. The one paragraph of Jensen's article that foreshadowed later events said only this:

I find little information about the extent to which Negro-white differences have a genetic basis Therefore, statements . . . can at present be nothing but conjecture and surmise But the question arises whether there has been an official decision to create the impression that such hypotheses have already been scientifically tested with conclusive results. (pp. 22-23)

He cited particular U.S. Government releases asserting a conclusion about biological equality that had been for Klineberg and other scholars no more than a working assumption.

During the same year, Jensen (1967) produced a noteworthy but little-known paper on compensatory education which entirely denied the view that came to be known as "jensenism." On comparative education, he said: "Action programs are obviously needed immediately" (p. 4). On race (a single page in the article):

[T]he fact that Negroes and Mexicans are disproportionately represented in the lower [classes] . . . cannot be interpreted as evidence of poor genetic potential. . . . [I]t seems a reasonable hypothesis that their low-average IQ is due to environmental rather than genetic factors. (p. 10)

Jensen's well-known 1969 article arose out of the journalistic enterprise of the student editors of the *Harvard Educational Review*, abetted by his missionary zeal. Initially, an editor asked Jensen to submit the manuscript of a 1968 talk on "learning ability and socioeconomic status." Jensen meanwhile had drafted another lecture arguing that IQs do not respond to educational intervention although they might be changed by modifying prenatal care. He provided both manuscripts to the *Harvard Review*, and after studying them the editors, in April 1968, asked Jensen to work up a complete statement of his views for their 12,000 readers (Jensen, 1972). Having outlined controversial con-

tent to be covered, the editors arranged for a panel to provide comments for publication alongside Jensen's piece. Jensen was to cover, in the context of intervention programs, the contribution of heredity and environment, "his own position" on social class and racial differences, and his research on learning. Jensen, even as you and I, let other work crowd his schedule until mid-September, and then put together 50,000 words in two months.

The elaborations and asides added in Jensen's hasty completion of the manuscript damaged the balance and clarity of the final version. The distribution of emphasis was not much like that of the 1967 and 1968 articles and seems not to follow closely the weighting suggested by the *Harvard Educational Review* editors. Even in this paper, race took up less than 10% of the space; but race somehow looms large because of the interest of blacks in compensatory education. The 1969 title came from the second 1968 speech, "How much can we boost IQ and scholastic achievement?" The 1969 reader would conclude that the answer was "Not at all; neither of them," whereas the 1968 listener had been urged to concentrate on improving achievement *rather than* IQ, a very different answer. Jensen's own research led him to the optimistic view that an alternative form of teaching could benefit the poor child whose IQ is low. Only the final 2 pages of the 120-page paper in the *Harvard Review* offered this positive approach to compensatory education. The message was a quaver compared to the stentorian opening: "Why has there been such uniform failure of compensatory programs wherever they have been tried?" (p. 3).

As soon as the article was in type the publicity broke. The *Harvard Review* made the article available to the press along with the remarks of the prearranged critics. Substantial excerpts appeared in *U.S. News* and the *New York Times*, and lesser accounts appeared in other media. Within two weeks, the Students for a Democratic Society were cruising the Berkeley campus with a sound truck whose chant was "Stop racism. Fire Jensen!" And on the Eastern seaboard, it was rumored that the Nixon cabinet had discussed whether the article could be used to justify reducing outlays to aid blacks. In denying the rumor Moynihan (1970) commented, "I know what Jensen is going through. I got the same treatment for almost exactly the opposite hypothesis" (p. 192). (A few years earlier, Moynihan had suggested that many black

difficulties in later life arise out of the conditions of child rearing.)

Style and substance of the controversy. There was no precise statement of a "position" in Jensen's sprawling and allusive argument. He defended at length the long-standing conclusion that a large fraction of the differences within the white population can be attributed to inheritance, and he acknowledged that differences between groups have some environmental sources:

But the possible importance of genetic factors in racial behavioral differences has been greatly ignored, almost to the point of being a tabooed subject, just as were the topics of venereal disease and birth control a generation or so ago. (Jensen, 1969, p. 80)

The statement that most nearly asserts racial genetic disadvantage is this:

[T]he discrepancy in . . . average performance [of the disadvantaged, compared to the norm] cannot be completely or directly attributed to discrimination or inequalities in education. It seems not unreasonable . . . to hypothesize that genetic factors may play a part. . . . The preponderance of the evidence is, in my opinion, less consistent with a strictly environmental hypothesis than with a genetic hypothesis, which, of course, does not exclude the influence of environment. . . . (Jensen, 1969, p. 82)

The language is equally moderate throughout the article, but the massive and one-sided presentation was read as an assertion.

The news media were not able to weigh matters as delicately as Jensen had. Fairly typical is *Newsweek's* ("Born dumb," 1969) summary: "Dr. Jensen's view, put simply, is that most blacks are born with less 'intelligence' than most whites" (p. 84). Journalists who had more space at their disposal did a remarkably good job of touching on the range of issues, though it was the epigrammatic or emphatic statements by the critics that were quoted.

It is impossible to summarize what was said. The first wave of solicited replies ran to the same length, in total, as Jensen's article. Further responses in the summer issue of the *Harvard Educational Review* ran to another 150 pages. The tone ranged from condemnation to applause, from polemic to technical analysis. Jensen made full use of his opportunity to reply to the first wave of criticism. Similarly, a *New York Times Magazine* profile of Jensen's personality and theory drew page upon page of correspondence. Jensen had the opportunity to set forth, letter by letter, just what he agreed or disagreed with.

One display of the academic community in action took place at Berkeley. In a spring when the radi-

cals were smiting the Establishment with any cudgel, no one took seriously the demands for Jensen's scalp. There was an attempted invasion of his classes and harassment of a research assistant that drew an editorial frown from the far-away *New York Times*, but most of the agitation was confined to the campus newspaper. A few faculty members opposed to Jensen's views decided that a public debate was needed to set the campus straight. The Administration assumed control and saw fair play, setting up a symposium before an audience limited to an equal number of observers chosen by each side. The three-hour proceedings were videotaped and shown to an evening meeting which drew a serious audience of several hundred, large even for Berkeley, on a spring evening just before finals.

Jensen made a brief, considered, and articulate statement. He spoke of genetic race differences as merely a hypothesis meriting investigation, saying nothing that can be read as even a tentative judgment that the gene pool of blacks is somehow inferior. Two geneticists, two sociologists, and an educational psychologist replied. Whatever matters were capable of being posed in intellectual and disciplined terms were so posed and debated for the Berkeley students. While panel members hit hard at various points, they did not invoke ideology or passion.

Meanwhile, there had been another type of formal reaction from academe. The executive board of the Society for the Psychological Study of Social Issues (SPSSI) was composed of 18 leading psychologists noted for their liberal social concerns. Several of them were deep into civil-rights and compensatory-education activities. Six weeks after the news broke they issued a measured statement (SPSSI, 1969) denying that there is any technique for investigating innate racial differences under current conditions, asserting that the failure of compensatory education was in the planning and scope and not in the idea, and decrying heredity versus environment statements as oversimplified. The American Anthropological Association made a less restrained attack, adopting an anti-Jensen resolution that even its writers considered to be ad hominem (see Jensen, 1972, p. 38).

Jensen was right about the failure of compensatory efforts, inasmuch as even now we have no compensatory method, reproducible on a large scale, of demonstrated value. Jensen's article seems to say that efforts toward compensatory education ought to be dropped, overdramatizing the failure.

But buried within the paper was recognition that intensive small-scale programs often succeed, and in actuality Jensen's position was a call for invention of effective educational procedures.

Heritability was taken up by Jensen as if it has some bearing on compensatory policies. But even the most hereditarian position does not hold that ranks in performance will remain stable when the initially low-ranking children are treated specially. Jensen let this point drop from sight. Jensen (1972) wrote strongly on heritability because he thought that other psychologists had spread a false cliché:

Speakers and writers on intelligence, mental retardation, cultural disadvantage, and the like, . . . state, often with an evident sense of virtue and relief, that modern psychology has overthrown the "belief in fixed intelligence." (p. 11)

Jensen did not manage to restore the balance. The conventional results that Jensen reiterated were taken by readers to be a new and controversial doctrine.

Jensen was also attempting to balance talk about race. He rightly protested the propaganda that had science proving races indistinguishable with respect to any psychoactive genes. It seems that Jensen intended to speak for openmindedness and continued research on racial characteristics. But as the spotlight excited him, he became more assertive. In the *New York Times Magazine*, Edson (1969) offered this fantastic quote:

There are no "black" genes or "white" genes; there are intelligence genes which are found in populations in different proportions, somewhat like the distribution of blood types. The number of intelligence genes seems to be lower, over-all, in the black population than in the white. (p. 43)

This apparently is a garbled version of a statement made by telephone, but Jensen did not set the record straight when he could have.

For many, the distressing part of the Jensen affair was that it misdirected attention. Consider the ambivalence of Lederberg (1969) in his newspaper column on science:

[T]he article is a thoughtful review [P]opular commentaries . . . have emphasized a few controversial [and I would say incautious] remarks at the expense of a great deal of Dr. Jensen's wisdom and scholarly reserve.

The genetic hypothesis is almost irrelevant to Jensen's most cogent point [about adapting instruction] If a 6-year old has a deficit in abstract thinking, it is relatively unimportant for educational policy whether this is the fault of the genes or a cultural maladaptation. . . . The genetic hypothesis does matter if it discourages educators and scientists from probing more deeply into the crucial early years. (p. 611)

Lederberg remarked on malnutrition and lead poisoning in the ghettos, cited a study in which poor children whose mothers had received dietary supplements during pregnancy averaged 8 IQ points above controls, and ended, "With effects like that, why are we discussing anything else?"

The controversy reached a wide audience and Jensen's name is today one that the informed citizen is expected to recognize. But outside the professional literature the dispute got little space. The usual magazine editor covered the controversy just once; *The Reader's Guide to Periodical Literature* lists just 12 articles in 1969 and only 2 in 1970. To be sure, polemicists continue to seek and occasionally find space in print, some of them decrying attention to human differences as inhumane and others accusing the equalitarians of Lysenkoism.

THE *Atlantic Monthly* PROVIDES A LITTLE WAR

The Jensen episode had a peculiar echo. Richard Herrnstein (1971) offered *Atlantic Monthly* an article discussing IQs, squarely within the tradition of the magazine's essays on scientific themes. The editor accepted the essay provisionally but asked Herrnstein to make it livelier and to bring in questions of policy. By mutual agreement, the topic of race differences was inserted at this time. In Herrnstein's final 16 pages (mostly on long-established psychological generalizations), less than 1½ pages touched on racial differences. Every paragraph of the editors' 1,200-word introduction, however, brought in the racial theme. This filter distorted Herrnstein's communication and in a sense politicized a paper that had originally been almost textbookish.

Herrnstein spelled out once again the facts indicating that IQ differences within the white population, under present social conditions, have a large hereditary component, adding that such estimates of heritability "say nothing" about differences between groups. Then Herrnstein offended some sensitive souls with one sentence stating that possible genetic differences affecting average IQs of groups should be open to investigation.

Much louder criticism was stirred up by Herrnstein's parsing of a simple syllogism: Make conditions wholesome and uniform for everyone, and the power of heredity will in the course of generations stratify any achievement-based society so that a child remains in his father's status. Herrnstein did not suggest—as many readers thought—that

the offspring of the blacks or of the present poor are doomed to form the low caste of the future. His position was simply that if there is no variance in environment, nothing save heredity (and luck) can influence status.

The Jensen drama was replayed on a small stage. The editors supplied copies of the finished article to many academics and invited publishable letters. Two issues later, the *Atlantic Monthly* carried seven pages of letters, some condemning Herrnstein and some endorsing his facts and logic. The Harvard radicals mobilized to demand Herrnstein's dismissal. Herrnstein found himself challenged to debate about a position that seems as guarded as a statement about human affairs can be. The range of complaints was confusingly broad, but there was, of course, no logical rejoinder to Herrnstein's syllogism.

The only fault of the Herrnstein paper is its silence on matters of social philosophy (see Wolff, 1971). It is assumed that those social contributions that derive from mental ability are the proper basis for distributing rewards and respect. But the article is not factually or logically deficient, and it is not at all racist.

The Herrnstein affair did not live beyond three issues of the *Atlantic Monthly* and received only the briefest of attention elsewhere. Agitation on the Harvard campus continued for a year, fomented by radicals needing a target. At base, the Herrnstein affair was a journalistic stunt, in the tradition of Hearst to Remington: "You provide the pictures, I'll provide the war." The *Atlantic Monthly* provided no more than a pop-gun fusillade, but liveliness sells papers.

PROPOSALS TO RESTRICT INQUIRY

Whereas the social scientist has insisted upon his right and duty to inquire into all human affairs, the community is increasingly unsympathetic. It has become difficult for white social scientists to work in the black community; black distrust was surely increased by the Jensen affair.

In the America of 1969, to make a statement about race differences even at the level of hypothesis was to offend blacks and threaten their political interests (Rowan, 1970). Many laymen and scholars condemned Jensen not for false impressions he might have given but for making *any* statement about race. See one letter to the *New York Times Magazine* (Greenwald, 1969): "The wise scientist will not devote himself to research

on the relation between race and ability; the wise university will not honor those who do, or disseminate their work" (p. 4).

Jensen's (1970) stance was and is that society should encourage the interested scholar to investigate *any* hypothesis:

[U]nnecessary difficulties arise when we allow the scientific question to become mixed up with the social-political aspect of the problem, for when it does we are less able to think clearly about either set of questions. The question of whether there are or are not genetical racial differences in intelligence is independent of any questions of its implications, whatever they may be. (p. 149)

The root of Jensen's martyrdom is that word "Unnecessary."

Law professor J. D. Hyman (1969) spoke for the other side:

Sensitivity to the social problems of our day requires a clear showing of the high probability of their truth before hypotheses are advanced which reinforce the stereotypes on which our caste system has been built. (p. 31)

One infers that the social scientist with a disturbing hypothesis should pursue it privately, keeping his dark suspicions secret until he has a solid case. Given the social nature of the scientific enterprise, this seems as inhospitable to heterodoxy as an outright embargo on a research topic.

PYGMALION IN THE HEADLINES

After Jensen, the work on testing that received most attention from the press was *Pygmalion in the Classroom* (Rosenthal & Jacobsen, 1968). This work too was a matter of heated controversy, but the controversy was wholly concealed from public view. We had better wonder why.

Robert Rosenthal had ingeniously demonstrated that the psychologist experimenting with animals tends, all unconsciously, to make his results favor the hypothesis he initially tended (or wanted) to believe. Perhaps he makes nonrandom errors of observations, or perhaps he preconditions subjects differently. In time, Rosenthal launched a study to demonstrate that similar effects occur in the school when teachers "know what to expect." Teachers in a California school were handed a list of pupils who supposedly had shown exceptional promise on a test but who in fact had been picked at random. Months later, we are told, these children had progressed more than their unheralded classmates. The teacher supposedly created talent by his own ministrations to fulfill the prophecy.

In my view, *Pygmalion in the Classroom* merits no consideration as research. The "experimental

manipulation" of teacher belief was unbelievably casual—one sheet of paper added to the teacher's in-basket, which apparently moved within seconds to the wastebasket. The technical reviews indicate that the advertised gains of the "magic" children were an artifact of crude experimental design and improper statistical analysis. (No doubt there are expectancy effects in the classroom. The question ought to be whether tests add to bias or instead bring expectations closer to the truth. On that there is no direct evidence.)

The *Pygmalion* study was carried out from 1964 to 1966 and was given a preliminary report in a few pages of a scholarly book on experimenter effects, followed by fuller scholarly reports and news stories. By 1968 Rosenthal had produced *Pygmalion in the Classroom* (with Jacobsen), plus articles in *Psychology Today* and *Scientific American*. The publisher of the book sent the manuscript to referees; the opinions ranged from high praise to sharp criticism. The decision to issue the book was not unreasonable, but one might have expected the finished book to at least acknowledge that the methodology was in dispute. Worse, while scores of pages were given to the desired "findings" from California, fewer than 500 words were used to report a Massachusetts study by Rosenthal in which a "significant" difference favored the control (1) group.

We need not trace here the controversy within the profession or spell out the objections. Much of the professional debate is conveniently collected in the Elashoff and Snow (1971) volume, a project of the National Society of the Study of Education. The vehemence of the criticism is only incompletely suggested by this remark from R. L. Thorndike's review (there reprinted): *Pygmalion* "is so defective technically that one can only regret that it ever got beyond the eyes of the original investigators!" (Elsahoff & Snow, 1971, p. 65).

Now what was the public being told? They were told nothing about the controversy but heard much about the study as evidence that mental tests are doing harm. It is significant that the media ignored the loud but cloistered battle over the quality of the evidence. The *New York Times* gave Rosenthal's story sympathetic frontpage play on August 8, 1967, when the first book appeared. Out of 30 column inches, 1 inch on the carry-over page mentioned that studies in Massachusetts and Ohio "have not clearly corroborated" the California findings. Nothing was said about evidence in the opposite direction; indeed the next paragraph re-

turns to restating that teacher expectations are a powerful determinant.

The later book got more public notice. Robert Coles (1969), retelling the story for the *New Yorker*, went out of his way to praise what the book said about the "ethical dimensions of scientific work." If Coles had heard of doubts about the scientific dimensions of the work, no hint appeared. In the *New York Review of Books* Herbert Kohl (1968) was equally sure of the conclusions, but he found the Rosenthal work defective ethically because the experimenters had "assumed god-like" roles and acquired knowledge "through deceit and bad faith" (p. 31). Coles's (1969, p. 177) asides—about "experts, the secular gods of the twentieth century," "the effect a pair of scientists armed with tons of paper can have," and the "dense, muddled language" of theorists—are characteristic of an anti-intellectual tone present in nearly all of the recent attacks on testing and research into human differences. To be sure, some social scientists invite such criticism by writing as if their numerical analyses are sufficient to settle issues of social policy.

Earlier Challenges to Mental Testing

My thesis that the hearing given to social evidence and issues depends on the times and the audience is best demonstrated by stepping back to decades when the audience was attuned differently. The most useful examples are the arguments of Allison Davis in the late 1940s and the attacks of Lippmann and Bagley in the 1920s.

SOCIAL-CLASS BIAS: A CHARGE FAILS TO IGNITE

From 1945 to 1953, Davis, a sociologist, contended that existing tests underestimate the abilities of children of the working class. His challenge was not accepted by the profession at the time, and, despite its significance, was ignored by the public. The basic scholarly document combined the doctoral study of Kenneth Eells with short essays written by Davis and others to place the results of Eells in a policy context (Eells, Davis, Havighurst, Herrick, & Tyler, 1951). Eells compared children from higher and lower classes and found the latter weaker on nearly all test items. But verbal items, especially those referring to uncommon words and objects, magnified the group difference. Poor children would rank higher, said Davis, if the test were limited to "fair" items. He tried out tests

of this kind and then reported that according to such tests "hereditary ability" is about the same in all socioeconomic groups.

Davis and his colleagues sought attention for their thesis by means of numerous semipopular presentations to lay and professional groups and in articles in *Scientific Monthly*, the *Journal of the National Education Association*, etc. Their argument was not given much attention by educators. Benjamin Fine, the education writer for the *New York Times*, did run two pieces about a Davis speech to school superintendents. Six months later, Fine (1949) included Davis's work in a pro-and-con piece on tests for the *New York Times Magazine*. Fine concluded only that tests should be used sensibly. Mild as this essay was, it provoked outraged howls from psychologists; but the published letters, while countering other criticisms of tests that Fine had mentioned, ignored the Davis material. Likewise, Davis's own popular articles and speeches were left unanswered, hence the challenge withered.

Such statements as "Half the ability in this country goes down the drain because of the failure of intelligence tests . . ." (Davis, 1949, p. 294) were not ignored *within* the testing profession. Professional symposia featured attacks on the new Davis tests. Although Davis expected his tests to eliminate class differences in IQ and found this in his initial studies, other investigators found as much social-class difference on the Davis tests as on traditional ones. Moreover, his tests were not good predictors of success in school, and so they were dismissed. Davis and his group were so discouraged that they allowed striking results they obtained on a "fair" individual test (Hess, 1955) to drop from sight.

As the argument developed, the central policy issue was misperceived. Davis's challenge was actually not to tests but to traditional schooling, as some material in the Eells volume shows. Traditional tests predict who will succeed in a schooling that makes use of abstract ideas and experiences the middle-class child is likely to have. At base, the Davis argument was that there are reasoning abilities in the lower class that schooling could capitalize upon if it were redesigned to be less verbal and less culture-laden. (This is analogous to Jensen's view that there are memory abilities in the lower class on which schooling could rely.) The Davis campaign failed for several reasons. He challenged the testers when they were in public favor. He concerned himself with "persons of low status," and there were no militant voices to take

up that cause. And if it were true that his new tests identified potentially able children for whom someone ought to invent better schooling, that advice was too abstract for public debate or action.

MORONS, IMMIGRANTS, AND EFFICIENCY EXPERTS

A principal tenet of the Progressive movement at the turn of the century was the power of social science to redirect and reshape society. Darwin and Comte, between them, had made the point that man and his institutions are subject to scientific, iconoclastic analysis. Within the Darwinist camp, such men as T. H. Huxley and L. F. Ward read the evolutionary message as a call for social activism to redesign institutions so that men would prosper in them. The American reformers expected factual analyses to free society from ills that ranged from political corruption to prostitution, from despoliation of the environment to child labor. The social survey took its place alongside the journalistic exposé as a way of determining what should be set right and how. Efficiency and scientific management like that of industry were to be brought to social institutions (Callahan, 1962).

In education, few studies had more influence than Ayres's (1909) *Laggards in the Schools*, which counted how many children were below the grade normal for their age and tallied up the economic and social costs of such retardation. At the same time, the child-study movement was deploring the emotional costs of a Procrustean school system. With child labor going out, the junior high school and vocational education coming in, and high school enrollment booming, superintendents were greatly concerned about coping with the range of abilities. The mental test, fresh from its triumphs in the Army, promised to sort out pupils who should move fast, those who should move slow, those who should go to college, and those who should not. The testers and their audience were sympathetic with the bright child forced to poke along at the average pace and with the dull child pressed to keep up. The tests would allow grouping by ability and would pick out the talented children for special encouragement.

Virtually everyone favored testing in schools; the controversies arose because of incautious interpretations made by the testers and, even more, by popular writers. Debate was touched off by the official memoir on the Army data, specifically by the 2 pages (out of 800) describing the distribution of soldiers' scores. Scores had been converted to a

mental-age scale, and after reporting the average score of 13 for the white draft the writers added two sentences by way of popularization: "A moron has been defined as anyone with a mental age from 7 to 12 years. [By] this definition . . . almost half of the white draft would have been morons" (Yerkes, 1921, p. 789). The comparability of the sample to 13-year-olds need not have been too surprising, at a time when most adults had left school by age 14. But the theoretical standard of maturity had been taken as 16 years of mental age, which argued that the typical soldier was markedly deficient.

Popular writers, especially those associated with the eugenics movement, drew broad elitist conclusions from this "finding." Popenoe (1921), for example, wrote, "Can we hope to have a successful democracy where the average mentality is 13?" (p. 233), and he went on to quote Goddard's reading of the data: "Obviously there are enough people of high intelligence to guide the Ship of State, if they are put in command. The disturbing fear is that the masses . . . will take matters into their own hands" (Goddard, 1920, p. 97). Albert Wiggam (1922), a popularizer of science, wrote a deliberately outrageous piece for the *Century Illustrated Magazine*. He declared that efforts to improve standards of living and education are folly because they allow weak elements in the genetic pool to survive, that men are born equal is "a great sentimental nebulosity," that social classes are "ordained by nature," and that "slum-people make the slums" (p. 645). In the clever talk of the twenties, "moron" claimed as large a place as "Babbitt." As Vernon Parrington (1930) mourned, "morons . . . jar one's faith in human perfectibility. In the light of intelligence tests perhaps the whole romantic theory of democracy was only a will-o'-the-wisp . . ." (p. xxviii).

Racism, directed especially against immigrants from Southern and Eastern Europe, was active, and the Army data provided ammunition for it. The psychologist C. C. Brigham (1923) was persuaded by one of the advocates of racial purity to rework the Army data on ethnic groups into a book. Brigham did warn that the Italians, for example, recently come to this country were probably not representative of the population of Italy; but his attempt to confine his conclusions to immigrants was not made prominent, and it escaped the attention of most persons who cited or attacked his work. Likewise, Brigham acknowledged that tests were not pure measures of innate ability, but he was

sure that innate racial differences had been proved to exist. So closed was his mind that in what purported to be the definitive tabulation of the Army data by ethnic origin, he presented no cross-tabulation by years of schooling or the like. Social scientists quickly came to realize that the evidence proved nothing about group differences, and Brigham was the subject of criticism (see below). But Brigham let his book stand for years. His only published acknowledgment of criticism was a side remark: psychologists should work on their data and ignore armchair challengers (Brigham, 1926). Brigham (1930) did disavow his studies in a professional journal, and since that time psychologists have been happy to point to this as evidence of the scientist's openness to correction. But Brigham's "retraction" was on the largely irrelevant grounds that the Army tests were inhomogeneous; he bowed not at all to the professional consensus that ethnic comparisons themselves are meaningless.

Occasionally a psychologist took to a popular magazine to correct false impressions. In the *Atlantic Monthly* Robert Yerkes (1923), a principal figure in the Army testing, condemned Wiggam's distortions and tried to clarify the draft data. He claimed that the mental-age interpretation was "merely a sort of by-product" intended to simplify matters for the public and should be dismissed. But Yerkes himself then explained various Army results in ways that only fed the flames:

[It] is quite commonly believed that intelligence increases with schooling. This, however, is flatly contradicted by results of research, for it turns out that the main reason that intelligence status improves with years of schooling is the elimination of the less capable pupil. (p. 362)

Not more than 10% of the population are capable of earning the BA, as their earlier departure from school proves. (!)

Two months later (May 1923), Walter Lippmann lambasted Yerkes and Brigham in the *Century Illustrated Magazine*. He objected particularly to the claim that the tests measured innate abilities and foretold who could profit from education; also, he objected to the comparisons of ethnic groups. The *Century* essay was Lippmann's second position paper. The first appeared in installments in *The New Republic*; with rejoinders and extensions, he published 10 articles there between October 1922 and May 1923. There, too, the foe was the elitist who argued that democracy, buttressed by education, could not work. While Lippmann was basically favorable to the tests as an aid in school management, he became increasingly

vehement in attacking psychologists. By the end of the series, his language was strong indeed: "men of science presume to dogmatize," "purely statistical illusion," "behind the will to believe, . . . the will to power," "self-deception as a preliminary to public deception." The "deception" was the claim that native intelligence was being measured. (See Lippmann, 1922a-g, 1923a-c.)

Terman (1922a), who had been a special target, replied to the *New Republic* articles with a heavy-handed sarcasm that gave thoughtful consideration to none of Lippmann's points. Elsewhere Terman himself acknowledged the debatability of some working hypotheses, but he apparently saw Lippmann as a presumptuous layman to be routed. Lippmann, for example, had said (as we would say today) that early experience in the home might account for much of the correlation of IQ with social class. Here is a bit of Terman's reply:

[If so, it would be] high time that we were investigating the IQ effects of different kinds of baby-talk, different versions of Mother Goose, and makes of pacifiers and safety pins. . . . Does not Mr. Lippmann owe it to the world to abandon his role of critic and to enter this enchanting field of research? (p. 119)

This served only to allow Lippmann an ad hominem rejoinder, followed by an exchange of sharp letters. Nothing was clarified, and readers were annoyed with Terman for refusing to discuss issues soberly.

Lippmann (1923) was confused at several points, and his tactics as a debater were not above reproach. Here is an excerpt from the *Century Illustrated Magazine*:

They are determined that education and opportunity shall not count, "for we must (*sic*) assume," says Dr. Brigham, "that we are measuring native or inborn intelligence." To this we can reply that there is no law compelling professors to assume the very thing which they set out to prove. They are quite free to assume nothing and to conclude, if the facts point that way, that they are measuring very crudely some aspects of the mixture of native ability and acquired habits. That is in fact all that modest and critical psychologists claim for the tests. But, unfortunately, the modest and critical psychologists . . . have remained in the background, while their rasher colleagues have offered to the public a yellow science. The headline professors, be it said, to the dismay and chagrin of the true scientists in this field, have succeeded for the moment in producing something like a panic, using misleading statistics to destroy confidence in the value and possibilities of education. (Lippmann, 1923d, p. 103)

If yellow science there be, it inspires yellow journalism. The quoted words on which the sarcasm focuses were a target of Lippmann's own making. In Brigham's context, the words clearly meant "We must assume this *or* the contrary." Brigham

had used his sentence to launch a 10-page examination of alternative explanations, not to close off thought.

While Terman and a few other psychologists were chopping savagely at every critical head, there was also a dignified response. Just two months after Lippmann's first essay objecting to the "13-year-old" interpretation, Freeman (1922) had gotten into the pages of a professional journal a note showing that the mental-age statement was indefensible and saying to his colleagues, in effect, "We must stop talking nonsense about these important matters." As the criticisms continued, Freeman put the central issues into a series of questions, obtained answers from leading testers including Terman and Yerkes, and published a consensus in the *Century Illustrated Magazine* (Freeman, 1923) that should have satisfied Lippmann on every point. It said, among other things, that there was no logical way to judge the native ability of groups that had had dissimilar upbringings. (It is significant that a close reading of Brigham shows that his book had included every one of the specific qualifying remarks required to keep his statements literally in line with Freeman's consensus. Yet the message Brigham transmitted was not muted a bit by his grudging caveats.)

With the appearance of Freeman's consensus statement and an equally judicious piece in the *Atlantic Monthly* by Henry Link (1923), the controversy vanished from the press. These papers aroused no comment; they were water on a bonfire that had already burned out. The entire debate had come and gone in an 18-month period. It provided some lively reading, but it changed few minds. At most, it led psychologists of the time to be as circumspect in public as in their professional writings. This apparently was the only aim of the critics; Lippmann endorsed such use of tests as Freeman's consensus called for. A more restrictive immigration law was passed, as would have happened without the Army data on ethnic groups. The mental-age scale for adults was adjusted but not abandoned even in the face of continuing intraprofessional criticism. The heredity-environment issue remained the subject of heated intraprofessional debate. Reviews by the National Society for the Study of Education in 1928 and 1940 settled little, because the parties put a quite different coloration on the evidence (Whipple, 1928, 1940). Only stray echoes of this scholarly dispute reached the public, however, until the issue was politicized in the 1960s.

William James had warned psychologists that to understand man was not to write his biography in advance, but the testers came very close, in their estimate as to how much education a man could use and what careers he could thrive in. More serious, when the tests determined who would enter the college preparatory program and before that determined who would go into the "fast" section of an early grade, the tests began to determine fates. The testers' sorting process was to shield the child destined to be a worker from the rigors of an academic curriculum. Such a plan would reduce distaste for schooling, prevent failure, and retain him in school longer. Testers said that the IQ was constant; hence to make decisions early was merciful and just. Hostility to this determinism was a central theme of many essays by William Bagley (1925), a conservative educational theorist, and it entered tangentially into Lippmann's writings. This attack too was answered by Terman (1922b), Whipple (1923), and a few others, again in a rhetoric which tended to beg the issues. Bagley's is a "Christian Science Psychology," said Terman, and his vision "blurred by the moist tears of sentiment" (p. 58). The peroration of Whipple's vice-presidential address to the American Association for the Advancement of Science was:

I am not ashamed to find myself supported by Mr. Hoover, who declares that, "We in America have too much experience of life to fool ourselves into pretending that all men are equal in ability, in character, in intelligence, in ambition; that was part of the claptrap of the French Revolution." (pp. 603-604)

This argument requires only passing attention here. It took place in educational and academic circles more than in public, and Bagley got too little support to keep the challenge alive. It is certainly a most important theme, of the 1920s and of the present, for it questions whether pupils who differ should be sorted into educations different in kind, or different in pace, or different in method of instruction. (The constancy issue as such was less important even in the 1920s than the debate suggested, for in their writings to the profession Terman and the others advised periodic retesting, and they advised that decisions be adjusted when the IQ changed.) The wisdom of sorting pupils into distinct programs deserved public scrutiny it did not get. Streaming was the obvious answer to the problem of laggards and dropouts in an efficiency-minded decade. Today, with other social

priorities, there is again uneasiness about school sorting that shapes life chances. But philosophical issues remain unvoiced while loaded, overblown issues such as heritability get all the attention. Society needs to think once again about the kind of equality it would prefer and about the desired relation between productivity, social status, and standard of living. Most of all, it needs to distinguish between education as preparation for service to society, education as preparation to get more out of living, and education as a means of certifying social status.

Some Possible Generalizations

THE *Zeitgeist* AND THE MEDIA

Controversies over social science are not created by findings as such. At any time, the professional literature is full of socially important results that are potential raw material for journalism. The journalist, by and large, controls what becomes public and when. The topics of controversy about tests were always there to be exploited after 1905. More facts became available each year, but there were few surprises. There is complete harmony between the consensus statement of 1923 and the SPSSI statement of 1969. It is hard to see that any one of the controversies changed any minds, or even enlightened its public. At best, controversy corrected an incautious and unpopular statement, here or there. Journalism has no corrective for misstatements that are popular.

Repeatedly, we have seen journalists mining scholarly reports for controversial copy, distorting the original to make it more exciting, pointing up disagreements, and sometimes reporting only the iconoclastic side. Man bites establishment is news. But the journalist cannot keep alive a message the public is not attuned to hear.

There is a tide in the affairs of issues. When the nation had ordered up businesslike public administration, Bagley could not get attention for a philosophical dissent. Around 1960, when eyes were on international competition and schools were charged with slackness, John Gardner's (1961) *Excellence* had a ready market. The same market, short years later, is avid for attacks on the competitive and on the work ethic itself. The public of the twenties was sour on immigrants; hence Brigham's racial comparisons were popular and his critics got little hearing. In the 1970s a proposal merely to do research on ethnic differences

is howled down. The times were against Davis and Bagley; the times favored Rosenthal and Terman. Jensen scorned the Zeitgeist and became a target of scorn himself.

The usual controversy fades away quickly and seems not to alter the trend of social policy. Perhaps the controversies, more than determining policy, are evidence that broad changes in social thought and policy are taking place. When a fresh spring air mass sweeps over the plains of the Midwest, brief lightning and hail storms play upon scattered localities along its edge. The storms are dramatic, but the energy within them is trivial in proportion to the momentum of the oncoming mass and the inertia of the one in possession of the territory. In the first half of this century the American system turned from *laissez faire* to rationalization, system, and an increasingly managerial government. Hoover's "individualism" and Parrington's "romantic idea of human perfectability" were both pushed aside, and the controversy over tests was a symptom of the shift of forces. The world view entrenched before World War II is now under attack, and an alternative scheme that cherishes pluralism, affiliations within local communities, and fulfillment rather than "perfection" is taking shape. It is too early to judge whether the new view or the old one will dominate the next generation, but, I suggest, it is this struggle for the minds of men, and not concern for specifics such as mental tests, that has generated the recent controversies.

HOW THE SCIENTIST CONDUCTS HIMSELF IN PUBLIC

Once caught up in a public issue, scholars who are not journalists at heart lose their composure, their clarity, and their judgment. Scholars typically welcome the widest possible attention to their views because they cherish the ideas and because they prize visibility and influence. What wonder, then, that a scholar given his once-in-a-lifetime moment in the public eye will seek to make the moment memorable?

There is a fundamental difference between the style of the advocate, in law and in journalism, and the style of the scholar. An advocate tries to score every point, including those he knows he deserves to lose. The advocate who bridles his partisanship places his side at an "unfair" disadvantage. Our scholars chose to play advocate when they went before the public, and they abandoned schol-

arly consistency. Terman spoke to the public only of the "constant" IQ, but to the profession he recommended periodic retesting. Brigham minimized the cautionary statements that cast doubt on his conclusions. And the public Jensen remained silent from 1969 to 1973 about the environmentalist positions the professional Jensen set forth in 1968. (In 1974 he answered a challenge by stating that his studies in 1967-1969 had changed his views [Jensen, 1974], but the manuscripts lead me to think that most of the change was subsequent to 1969.)

The American academic is ill trained to cope with the media and the public. In his normal life he speaks to a captive, note-taking audience. He writes for archives where those who want his thoughts *in extenso* can find them, and where the reader can be trusted to weigh sentences in context. But the public reads the headlines and the snappy quotes, and only half-remembers even them. In the dispute of the 1920s the journalists were the first offenders, with Wiggam's flamboyance and Lippmann's muckraking tone and irreverent pinpricks. Terman and Whipple tried to play the same game and were hopelessly overmatched. Their sarcastic jabs appeared to be a frivolous and evasive response to serious if ill-specified charges. At best, they answered peripheral points while leaving main issues in confusion. The academic needs writing skills of an entirely unaccustomed order if he is to make sure that no unwanted implication will be drawn from a buried sentence, that no sentence quoted out of context will advocate what he does not believe, and that no colorful aside will be remembered instead of his main message. We may rail against the journalist for relaying what we said instead of what we meant to say, but mind-reading is not his job. We may rail against the public for not studying the text we put before them. But the writer in public is the servant, not the master, of the reader.

Academics in public take themselves too seriously. They write too much; one thinks of Jensen solemnly answering in print letter after letter, diffusing the force a few sharp retorts on central issues might have had. In public discourse, the more one says, the more trouble he creates for himself. That lesson comes hard to the academic.

The academic cannot control the timing of controversy. When the media offer space, they write a time for delivery into the contract. And the academic who on his own decides that the time has arrived for a public report must speak while the

ear is turned his way. Hence we had one party to our controversies scribbling last-minute additions in airplanes, trying to meet a deadline and still make the most of his moment on stage. If it took a full year for Yerkes to get his reply to Wiggam into print, one can scarcely fault him for not obtaining reviews of his draft; yet the logical holes in his reply made it as damaging to his cause as helpful.

The associations and informal devices of the scholarly community are tested in these conflicts. Some acquit themselves well. The National Society for the Study of Education has a particularly good record in providing forums; its symposia on tests, inheritance, and expectancy effects have clarified disputes in 1922, 1928, 1940, and 1970. Even a politically active group such as SPSSI can at times make a judicious and stabilizing statement. But associations also can be rash and one-sided—witness the Anthropological Association. Hence one cannot assume that scholars collectively will bring a debate to a sober finale. Nor does a balanced and unexciting summing-up get the public ear.

CONSEQUENCES OF KNOWLEDGE AND INNOCENCE

Since Eden, there have been uncertainties about whether knowledge is good. In the scientific ethic, and even more in the vision of social science held by the Progressives, knowledge is created to be made available. But there is a higher knowledge that records the effects of knowledge, and there is a social science still to be built that will clarify when and how knowledge is likely to be used to exploit or corrupt or dehumanize.

Inquiry is best left unrestricted. But the person publishing or popularizing a study does have responsibility for anticipating what his words will suggest to the rightists and the leftists, the exploiters and the *descamisados*. He is not irresponsible when his conclusions sway public decisions; he is irresponsible if his careless writing does so.

Our greatest difficulty is our innocence. To spotlight one question, pleased that social science can answer it, often casts closely related questions into a deeper shadow. The testers of the 1920s were innocently pleased with their new powers. Tests improved the schools' judgments about academic promise; hence they were socially good. But were they wholly good? Whose children gained by the tests? Was there any risk of undue reliance on the new "scientific" judgments? Do the tests

really increase mobility? And what does the very existence of social mobility mean for the health of a democratic society?

Testers as a profession have been accused of being servants of the interests, specifically, of "corporate liberalism" (Cohen & Lazerson, 1972). While it is true that there was a natural affinity between their ideas and a hierarchically organized, differentiated society, testers worked with the social structure, not for it. What distorted their public remarks was chiefly their conviction that they were right. It was so obviously cruel and inefficient to instruct everyone in the same things at the same pace, and so obvious that systematic measurement was providing better information for teachers, that the testers of the 1920s could conceive of no risk or error save that of failure to take the tests seriously. The spokesmen for tests, then and recently, were convinced that they were improving social efficiency, not making choices about social philosophy. Their soberly interpreted research did place test interpretation on a more substantial basis. But they did not study the consequences of testing for the social structure—a sociological problem that psychologists do not readily perceive.

The social scientist is trained to think that he does not know all the answers. The social scientist is not trained to realize that he does not know all the questions. And that is why his social influence is not unfailingly constructive.

REFERENCES

- Ayres, L. D. *Laggards in the schools*. New York: Russell Sage Foundation, 1909.
- Bagley, W. C. *Determinism in education*. Baltimore: Warwick & York, 1925.
- Born dumb? *Newsweek*, 1969, 73(13), 84.
- Brigham, C. C. *A study of American intelligence*. Princeton, N.J.: Princeton University Press, 1923.
- Brigham, C. C. Value of tests in examinations of immigrants. *Industrial Psychology*, 1926, 1, 413-417.
- Brigham, C. C. Intelligence tests of immigrant groups. *Psychological Review*, 1930, 37, 158-165.
- Callahan, R. *Education and the cult of efficiency*. Chicago: University of Chicago Press, 1962.
- Caspari, E. W. Genetic endowment and environment in the determination of human behavior: Biological viewpoint. *American Educational Research Journal*, 1968, 5, 43-56.
- Cohen, D. K., & Lazerson, M. Education and the corporate order. *Socialism Revisited*, 1972, 2, 47-71.
- Coles, R. What can you expect? *New Yorker*, 1969, 45 (9), 169-177.
- Davis, A. Poor people have brains, too. *Phi Delta Kappan*, 1949, 30, 294-295.
- Edson, L. Jensenism, *n*. The theory that I.Q. is largely determined by the genes. *New York Times Magazine*, August 31, 1969, pp. 10-11.

- Eells, K., Davis, A., Havighurst, R. J., Herrick, V. E., & Tyler, R. *Intelligence and cultural differences; a study of cultural learning and problem-solving*. Chicago: University of Chicago Press, 1951.
- Elashoff, J., & Snow, R. E. (Eds.). *Pygmalion revisited*. Worthington, Ohio: C. A. Jones, 1971.
- Fine, B. More and more, the IQ idea is questioned. *New York Times Magazine*, September 18, 1949, p. 7ff.
- Freeman, F. N. The mental age of adults. *Journal of Educational Research*, 1922, 6, 441-444.
- Freeman, F. N. A referendum of psychologists. *Century Illustrated Magazine*, 1923, 107, 237-245.
- Gardner, J. *Excellence*. New York: Harper, 1961.
- Goddard, H. H. *Human efficiency and levels of intelligence*. Princeton N.J.: Princeton University Press, 1920.
- Greenwald, E. Letter. *New York Times Magazine*, September 21, 1969, p. 4.
- Herrnstein, R. J. I.Q. *Atlantic Monthly*, 1971, 228(3), 43-64.
- Hess, R. D. Controlling culture influence in mental testing: An experimental test. *Journal of Educational Research*, 1955, 49, 53-58.
- Hyman, J. D. Letter. *New Republic*, 1969, 161(17), 30-31.
- Jensen, A. R. The culturally disadvantaged: Psychological and educational aspects. *Educational Research*, 1967, 10, 4-20.
- Jensen, A. R. Social class, race, and genetics: Implications for education. *American Educational Research Journal*, 1968, 5, 1-42.
- Jensen, A. R. How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, 1969, 39, 1-23.
- Jensen, A. R. Can we and should we study race differences? In J. Hellmuth (Ed.), *Disadvantaged child*, Vol. 3. *Compensatory education: A national debate*. New York: Brunner/Mazel, 1970.
- Jensen, A. R. *Genetics and education*. New York: Harper & Row, 1972.
- Jensen, A. R. The strange case of Dr. Jensen and Mr. Hyde? *American Psychologist*, 1974, 29, 467-468.
- Klineberg, O. Racial and national differences in mental traits. In W. Monroe (Ed.), *Encyclopaedia of educational research* (Rev. ed.). New York: Macmillan, 1952.
- Kohl, H. Great expectations. *New York Review of Books*, 1968, 11(4), 31-33.
- Lederberg, J. Racial alienation and intelligence. *Harvard Educational Review*, 1969, 39, 611-615.
- Link, H. C. What is intelligence? *Atlantic Monthly*, 1923, 132, 374-385.
- Lippmann, W. The mental age of Americans. *New Republic*, 1922, 32, 213-215. (a)
- Lippmann, W. The mystery of the "A" men. *New Republic*, 1922, 32, 246-248. (b)
- Lippmann, W. The reliability of intelligence tests. *New Republic*, 1922, 32, 275-277. (c)
- Lippmann, W. The abuse of tests. *New Republic*, 1922, 32, 297-298. (d)
- Lippmann, W. Tests of hereditary intelligence. *New Republic*, 1922, 32, 328-330. (e)
- Lippmann, W. A future for tests. *New Republic*, 1922, 33, 9-11. (f)
- Lippmann, W. The great confusion. *New Republic*, 1922, 33, 145-146. (g)
- Lippmann, W. Mr. Burt and the intelligence tests. *New Republic*, 1923, 34, 263-264. (a)
- Lippmann, W. Rich and poor, girls and boys. *New Republic*, 1923, 34, 295-296. (b)
- Lippmann, W. A judgment of tests. *New Republic*, 1923, 34, 322-323. (c)
- Lippmann, W. A defense of education. *Century Illustrated Magazine*, 1923, 106, 95-103. (d)
- Moynihan, D. P. Comment: Jensen not 'must reading' in the Nixon Cabinet. *Journal of Social Issues*, 1970, 26, 191-192.
- Parrington, V. L. *Main currents of American thought* (Vol. 3). New York: Harcourt, Brace, 1930.
- Popenoe, P. Measuring human intelligence. *Journal of Heredity*, 1921, 12, 231-236.
- Rosenthal, R., & Jacobsen, L. *Pygmalion in the classroom*. New York: Holt, Rinehart & Winston, 1968.
- Rowan, C. T. How racists use "science" to degrade black people. *Ebony*, 1970, 25(7), 31-40.
- Society for the Psychological Study of Social Issues. SPSSI on current IQ controversy: Heredity versus environment. *American Psychologist*, 1969, 24, 1039-1040.
- Terman, L. M. *Measurement of intelligence*. Boston: Houghton Mifflin, 1916.
- Terman, L. M. *The intelligence of school children*. Boston: Houghton Mifflin, 1919.
- Terman, L. M. The great conspiracy. *New Republic*, 1922, 33, 116-120. (a)
- Terman, L. M. The psychological determinist; or Democracy and the I.Q. *Journal of Educational Research*, 1922, 6, 57-62. (b)
- Thorndike, R. L. Review of *Pygmalion in the classroom*. *American Educational Research Journal*, 1968, 5, 708-711.
- Whipple, G. M. The intelligence testing program and its objectors—Conscientious and otherwise. *School and Society*, 1923, 17, 561-568, 596-604.
- Whipple, G. M. (Ed.). *Nature and nurture. Twenty-seventh yearbook of the National Society for the Study of Education*. Bloomington, Ill.: Public School Publishing Co., 1928.
- Whipple, G. M. (Ed.). *Intelligence: Its nature and nurture. Thirty-ninth Yearbook of the National Society for the Study of Education*. Bloomington, Ill.: Public School Publishing Co., 1940.
- Wiggam, A. E. The new decalogue of science. *Century Illustrated Magazine*, 1922, 103, 643-650.
- Wolff, R. P. Letter. *Atlantic Monthly*, 1971, 228(6), 106.
- Yerkes, R. M. (Ed.). *Psychological examining in the United States Army. Memoirs of the National Academy of Sciences*, 1921, 15, 1-890.
- Yerkes, R. M. Testing the human mind. *Atlantic Monthly*, 1923, 131, 358-370.