

Either a Paradigm Shift or No Mental Measurement: The Nonscience and the Nonsense of The Bell Curve

ASA G. HILLIARD, III
Georgia State University

In the spring of 1994, I was a panelist at the American Educational Research Association/National Council on Measurement in Education Symposium entitled, "Whatever happened to the measurement of intelligence?" In his introductory remarks, the chairman of the symposium suggested that IQ testing was in decline. Part of the evidence cited for this view came from a review of the programs at the annual meetings of organizations like the American Educational Research Association and the American Psychological Association and the National Council on Measurement in Education. In the chairman's review of these programs at annual meetings, he had seen virtually no sessions on IQ testing in the past 7 years, and there was no session at the 1994 Ameri-

can Educational Research Association or the National Council on Measurement in Education other than our panel.

To some, this seemed to suggest a declining interest in the subject, and perhaps a weakening of the hold that IQ testing and thinking has on the minds of psychologists and others. I then challenged that view, indicating that in my experience, IQ testing and thinking was alive and well. In fact, for the entirety of my nearly 40-year career as a psychologist and educator, I have been involved in one struggle or another over the validity and use of IQ tests and the construct of intelligence in education and in related fields. It is interesting that this questionable professional practice has faded to the background of consciousness of so many psychologists.

-
- ***Editor's Note:*** *The publication of The Bell Curve has fueled a controversy regarding racial differences in psychological testing. This invited article documents the political and economic bias permeating the ideas presented in the book. I thank Maria P. P. Root for her instrumental role in facilitating the publication of this article.*
 - *Dr. Hilliard is the Fuller E. Callaway Professor of Urban Education at Georgia State University. This article is based on Dr. Hilliard's address at the annual meeting of the American Psychological Association, August 13, 1995, in New York City.*

Reprint requests should be directed to Asa G. Hilliard, III, Ph.D., Educational Policy Studies, Georgia State University, Atlanta, GA 30303.

Perhaps *The Bell Curve* will have the effect of causing the light of science to shine in corners that seldom, if ever, see the light of day.

Within 2 months or so after this friendly discussion about the importance of IQ, Murray's and Herrnstein's book, *The Bell Curve* (1994), was published. I tried to get a copy at the Oxford Bookstore in Atlanta on several occasions; each time the book was sold out. I tried other bookstores, with the same result. Finally, I was able to locate a copy while traveling, at a bookstore in the Cincinnati airport. The expensive, hardback copy of *The Bell Curve* was being sold and prominently displayed there.

Within a very short period of time, this massive hardbound book of some 845 pages had found its place on the *New York Times* bestseller list where it remained for several weeks. It was introduced with great fanfare, with the surviving author, Charles Murray, being featured on virtually all the major talk shows and in popular magazines. In fact, he was apparently able to dictate the format for such talk shows as the Phil Donahue show, where he was interviewed atypically, one-on-one, by Phil Donahue.

Following the well-hyped introduction of *The Bell Curve*, I was invited to give reactions to the book on several university campuses. In most of those audiences, several professors of psychology were present. I queried each audience as to who had read the book. To my surprise, in audiences of graduate students and faculty at major universities, very few in attendance at my presentations indicated that they had read the book! That made me wonder: Who has made the book a bestseller, and why? What deep feelings are touched by it?

One of the other things that I have noted is that, aside from the shrill outcry of those who feel maligned by the conclusions of the book and who challenge it largely on grounds of morality and fairness, the general scholarly reaction to the book has been comparatively muted. Most of the rare critiques are uninspired and, in my opinion, miss the main points.

It is fair to say that the overwhelming response to the book in the editorial pages of newspapers and popular magazines is that *The Bell Curve* is a work of science, controversial, and says some hard things that need to be said, even though they may not be "politically correct," that cute little popular quip that seems intended to quell debate and criticism. There is also the suggestion that those who are maligned by the book are fearful of its "truths." But the real question is: Is it scientifically correct?

Unlike the reception of Arthur Jensen's and William Shockley's work years ago, Murray's and Herrnstein's opinions seem to have been much more openly welcome within the ranks of the conservative elite and even quite well tolerated by the liberal elite, both lay and scholarly. In fact, surveys show that *The Bell Curve* views on IQ differences between and among "races" is mainstream, both in the general public and in the profession of psychology as well (Duke, 1991; Snyderman & Rothman, 1990). I am aware that Murray and Herrnstein also attacked people who score in the bottom quartile on IQ tests, whether African or not, and that the primary focus of the book is not on race.

Over the years, I have been interested in the study of several interrelated topics. I have been interested in the validity of mental measurement in its relationship to education. I have been interested in the *validity of teaching* and the *validity of schools*, looking especially at those schools that get excellent achievement from students without regard for race or socioeconomic status (Backler & Eakin, 1993; Hilliard, 1991; Sizemore, Brosard, & Harrigan, 1982). I have been interested in the study of history and culture as it relates to assessment and teaching and learning, especially the history and culture of people of African descent. And finally, I have been interested in the study of racism/white supremacy in science, including the scientific study of oppression and its dynamics. All of these things are a part of the context within which the study of IQ, intelligence, and their application take place. IQ test results are confounded and uninterpre-

table in the absence of an understanding of all of these.

Literature and clinical experiences have supported the conclusion that human problems do not divide themselves neatly into the academic disciplines represented in the structure of our universities, such as psychology, sociology, anthropology, and so on. Human problems present themselves as wholes. It is highly unlikely that those problems can be understood from the perspective of a single discipline. It is even less likely that the problem of how we think and learn and how much we can think and learn can be understood from the perspective of a small segment of a single discipline. Specifically, applied psychometrics within the broad field of psychology—independent of related social disciplines, even cognitive psychology—is insufficient to understand the problem of human potential. Aspects of this very complicated problem, the understanding of human potential as it is situated in a context, can and must be approached from a variety of academic disciplines, simultaneously and in a coordinated way. I am aware of no such efforts that claim the attention of the mainstream in psychometrics. The professional literature in the field of IQ psychology is ultra narrow and restricted.

Unfortunately, mental measurement psychologists have manifest little interest in the related and necessary social science disciplines. In fact, it appears that some mental measurement specialists are fearful of opening the field to scrutiny by other scientists, such as anthropologists and linguists, to the profound detriment of the scientific study of human potential and its application to teaching and learning and related areas.

And so in this article, I want to raise issues of science, not issues of morality or fairness, although these are important. I want to look at the foundation upon which the structure of *The Bell Curve* is erected. I believe that any scientific look at the foundation of psychometrics and at the structure of *The Bell Curve* will reveal that there is no science in *The Bell Curve* at all.

The Bell Curve mimics a narrow-minded

approach to science. The current approach to the study of human potential with IQ tests and to their application is analogous to trying to do space travel relying only on the knowledge of the physicist. On the contrary, scientific problem solving at NASA is unavoidably a matter of multidisciplinary teamwork. The problem of how to know about human potential, on the other hand, is to present a virtually solitary quest, largely left to the tinkering of the psychometrician and statistician. There is a profound paradigm problem here for the field of psychology. It is not a problem for Murray's and Herrnstein's IQ uses alone.

There is also a political problem here. We have not yet overcome problems of white supremacy belief and behavior in the United States. To what extent does it intrude in the work of psychologists? Can we feel comfortable with any results when we have not made a systematic analysis of the effects of the well-documented racial politics of psychology on mental measurement (Gould, 1981; Kamin, 1974; Thomas & Sillen, 1972)?

Performing the traditional mental measurement or IQ test construction rituals (routines) better will not address the basic problems with mental measurement, problems that stem from our failure to study potential sources of significant variation in test-taker performance, such as culture and political treatment. We cannot do this merely by trying to refine factor analysis or item analysis techniques. Within the field of psychology itself, within traditional psychometrics, and within related fields, we can find the evidence to challenge the scientific validity of *The Bell Curve*.

Murray and Herrnstein raised a number of questions with their presentation.

1. They assumed that psychology, at present, can measure mental capacity accurately.
2. They assumed that the devices (IQ tests) created for mental measurement are universally applicable, in a culturally plural and highly politicized world.

3. They assumed that correlation is causation.
4. They assumed that human potential is correlated with certain human behaviors of interest, such as crime, school achievement, welfare dependency, teenage pregnancy, and so forth, and that these are explain by IQ.
5. They assumed that “intelligence,” which is supposedly measured by IQ, is stable and does not change.
6. They assumed that there is equal opportunity to learn and common exposure to cultural experiences for all because there are no controls in their research for variation in opportunity or exposure.

In looking at these and other implicit assumptions, we may develop very quickly a list of at least eight “scientific cracks” in *The Bell Curve*. In other words, there are at least eight fundamental scientific flaws in the Murray–Herrnstein bell curve. This is in no way a definitive list of the problems. The purpose of presenting such a list and the accompanying discussion is to illustrate what is missing and what type of scientific work still has to be done before the IQ testing used by Murray and Herrnstein and other forms of mental measurement can mean anything important, if it ever can.

The last two chapters in the Murray and Herrnstein book, in particular, are matters of politics or public policy. They are an attempt to apply the results of IQ testing, which they assumed were “mental measurements,” to the solution of a whole complex of social problems. However, those two chapters, although interesting, are wholly invalidated by virtue of the fact that they are linked to the validity of mental measurement by reliance on IQ tests. According to Murray and Herrnstein, low IQ is synonymous with low intelligence and causes crime, early pregnancy, welfare dependency, and social failure. For them, none of these things can be reversed, because IQ cannot be changed. It is not my purpose here to

deal with these complicated matters. I merely note that these matters have been the perennial concerns of Murray and Herrnstein and their elite professional allies (“Mainstream Science,” 1994), and their political sponsors for many years. The only thing that is different here is that their views are now benefiting from a well-funded and well-planned sophisticated right-wing propaganda blitzkrieg.

**Spin disguised, period.* Murray’s work on *The Bell Curve* was underwritten by a grant from the Bradley Foundation, which the *National Journal* in 1993 described as “the nation’s biggest underwriter of conservative intellectual activity.” Bradley is a respectable foundation about whose financial support no author need apologize. But Bradley backs only one kind of work: that with right-wing political value. For instance, Bradley is currently underwriting William Kristol, a former adviser to President Bush and director of the Project for a Republican Future. *The Bell Curve* identifies Murray as a “Bradley Fellow” but gives readers no hint of the foundation’s ideological requirements. Telling readers this would, needless to say, spoil the book’s pretense of objective assessment of research.

Slipping down the slope from the respectable Bradley Foundation, Herrnstein and Murray praise some research supported by the Pioneer Fund, an Aryan crank organization. Until recently, Pioneer’s charter said it would award scholarships mainly to students “deemed to be descended from white persons who settled in the original 13 states.” Pioneer supports Rushton and backed the “Minnesota Twins” study, which purports to find that identical twins raised apart end up similar right down to personality quirks. The Aryan crank crowd has long been entranced by the Minnesota Twins project, as it appears to show that genes for mentation are entirely deterministic. Many academics consider the protocols used by the Minnesota Twins study invalid.

Lesser examples of disguised ideological agenda are common in *The Bell Curve*. For example, at one point Murray presents an extended section on problems with the D.C. Police Department, saying their basis lies in

“degradation of intellectual requirements” on officer hiring exams. Information in this section is attributed to “journalist Tucker Carlson.” No one who lives in Washington doubts its police department has problems, some of which surely stem from poor screening of applicants. But who is the source for the particularly harsh version of this problem presented in *The Bell Curve*? “Journalist Tucker Carlson” turns out to be an employee of the Heritage Foundation; he is an editor of its house journal *Policy Review*. Heritage, for those who don’t know it, has a rigid hard-right ideological slant. Its *Policy Review* is a lively and at times insightful publication, but anyone regarding its content as other than pamphleteering would be a fool. The article *The Bell Curve* draws from lampoons the intelligence of D.C. police officers because some cases have been dismissed owing to illegible arrest records. And just how many high-IQ white doctors have unreadable handwriting? If an article in *Policy Review* were an impartial source of social science observations, Murray would simply come out and say where his citation originates. Instead he disguises the source, knowing full well its doctrinaire nature. (Easterbrook, 1995, pp. 40–41)

I do not dispute the fact that Murray and Herrnstein found what they said they found, the IQ correlations with certain items. I dispute their interpretation, an interpretation of the meaning of the relationships that is based on fundamental scientific flaws, each sufficient to invalidate the results reported in *The Bell Curve*.

We can ask at least eight major questions about *The Bell Curve* to determine if it reflects the state of the art in these eight scientific areas, areas of study where there is a body of literature relevant to the question of mental measurement. When one looks carefully at these areas, it will be clear that the authors of *The Bell Curve* have restricted themselves to safe ground, reflecting little, and usually no, awareness of relevant scientific information in many areas. The mental measurement enterprise is barely in its infancy, certainly far from its maturity. If it fails to incorporate the relevant findings from

the sister sciences, it never will mature. We cannot have confidence in grandiose inferences from invalid IQ tests and an invalid model of application of mental measurement to human problem solving.

The First Crack

The Bell Curve is bad psychology; it does not reflect the state of the art in mental measurement. To the best of my knowledge, the most recent international meeting of scholars in mental measurement who were attempting to develop a state-of-the-art synthesis was held in 1988 in Melbourne, Australia. Helga Rowe, president of the Australian Research Association, edited a book of the important papers from that conference and published it in 1990. Many interesting items were mentioned in this conference summary. However, at least three important points were made relative to the validity of the construct of intelligence and IQ instruments.

1. First, attendees were unable to come to a consensus on a definition of intelligence. To put it mildly, there is a major construct validity problem that persists.
2. But even more serious was the second matter. In fact, it was so serious that this second topic dominated the introduction to the report by Helga Rowe:

Erickson’s (1984) overview of research from an anthropological view shows, for example, that mental abilities (including language and mathematical abilities) that were once thought to be relatively or even totally (as presumed by classical learning theory and Piagetian developmental theory) independent of context, are much more sensitive to context than traditionally thought. Cognitive processes such as reasoning and understanding develop in the context of personal use and purpose. The demand characteristics of a learning task

can be changed by altering the context within which it is presented. (Rowe, 1991, p. 6)

For the first time, I believe, a major gathering of mental measurement psychologists acknowledged the profound meaning of *context* in mental measurement. In effect, they said that when and if ever intelligence is conceptualized and measured, it will, of necessity, have to take into account the *context* within which individuals exist and within which mental measurement efforts are conducted. This is because *meaning* of communications is always context specific, never universal. Consequently, if a universal mental operation is to be manifest (e.g., inference of syllogistic reasoning), to detect it, the mental measurement expert *must also be expert in context*. That expert must develop valid instrumentation and processes in response to context and salient for it in order to assess it. Pragmatically, this means that the mental measurement expert must possess the same kind of expertise as anthropologists, linguists, and other behavioral scientists. If not, he or she must collaborate with those who have such expertise. Otherwise, the context cannot be understood in any scientifically valid way.

This matter of the importance of context to measurement was also the main ingredient in the farewell article by Gavriel Salomon, former editor of the *Educational Psychologist* (Salomon, 1995). Looking at measurement-research technical writing for 4 years, he articulated the scientific issues, problems, and imperatives brilliantly. I might add that many researchers, especially ethnic minority researchers, have been making these same arguments for years.

Gone are the one-shot trial, short experiments carried out under highly contrived conditions; gone are the simple statistical analyses, and gone is the exclusivity afforded to the quantitative approach. Gone also are the simple-minded questions and with them the simple, two-group horse-race comparisons among ecologically strange treatments (e.g., Pintrich, 1994). . . .

My second observation is that at least two traditionally espoused assumptions underlying much of the work in educational psychology need to be seriously revised. These two assumptions are (a) that most if not all that is important and interesting to educational psychology lies in the study of the (decontextualized) individual; and (b) that complex phenomena concerning learning, development, and other educationally relevant psychological phenomena are to be broken down into simpler and more easily controllable elements to be studied as discrete elements. The need for revision of these assumptions stems from several developments: from the expectation for significantly greater ecological validity of our research, from implications emanating from the cognitive revolution, from the newly accepted research paradigms, and from the demand for greater practical relevance.

Two things are wrong with these assumptions. First, countering the exclusive focus on the individual, we have come to accept the premise that learning is social. It is as much an interpersonal as an intrapersonal process, and it is a situated, culturally, disciplinarily, and contextually anchored process. To paraphrase Sarason (1981), learning is a socially based process, and this renders suspect explanations that focus solely on the individual learner.

Second, we have gradually come to realize that phenomena of interest (e.g., the functioning individual and the learning environment), once broken down into their more basic elements such as discrete cognitive processes, motivational attributions, or computer-related activities, cease to resemble or represent the real-life phenomena of interest (e.g., Bronfenbrenner, 1979; Bruner, 1991). Many of the phenomena we study—the learning individual, classroom activities, anxiety as it affects learning, social relations with respect to individual's well-being, the development and function of learning strategies, overcoming students' misconceptions, or individual, gender, and cultural differences—are in fact *composites*. . . . And because composites are always greater than and have a different meaning than the sum of their components, one cannot study compo-

nents and expect the findings to apply to the composites.

Proposition 1: Our main (though not necessarily exclusive) focus needs to change from the study of isolated and decontextualized individuals, processes, states of mind, or interventions to their study within wider psychological, disciplinary, social and cultural contexts (see, e.g., Goodenow, 1992).

Proposition 2: Based on the premise that individuals are themselves composites and interact with composites, not with isolated variables, states, or processes, our models, experimental designs, and measures should ultimately (although not immediately) reflect the composites of such real-life settings. In other words, the atomic fabric of our trade ought to be molecules. (Salomon, 1995, pp. 105–106)

To accept this principle of the meaningfulness of context would revolutionize the practice of the approaches to psychological assessment. I believe that this is why mental measurement experts have avoided such specialties as cultural linguistics as if they were the plague. I do not believe that they are ignorant of the criticisms of mental measurement practices over many years that have been based on this very principle of attention to context. Of course, it will be extremely difficult to stop doing what is merely *economically* correct and to do what is *scientifically* correct. It will be costly because valid mental measurement will be labor intensive, initially. But there is no alternative if we seek validity.

3. Finally, at the Australian meeting, papers were presented that questioned the long-assumed relationship between IQ and school achievement.

As can be seen in table 13.1, the overall results failed to fulfill the expectation of a close positive relationship of intelligence test scores and performance scores derived from system control. The reported correlation coefficients are remarkably low; in most cases they are close to zero. Few coefficients reach values of .4–.5. Only in

four studies . . . can correlation coefficients of this size be found. Thus the reported results from these studies do not support the general assumption that intelligence tests are good predictors of an individual's performance when operating a complex system. . . . In addition, most studies agree with respect to the interpretation of the results in two important ways. One, it is argued that the task (i.e., simulated systems) have higher ecological validity and are closer to reality than problem situations, such as intelligence tests' items, or the Tower of Hanoi . . . which have traditionally been studied by cognitive psychologists. Two, the low correlation coefficients allow us to infer that intelligence test scores cannot be regarded as valid predictors for problem solving and decision making in complex, real-life environments. (Kluwe, Misiak, & Haider, 1991, pp. 228, 232)

It appears that complex, conceptually oriented problem solving in novel situations may be unrelated to IQ. Coaching companies in the United States make good profits teaching test-taking routines and strategies to those who are able to afford the classes. Scores are raised on IQ-like tests. But what about problems for which algorithms have yet to be invented, and which take context into account? We really don't know the answer to that question do we? That is the point.

These and other matters had a thorough hearing at the Melbourne conference. Of course, neither this conference nor that body of literature in the social sciences that is related to these findings is reflected in *The Bell Curve*.

In other words, it is impossible to reconcile the state-of-the-art conversation that took place among mental measurement psychologists in 1988 in Melbourne, Australia with the approach taken by Murray and Herrnstein. Their work is clearly inconsistent with the state of the art in mental measurement as reflected in the reports of the

work of these psychologists. I might add that the Melbourne meeting was an open, professional meeting and was not restricted to a clique of psychologists, sponsored by radical right-wing advocacy foundations (Easterbrook, 1995; Lane, 1995; Sedgwick, 1995). For example, Lane reported the following:

No fewer than seventeen researchers cited in the bibliography of *The Bell Curve* have contributed to *Mankind Quarterly*. Ten are present or former editors, or members of its editorial board. This is interesting because *Mankind Quarterly* is a notorious journal of "racial history" founded, and funded, by men who believe in the genetic superiority of the white race.

Mankind Quarterly was established during decolonization and the U.S. civil rights movement. Defenders of the old order were eager to brush a patina of science on their efforts. Thus *Mankind Quarterly's* avowed purpose was to counter the "Communist" and "egalitarian" influences that were allegedly causing anthropology to neglect the fact of racial differences. "The crimes of the Nazis," wrote Robert Gayre, *Mankind Quarterly's* founder and editor-in-chief until 1978, "did not, however, justify the enthronement of a doctrine of *a-racialism* as fact, nor of egalitarianism as ethnically and ethically demonstrable.

Undaunted, *Mankind Quarterly* published work by some of those who had taken part in research under Hitler's regime in Germany. Ottmar von Verschuer, a leading race scientist in Nazi Germany and an academic mentor of Josef Mengele, even served on the *Mankind Quarterly* editorial board.

Since 1978, the journal has been in the hands of Roger Pearson, a British anthropologist best known for establishing the Northern League in 1958. The group was dedicated to "the interests, friendship and solidarity of all Teutonic nations."

Pearson's Institute for the Study of Man, which publishes *Mankind Quarterly*, is bankrolled by the Pioneer Fund, a New York foundation established in 1937 with the money of Wickliffe Draper. Draper, a textile magnate who was fascinated by eugenics, expressed early sympathy for Nazi Germany, and later

advocated the "repatriation" of blacks to Africa. The fund's first president, Harry Laughlin, was a leader in the eugenicist movement to ban genetically inferior immigrants, and also an early admirer of the Nazi regime's eugenic policies. (Lane, 1995, pp. 126-127)

The Second Crack

The Bell Curve is bad biology/anthropology. Murray and Herrnstein used the term, *race*, however, not in any scientific way. Some of their conclusions are related to *race*. In fact, perhaps sensitive to the difficulty with the term *race*, they arbitrarily shift to the term *ethnicity* as a substitute. Whereas *race* as used by Murray and Herrnstein is intended to have a biological meaning, the substitute term, *ethnicity*, is really cultural. Moreover, it has no construct validity for the psychologist. The problem is that Murray and Herrnstein did not articulate a scientific procedure for getting a *racial* or an *ethnic* sample. Of course, the genetic "heritability of ethnicity" makes no scientific sense at all. What do these writers do to document the construct validity of ethnicity? What are their scientific procedures for determining it?

But whatever the meaning of the terms, the important fact is that there is a body of literature in which scholars attempt to deal with the construct of *race* scientifically. Whereas the term *race*, mainly popularly associated with phenotypical variety, may have political meaning, neither Murray and Herrnstein nor the science that they cited or any science that they did not cite establishes an operational definition of race that is acceptable to the scientific community.

The obvious phenotypical variety notwithstanding, who is the psychologist who does racial comparisons in "scientific studies" who is prepared to offer valid criteria for racial or ethnic sample selection? Yee (1983) has been calling us to task on this matter for years. Psychologists generally have simply ignored his arguments, looked the other way, and proceeded with arbitrary "racial" sam-

ples in their studies as if no problem existed. Whatever Murray and Herrnstein did with race in their book, it was not science. Once again, this problem is not theirs alone; it is endemic to behavioral science research.

There is a large body of literature on the matter of race and science. A sample of that body of literature would include the following: Benedict (1959), Gossett (1973), Guthrie (1976), Montagu (1964, 1974), and Yee (1983). An examination of the bibliography and the discussion in *The Bell Curve* will show that the relevant literature in this area was not reflected. That is a scientific error.

The Third Crack

The Bell Curve is bad pedagogy. There is an extensive body of literature now that documents the power of schools to change student achievement in significant ways. Although all schools do not, some schools do. As mentioned previously, some teachers and some schools seem to have no trouble whatever producing the highest levels of academic achievement *in spite of IQ, socioeconomic status, single parent families, drug-infested neighborhoods, gang-banging neighborhoods, and so on*. These types of schools have always existed.

One of the researchers who has been interested almost exclusively with such schools is Sizemore. The Vann School and the Madison school in Pittsburgh are but two examples. These are the highest achieving schools in the city. The Vann school has been at, or near the top of, that city for nearly 20 years. Benton Harbor, Michigan offers a case where a whole school district was identified as the worst district in America in *Education Week* in 1993. Within 1 year, it was cited by the state of Michigan as the most improved district in the state, surpassing in academic achievement many middle-class suburban districts in the state.

There are research and evaluation data to substantiate the point that some schools succeed where failure or low performance is

predicted by IQ or previous achievement. The students are not supposed to be able to do these things, according to the theories of Murray and Herrnstein. Yet they most certainly do. Such achievement is not a surprise at all to those who are familiar with education in the African American communities over the years, especially in the historically Black colleges and universities.

The Bell Curve does not reflect the state of the art in power schools literature.

The Fourth Crack

The empirical record on inequity among schools is quite clear, an inequity in school treatment that puts minority groups and poor students at a disadvantage. Simply put, it is obvious that schools do not all offer the same quality of service. This is a part of the *context that varies*. There are also well-established approaches to the scientific study of school equity/inequity and a vast literature on it. It is a scientific error in a book such as *The Bell Curve* to fail to consider treatment variation, the intervening variable between IQ and achievement. The quality of the treatment that students receive is a major variable. Based on nearly 40 years of observing schools, I am convinced that this variation in the quality of teacher and school treatment is the major variable in student achievement. The ethnographic research literature in particular supports this (Heath, 1983; Kozol, 1991; Lewis, 1995; Oakes, 1985; Rist, 1973).

Unfortunately, few mental measurement professionals seem to concern themselves with this significant variability. Dramatic case studies such as that reported by Kozol (1991) reveal just how massive these inequities are. It borders on the irresponsible for scientists to blame children for low achievement in the face of the undocumented "sav-
age inequalities" in their treatment. But it is bad science, not just bad morality, to fail to control for known sources of variation. If

they are unknown, then the scientist is not prepared to do the research.

The Bell Curve fails to reflect the state of the art in school inequity research. This is a gross scientific error.

The Fifth Crack

The Bell Curve reflects bad cultural linguistics and cultural anthropology. Culture is an unavoidable part of the context for mental measurement. Cultural diversity is an empirical reality. However, valid and sophisticated cultural observation requires expertise. Because many psychometricians lack expertise in culture, they tend to downplay its significance in mental measurement. There is a voluminous literature on culture, and even on the relationship of culture to cognition and measurement. Culture is context, a part of the context that professional consensus now acknowledges must be considered. However, it must be considered in a *systematic* and *sophisticated* way to be scientific. *The Bell Curve* does not reflect the state of the art in culture and mental measurement. It reflects an ignorance of cultural linguistics and cultural anthropology.

There are several types of literature on culture relevant to measurement questions. It must be understood that to acknowledge culture is to complicate the measurement problem significantly. It makes matters "messy," necessarily so. To acknowledge culture is to destroy the easy universalism in instrumentation and interpretation. Yet to ignore culture is to ignore reality itself and to lose all chance of producing valid science; see Cohen (1969), Cole, Gay, Glick, Sharp (1971), Helms (1992), Hoover, Politzer, and Taylor (1995).

The Sixth Crack

There is a well-documented history of racism/white supremacy in psychology and, in particular, in psychological research. Elites

within psychology, as in all other academic disciplines, far from being invulnerable to racism and white supremacist ideology, reflect it in virtually the same proportions as in the population at large. There is a voluminous literature on racism in psychology and psychological research (Ani, 1994; Chase, 1977; Gould, 1981; Guthrie, 1976; Kamin, 1974; Thomas & Sillen, 1972; Weinreich, 1946). *The Bell Curve* does not reflect the state of the art in this literature. I have deliberately added a large special section of a selected bibliography on racism and white supremacy in science to this article.

The Seventh Crack

IQ testing is bad measurement; therefore, *The Bell Curve* relies on bad science. Many scientists who study human behavior challenge the very use of the word "measurement" when it is applied to IQ testing. Physicists and chemists understand measurement to mean the application of an interval scale to phenomena. The issue with IQ testing is whether monocultural material in a multicultural world can be used to construct measuring instruments, with an interval scale: Specifically, in a multilingual world, can a single language be used to construct an IQ test? Whether one agrees or not with cultural sociolinguist Shuy (1977), who asserted the folly of trying to do so, those who would use language to construct IQ tests must reflect a sophisticated understanding of linguistic and cultural dynamics.

The Bell Curve reflects no awareness of the existence and meaning of linguistic and cultural diversity and its meaning as a threat to validity (Chomsky, 1971; Cohen, 1969, 1971; Cole et al., 1971; Helms, 1992; Hoover, Politzer and Taylor (1995); Smith, 1979). Armchair cultural evaluation will not do.

It cannot be overemphasized that the challenge to the measurability of the unknown construct "intelligence" is a challenge not merely based on cultural diversity among ethnic groups, it is a more funda-

mental challenge to the validity of IQ tests as measuring devices for anyone. Although many authors have discussed this matter, perhaps the best collection of articles on this subject appeared in Houts's (1977) excellent book, *The Myth of Measurability*, which by its title, identifies the problem precisely. This book begs for a wide dialogue in the field of psychology that should be held in a multidisciplinary group of professionals: Can measurement instruments be constructed from cultural material, given the guaranteed diversity in human language, values, and experiences?

Zacharias, a physicist from MIT, was eloquent in making this challenge: *His challenge was to the quality of the database.*

The main defect in both sides, or either side, of this argument is that the protagonists pay so little attention to the *quality of the data base*. They revert to saying that the data are not very good, but let's use them anyway because they are all we have. . . .

The worst error in the whole business lies in attempting to put people, of whatever age or station, into a single, ordered line of "intelligence" or "achievement" like numbers along a measuring tape: eighty-six comes after eighty-five and before ninety-three. Everyone knows that people are complex—talented in some ways, clumsy in others; educated in some ways, ignorant in others; calm, careful, persistent, and patient in some ways; impulsive, careless, or lazy in others. Not only are these characteristics different in different people, they also vary in any one person from time to time. To further complicate the problem there is variety in the *types* of descriptions, the traits *tall*, *handsome*, and *rich* are not along the same sets of scales as *affectionate*, *impetuous*, or *bossy*.

As an old professional measurer (by virtue of being an experimental physicist), I can say categorically that it *makes no sense to try to represent a multi-dimensional space with an array of numbers ranged along one line*. This does not mean it's impossible to cook up a scheme that tries to do it; it's just that the scheme won't make any sense. It is possible to strike an average of a column of figures in a telephone

directory, but one would never try to dial it. Telephone numbers at least represent some *kind* of idea: they are all addressed like codes for the central office to respond to . . . implicit in the process of averaging is the process of adding. To obtain an average, first add a number of quantitative measures, then divide by however many there are. This is all very simple, *provided the quantities can be added*, but for the most part with disparate subjects, they cannot be. (Zacharias, 1977, pp. 69–70)

This is the discussion that we still must have. What is the quality of the database? Psychometry has chosen to ignore this question and to concentrate mainly on an approach that *assumes* the quality of the database, that *assumes* that language, and in particular, vocabulary, has *universal semantics*. No science of linguistics will support this idea.

Nearly 20 years ago, at the American Psychological Association's annual meeting in San Francisco, with David Wexler in the audience, I was on a panel that included the heads of The Educational Testing Service and The Psychological Corporation, and attorneys for The American Psychological Association, among others. The subject of the panel dealt with the measurement of IQ and aptitude. I raised a point then that has yet to have a full debate within the American Psychological Association or the broader community of measurement experts. I challenged the symposium to discuss the point on two separate occasions during that panel meeting. No one accepted the challenge. The point had to do with using the relevant scientific insights of cultural anthropologists and cultural linguists to evaluate the psychologists' use of culture and language in the construct of all mental measurement devices. I have had several occasions in other forums since that time to raise this same issue (Hilliard, 1990c). In each case, the issue has been ignored or avoided.

Nothing from the past leads me to suspect that the future is likely to be different. Nevertheless, I raise it again. For Zacharias (1977), measurement was a scientific procedure that followed rigorous rules. To him

and others observing the work of psychologists, those rules are routinely violated.

The Eighth Crack

The Bell Curve is bad genetics. The question we must ask is, did Murray and Herrnstein use the science of genetics in discussion of heritability in a way that would be reflective of the state of the art in genetics today? Once again, there is a literature on this that must be reviewed. It was not. Samples of that literature include Cavalli-Sforza, Menozzi, and Piazza (1994) and Subramanian (1995). Murray and Herrnstein seemed to prefer citations by authors like Rushton:

The sole researcher asserting a hypothesis in this category is J. Philippe Rushton, a psychologist at the University of Western Ontario. *The Bell Curve* makes a point of praising Rushton as "not . . . a crackpot." But a crackpot is precisely what Rushton is. He believes that among males of African, European, and Asian descent, intellect and genital size are inversely proportional, and that evolution dictated this outcome in an as-yet-undetermined manner. Sound like something the sixteen-year-olds at your high school believed? That should not stop Rushton or any researcher from wondering if there might have been different selection pressures on different racial groups. But Rushton's "research" methods, defended by *The Bell Curve* as academically sound, are preposterous. For instance, Rushton has conducted surveys at shopping malls, asking men of different races how far their ejaculate travels. His theory is the farther the gush, the lower the IQ. Set aside the evolutionary absurdity of this. (Are we to presume that in prehistory low-IQ males were too dumb to find pleasure in full penetration, so their sperm had to evolve rocket-propelled arcs? Give me a break.) Consider only the "research" standard here. Is it possible that one man in a hundred actually knows, with statistical accuracy, the average distance traveled by his ejaculate? Yet *The Bell Curve* takes Rushton in full seriousness. (Easterbrook, 1995, p. 36)

The Bell Curve is not merely biased science, *The Bell Curve* is not science at all, given what has been presented here. At best, it is political opportunism, yet it raises fundamental issues for psychology as a discipline. For example, it raises these questions: Why should psychology be used in the schools at all? What is the *benefit* of psychological assessment? But especially, what is the *benefit* of mental measurement to the teaching and learning process for children. Why should mental measurement be used in criminal justice? What is its benefit? Why should mental measurement be used in public welfare? What is its benefit? The first benefit that psychology could offer is to clean its own house, to perform a valuable psychological service by eliminating irrational science, by withholding legitimation from those who practice it. This is not only a matter of science and professionalism, it is also an ethical issue.

So, the fundamental problem with *The Bell Curve* is that it is not science at all, and there is no complicated mystery about where the deficiencies of science are manifest. It is superstition. It is superstition because it is *bad psychology, bad biology, bad genetics, bad anthropology, bad pedagogy, and bad linguistics*, and follows a long tradition of science in support of white supremacy, among other things. In an increasingly radical right-wing political environment it has found a welcome reception. What is more significant, however, is that, just as in the past, *Bell Curve* thinking has met with a serious reception by some members of the psychological elite, and by large numbers of them. *Bell Curve* thinking is mainstream psychology, as far as mental measurement is concerned, if we take a recent authoritative survey of the beliefs of the psychological elite as representative (Snyderman & Rothman, 1990).

Up to this point, I have been addressing matters pertaining to the construction of the practice of the mental measurement enterprise, to validity matters. But these are validity matters that assume the validity of the larger applied mental measurement

paradigm. In the *Bell Curve* and in arguments about the validity of mental measurement, the arguments are bound to the validity of *applied mental measurement*. Diagnosis, classification, forecasting, and treatment recommendations are linked to the IQ tool, presenting a larger paradigm issue than the measurement paradigm alone.

The *Bell Curve* paradigm includes the assumption that the intelligence construct is valid, that intelligence in individuals is fixed or stable, that IQ tests are a valid measure of the construct, that culture and context are irrelevant, that the results of IQ testing have a valid and meaningful use in teaching and learning, and that the results of IQ testing have valid and meaningful uses in key areas of social policy, for example, criminal justice, welfare, teenage pregnancy, and schooling. Directly or indirectly, all of these uses of IQ are related to teaching and learning.

Virtually all of the validity research in this paradigm is correlational. In *Bell Curve* thinking, correlation is causation. Experimental studies that control for opportunities to learn, the quality of instruction, cultural and linguistic diversity, racism, and so forth, are virtually nonexistent, nor do many psychologists possess the expertise to do so. For example, cultural and linguistic expertise is missing, especially in relationship to various ethnic groups.

So as far as IQ testing and IQ thinking and IQ application are concerned, there is far less than meets the eye. Certainly the failure of psychology to balance correlational research with more rigorous controlled experimentation is a scientific deficiency in IQ work of the greatest magnitude.

But I emphasize, the basic problem is less with Murray and Herrnstein and their book, it is with the currently popular paradigm. Let's look at schooling for an example of an alternative paradigm. By the way, if there is no alternative assessment paradigm, then there is no instructionally meaningful use for mental measurement in the schools. The Academy of Sciences Panel on *Placing Children in Special Education: A Strategy for Eq-*

uity (Heller, Holtzman, & Messick, 1982) established an unambiguous and rigorous criterion for judging the value of the use of mental measurement in teaching and learning.

Our ultimate message is a strikingly simple one. The purpose of the entire process—from referral for assessment to eventual placement in special education—is to improve instruction for children. The focus on educational benefits for children became our unifying theme, cutting across disciplinary boundaries and sharply divergent points of view.

These two things—the validity of assessment and the quality of instruction—are the subject of this report. Valid assessment, in our view, is marked by its relevance to and usefulness for instruction (Holtzman, 1982, pp. x, xi).

While academic failures are often attributed to characteristics of learners, current achievement also reflects the opportunities available to learn in school. If such opportunities have been lacking or if the quality of instruction offered varies across subgroups of the school-age population, then school failure and subsequent E.M.R. referral and placement may represent a lack of exposure to quality instruction for disadvantaged or minority children. (Heller et al., 1982, p. 15)

The IQ test's claim to validity rests heavily on its predictive power. We find that prediction alone, however, is insufficient evidence of the test's educational utility. What is needed is evidence that children with scores in the E.M.R. range, learn more effectively in a special program or placement. As argued in more detail in Chapter 4, we doubt that such evidence exists, although we are not prepared as a panel to advocate the discontinuation of IQ tests, we feel that the burden of justification lies with its proponents to show that in particular cases the tests have been used in a manner that contributes to the effectiveness of instruction for the children in question. (Heller et al., 1982, p. 61)

Certainly it is unlikely that psychologists would argue that the use of mental measurement in teaching and learning should *not*

result in benefits to students. So it is this student-benefit criterion for applied mental measurement that changes the nature of the validity debate.

Since this criterion was presented, there have been many more studies of the instructional validity of popularly used mental measurement. The results are overwhelming. Skrtic's (1991) brilliant review is fully consistent with the earlier Nation Academy of Sciences report (Heller et al., 1982) done nearly a decade earlier. Skrtic deconstructed present dialogue and practice in the use of mental measurement in the schools. He provided overwhelming empirical data to challenge each of four grounding assumptions in American education, a system that is driven by IQ thinking. In an exhaustive review of the literature, Skrtic showed that in general, low achievement is not due to students' impairment, that diagnoses of pathology in students are not valid and objective and useful for instructional design, and that the problems of teaching and learning may have little to do with better diagnostics.

An alternative paradigm could be constructed on other assumptions, for example:

1. Cognitive/learning processes can be articulated and *assessed*.
2. Cognitive/learning processes can be changed. "Intelligence" can be taught (Hilliard, 1990a).
3. Culture and context are salient and useful in the assessment and design of valid teaching strategies.

Although far fewer psychologists have operated from a paradigm based upon these assumptions, there are some impressive beginnings (Feuerstein, 1980; Lidz, 1991). I have worked with Feuerstein and his associates and their approach to assessment and mediation and find it to be a powerful way to produce student benefits (i.e., cognitive growth and academic achievement). However, I realize that much more work has to be done, and is being done. Space will not

permit a review of the work that is being done.

The main point is this: Although there are those who will critique the designers of the alternative approaches created by psychologists such as Feuerstein (1979, 1980), Budoff (1974), Sternberg (1982), Dent (1995), Campione, Brown & Ferrara (1985), and so on, either this approach or other approaches must be evaluated according to the *benefits criterion*. Doing assessment for curiosity or for fun may not require the use of this criterion. But *applying* assessment to human problem solving or to public policy prescriptions means that the criterion cannot be avoided.

The Bell Curve gives no evidence whatsoever that the paradigm problem is recognized or understood. Given the fact that *The Bell Curve* and IQ thinking in its currently popular applied form has no track record of benefits, empirically demonstrable, then either there is a beneficial alternative paradigm or there should be no mental measurement in schools or in other areas where teaching and learning is implied. As I have said before, mental measurement at present is not only not beneficial, it is a burden on the educational process. There is a political constituency for it, but there is no valid pedagogical need for it.

Given that *The Bell Curve* is nonbeneficial bad science, it goes without saying that Murray and Herrnstein's voices on social policy are merely old voices that have turned up the volume. We heard these voices in Nazi Germany (Peukert, 1982; Weinreich, 1946). We have heard them from the early days of IQ testing and thinking, following the pre-IQ white supremacy science (Benedict, 1959; Chase, 1977; Gould, 1981; Guthrie, 1976; Kamin, 1974; Thomas & Sillen, 1972).

Then, as now, large numbers of elite psychologists developed a pessimistic view of human potential. Then, as now, they took activists' positions to sterilize the IQ-defined retarded, to save the sperm of high-IQ achievers in sperm banks, to create the

“master race” through eugenics, to segregate the schools, and so on.

Now we are asked to join in a rhetoric, a rhetoric that was heard before *The Bell Curve* was published and goes far beyond it into public policy debates in Congress and in government agencies. The “violence initiative” of the National Institute of Mental Health, the phenomenal growth of the population of children who are said to be impaired learners, the idea that the homeless and the welfare populations do not have the mental capacity to support themselves, the idea that low IQ causes criminal behavior, are all a part of a picture of potential that is being drawn, and is receiving a growing acceptance, in both the professional and lay populations. Many Americans now believe that there is a bottom quarter of us who are worthless people, people who are a drain on the resources of the whole. Murray and Herrnstein said so:

People in the bottom quartile of intelligence are becoming not just increasingly expendable in economic terms; they will sometime in the not-too-distant future become a net drag. In economic terms and barring a profound change in direction for our society, *many people will be unable to perform that function so basic to human dignity: putting more into the world than they take out* [italics added].

Perhaps a revolution in teaching technology will drastically increase the productivity returns to education for the people in the lowest quartile of intelligence, overturning our pessimistic forecast. But there are no harbingers of any such revolution as we write. And unless such a revolution occurs, all the fine rhetoric about “investing in human capital” to “make America competitive in the twenty-first century” is not going to be able to overturn this reality: *For many people, there is nothing they can learn that will repay the cost of the teaching* [italics added].

Is this the Fourth Reich emerging? Should we not be a little bit suspicious when such a large number of the references cited in this book come from research supported by what has

been referred to as the white supremacy Pioneer Fund, and from researchers associated with the white supremacist *Mankind Quarterly*? Should we not be concerned as scholars about the propaganda hype of *Bell Curve* debates on the college campuses funded by a right-wing foundation as a road show?

It is striking that in each age of the advent of *Bell Curve* thinking, what should be regarded as outrageous is tolerated as scholarship. For this, history will not be kind. Those who were silent in the days of Terman, Cyril, Burt, and others [Kamin (1974)] did not meet their academic responsibilities. It is neither Murray nor Herrnstein who are the main players in this drama. We are not, nor can we be, mere spectators in this folly. We must identify ourselves in this pivotal matter. Do more than half of us agree with Murray and Herrnstein that IQ tests are valid? Do we agree that they mean what Murray and Herrnstein said they mean for social policy?

As a former member of the Board of Scientific Affairs Committee on Psychological Testing, I challenge the American Psychological Association to take up the issue of *The Bell Curve* through the testing committee. The world should know where the scientific voice of the profession stands. Is Murray and Herrnstein’s use of IQ testing and thinking valid? Given the worldwide IQ propaganda campaign disseminating *The Bell Curve* currently being waged, we are both professionally and morally obligated to speak up.

Psychologists are, or, in my opinion, ought to be, healers first and foremost. *The Bell Curve* paradigm is not a healing paradigm. IQ as we know it is a tool whose time is past. Yet this is precisely where I entered the fields of psychology and education nearly 40 years ago. I am not shocked to see that Psychology as a political science persists. I am only shocked to see that a bankrupt tool, IQ and the failed paradigm of which it is a part, still holds such sway with no benefits to its subjects.

References and Selected Bibliography

- Ani, M. (1994). *Yurugu: An African centered critique of European thought and behavior*. Trenton, NJ: Africa World Press.
- Backler, A., & Eakin, S. (Eds.). (1993). *Every child can succeed: Readings for school improvement*. Bloomington, IN: Agency for Instructional Technology.
- Baller, W. R., Charles, D. C., & Miller, E. L. (1967). Mid-life attainment of the mentally retarded: A longitudinal study. *Genetic Psychology Monographs*, 75, 235-329.
- Benedict, R. (1959). *Race: Science and politics*. New York: Viking.
- Block, N. J., & Dworkin, G. (1976). *The I.Q. controversy: Critical readings*. New York: Pantheon.
- Bower, B. (1992). Infants signal the birth of knowledge. *Science News*, 142(20), 325.
- Bruer, J. T. (1993, Summer). The mind's journey from novice to expert: If we know the route, we can help students negotiate their way. *American Educator*, 6-46.
- Budoff, M. (1974). *Learning potential and educability among the educable mentally retarded*. Final Report Project No. 312. Cambridge: Research institute for educational problems.
- Campione, J. C., Brown, A. L., & Ferrara, R. A. (1985). Mental retardation and intelligence. In R. J. Sternberg (Ed.), *Human abilities: an information processing approach* (pp. 28-36). New York: W. H. Freeman and Co.
- Cavalli-Sforza, L., Menozzi, P., & Piazza, A. (1994). *The history and geography of human genes*. Princeton, NJ: Princeton University Press.
- Chase, A. (1977). *The legacy of Malthus: The social cost of scientific racism*. New York: Knopf.
- Chomsky, N. (1971). Deep structure, surface structure, and semantic interpretation. In D. Steinberg & L. Jakobovitz (Eds.), *Semantics*. New York: Cambridge University Press.
- Cohen, R. (1969). Conceptual styles, culture conflict, and non-verbal tests of intelligence. *American Anthropologist*, 71(5), 828-857.
- Cohen, R. (1971). The influence of conceptual rule sets on measures of learning ability. In *Race and intelligence*. Anthropological Association.
- Cole, M., Gay, J., Glick, J. A., & Sharp, D. W., (1971). *The cultural context of learning and thinking: An exploration in experimental anthropology*. New York: Basic Books.
- Dent, H. E. (1995). The San Francisco Public Schools experience with alternatives to IQ testing: A model for non-biased assessment. In A. G. Hilliard, III (Ed.), *Testing African American students* (pp. 121-137). Chicago: Third World Press.
- Dillon, R., & Sternberg, R. J. (1986). *Cognition and instruction*. San Diego: Academic Press.
- Donaldson, M. (1978). *Children's minds*. New York: Norton.
- Duke, L. (1991, January 9). Whites racial stereotyping persists: Most retain negative beliefs about minorities, survey finds. *The Washington Post*, p. A1.
- Easterbrook, G. (1995). Blacktop basketball and *The Bell Curve*. In R. Jacoby & N. Glaubergerman (Eds.), *The Bell Curve debate: History, documents, opinion* (pp. 30-43). New York: New York Times Books, Random House.
- Edmonds, R. R. (1979). Some schools work and more can. *Social Policy*, (March/April), 28-32.
- Everson, H. T., & Gordon, E. W. (1995). Cracks in the Bell: Two commentaries on *The Bell Curve: Intelligence and class structure in American life*. New York: The College Board.
- Fairchild, H. H. (1991). Scientific racism: The cloak of objectivity. *Journal of Social Issues*, 47(3), 101-115.
- Feuerstein, R. (1979). *The dynamic assessment of retarded performers: The learning potential assessment device*. Baltimore, MD: University Park Press.
- Feuerstein, R. (1980). *Instrumental enrichment*. Baltimore, MD: University Park Press.
- Feuerstein, R., Klein, P. S., & Tannenbaum, A. J. (1991). *Mediated learning experience (MLE) theoretical, psychological and learning implications*. London: Freund Publishing.
- Fuller, R. (1977). *In search of the IQ correlation: A scientific whodunit*. Stonybrook, NY: Ball-Stick-Bird, Inc.
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. London: Paladin.
- Ginsburg, H. (1972). *The myth of the deprived child: Poor children's intellect*. Englewood Cliffs, NJ: Prentice Hall.
- Glass, G. V. (1983). Effectiveness of special education. *Policy Studies Review*, 2(1), 65-78.

- Gossett, T. F. (1973). *Race: The history of an idea in America*. New York: Schocken.
- Gould, S. J. (1981). *The mismeasure of man*. New York: Norton.
- Guthrie, R. (1976). *Even the rat was white*. New York: Harper & Row.
- Hall, E. T. (1977). *Beyond culture*. New York: Anchor Press.
- Heath, S. B. (1983). *Ways with words*. Cambridge, England: Cambridge University Press.
- Hehir, T., & Latus, T. (Eds.). (1993). *Special education at the century's end: Evolution of theory and practice since 1970* (Harvard Educational Review Reprint Series, No. 23). Cambridge, MA: Harvard University Press.
- Heller, K. A., Holtzman, W. H., Messick, S. (Eds.). (1982). *Placing children in special education: A strategy for equity*. Washington, DC: National Academy Press.
- Helms, J. E. (1992). Why is there no study of cultural equivalence in standardized cognitive ability testing? *American Psychologist*, 47(9), 1083-1101.
- Hilliard, A. G., III. (1976). *Alternatives to IQ testing: An approach to the identification of "gifted" minority children*. Final report to the California State Department of Education, Special Education Support Unit. (ERIC Clearinghouse on Early Childhood Education ED 146 009)
- Hilliard, A. G., III. (1979). The pedagogy of success. In Sunderlin, S., *The most enabling environment* (pp. 43-52). Washington, DC: Association for Early Childhood International.
- Hilliard, A. G., III. (1981). IQ thinking as catechism: Ethnic and cultural bias or invalid science. *Black Books Bulletin*, 7(2), 99-112.
- Hilliard, A. G., III. (1983). Psychological factors associated with language in the education of the African-American child. *Journal of Negro Education*, 52(1), 24-34.
- Hilliard, A. G., III. (1984). IQ thinking as the Emperor's new clothes. In C. Reynolds & R. T. Brown (Eds.), *Perspectives on bias in mental testing* (pp. 139-169). New York: Plenum Press.
- Hilliard, A. G., III. (1987). The learning potential assessment device and instrumental enrichment as a paradigm shift. *The Negro Educational Review*, 38(2-3), 200-208.
- Hilliard, A. G., III. (1990a). Back to Binet: The case against the use of IQ Tests in the schools. *Contemporary Education*, 61(4), 184-189.
- Hilliard, A. G., III. (1990b). Misunderstanding and testing intelligence. In J. Goodlad & P. Keating (Eds.), *Access to knowledge: An agenda for our nation's schools*. New York: The College Board.
- Hilliard, A. G., III. (1989). "Discussant" in Pfliegerer, J. (Ed.), *The Uses of Standardized Tests in American Education*. Proceedings of the fiftieth ETS Invitational Conference. Princeton: Educational Testing Service, 27-36.
- Hilliard, A. G., III. (1991). Do we have the will to educate all children? *Educational Leadership*, 49(1), 31-36.
- Hilliard, A. G., III. (1994). Thinking skills and students placed at greatest risk in the educational system. In *Restructuring learning: 1990 Summer Institute Papers and Recommendations* (pp. 147-156). Washington, DC: Council of Chief State School Officers.
- Hilliard, A. G., III. (1995). *Testing African American students*. Chicago: Third World Press.
- Holtzman, W. H. (1982). Preface. In K. A. Heller, W. H. Holtzman, & S. Messick (Eds.), *Placing children in special education: A strategy for equity*. Washington, DC: National Academy Press.
- Hoover, M. R., Politzer, R. L., & Taylor, O. (1995). Bias in reading tests for Black language speakers: A sociolinguistic perspective. In A. G. Hilliard, III (Ed.), *Testing African American students* (pp. 51-68). Chicago: Third World Press.
- Horowitz, I. L. (1995). The Rushton file. In R. Jacoby & N. Glauberman (Eds.), *The Bell Curve debate: History, documents, opinion* (pp. 179-200). New York: New York Times Books, Random House.
- Houts, P. L. (Eds.). (1977). *The myth of measurability*. New York: Hart Publishing.
- Jacobs, P. (1977). *Up the IQ*. New York: Wyden.
- Jacoby, R., & Glauberman, N. (Eds.). (1995). *The Bell Curve debate: History, documents, opinions*. New York: New York Times Books, Random House.
- Jensen, A. (1969). How much can we boost IQ? *Harvard Educational Review*.
- Jensen, A. (1980). *Bias in mental testing*. New York: Free Press.
- Jensen, M. R. (1992). Principles of change models in schools psychology and education. In *Advances in cognition and educational practice*

- (Vol. 1B, pp. 47–72). Greenwich, CT: JAI Press.
- Kamin, L. (1974). *The science and politics of IQ*. New York: John Wiley & Sons.
- Kluwe, R. H., Misiak, C., & Haider, H. (1991). The control of complex systems and performance in intelligence tests. In H.A.H. Rowe (Ed.), *Intelligence, reconceptualization and measurement* (Australian Council for Educational Research. Hillsdale, NJ: Erlbaum.
- Kozol, J. (1991). *Savage inequalities: Children in America's schools*. New York: Crown.
- Labov, W. (1970). The logic of non-standard English. In F. Williams (Ed.), *Language and poverty*. Chicago: Markham.
- Lane, C. (1995). Tainted sources. In R. Jacoby & N. Glauberman (Eds.), *The Bell Curve debate: History, documents, opinions* (pp. 125–139). New York: New York Times Books, Random House.
- Lewis, M. G. (1995). "We're Family!": *Creating success in an African American public elementary school*. Unpublished doctoral dissertation, Georgia State University, Atlanta.
- Lidz, C. S. (Ed.). (1987). *Dynamic assessment. An interactional approach to evaluating learning potential*. New York: Guilford Press.
- Lidz, C. S. (1991). *Practioner's guide to dynamic assessment*. New York: Guilford Press.
- Lipsky, D. K., & Gartner, A. (1989). *Beyond separate education: Quality education for all*. Baltimore: Paul H. Brookes.
- Mainstream science on intelligence. (1994, December 13). *Wall Street Journal*.
- Montagu, A. (Ed.). (1964). *The concept of race*. London: Collier.
- Montagu, A. (1974). *Man's most dangerous myth: The fallacy of race*. New York: Oxford.
- Oakes, J. (1985). *Keeping track: How schools structure inequality*. New Haven, CT: Yale University Press.
- Peukert, D.J.K. (1982). *Inside Nazi Germany: Conformity, opposition, and racism in everyday life*. New Haven, CT: Yale University Press.
- Richele, M. N. (1991). Reconciling views on intelligence. In H.A.H. Rowe (Ed.), *Intelligence, reconceptualization and measurement* (Australian Council for Educational Research.) Hillsdale, NJ: Erlbaum.
- Rist, R. C. (1973). *The urban school: A factory for failure*. Cambridge, MA: MIT Press.
- Rowe, H.A.H. (Ed.). (1991). *Intelligence, reconceptualization and measurement* (Australian Council for Educational Research). Hillsdale, NJ: Erlbaum.
- Salomon, G. (1995). Reflections on the field of educational psychology by the outgoing journal editor. *Educational Psychologist*, 30(3), 105–108.
- Sedgwick, J. (1995). Inside the Pioneer Fund. In R. Jacoby & N. Glauberman (Eds.), *The Bell Curve debate: History, documents, opinions* (pp. 144–161). New York: New York Times Books, Random House.
- Shuy, R. W. (1977). Quantitative linguistic analysis: A case for and some warnings against. *Anthropology and Education Quarterly*, 1(2), 78–82.
- Sizemore, B. (1988, August). The algebra of African-American achievement. In *Effective schools: Critical issues in the education of Black children* (pp. 123–149). Washington, DC: National Alliance of Black School Educators.
- Sizemore, B., Brosard, C., & Harrigan, B. (1982). *An abashing anomaly: The high achieving predominantly Black elementary schools*. Pittsburgh, PA: University of Pittsburgh Press.
- Skrtic, T. M. (1991). The special education paradox: Equity as the way to excellence. *Harvard Educational Review*, 61(2), 148–206.
- Slack, W. V., & Porter, D. (1980). The Scholastic Aptitude Test: A critical appraisal. *Harvard Educational Review*, 50(2), 154–175.
- Smith, E. A. (1978). *The retention of the phonological, phonemic, and morphophonemic features of Africa in Afro-American ebonics* (Seminar Series Paper No. 40). Fullerton: Department of Linguistics, California State University.
- Smith, E. A. (1979). *A diagnostic instrument for assessing the phonological competence and performance of the inner-city Afro-American child* (Seminar Series Paper No. 41). Fullerton, CA: Department of Linguistics, California State University.
- Snyderman, M., & Rothman, S. (1990). *The IQ controversy: The media and public policy*. New Brunswick NJ: Transaction.
- Spitz, H. H. (1986). *The raising of intelligence: A selected history of attempts to raise retarded intelligence*. Hillsdale, NJ: Erlbaum.
- Sternberg, R. J. (Ed.). (1982). *Handbook of human intelligence*. New York: Cambridge University Press.

- Subramanian, S. (1995, January 16). The story in our genes. *Time*, pp. 54–55.
- Suzuki, S. (1984). *Nurtured by love: The classic approach to talent education*. Smithtown, NY: Exposition Press.
- Thomas, A. & Sillen, S. (1972). *Racism and psychiatry*. New York: Brunner/Mazel.
- Weinreich, M. (1946). *Hitler's professors: The part of scholarship in Germany's crimes against the Jewish people*. New York: Yiddish Scientific Institute-YIVO.
- Wigdor, A. K., & Garner, W. K. (Eds.). (1982). *Ability testing: Uses, consequences, and controversy*. Washington, DC: National Academy Press.
- Wiggins, G. P. (1993). *Assessing student performance: Exploring the purpose and limits of testing*. San Francisco: Jossey-Bass.
- Yee, A. H. (1983). Ethnicity and race: Psychological perspectives. *Educational psychologist*, 18(1), 14–24.
- Zacharias, J. R. (1977). The trouble with tests. In P. L. Houts (Ed.), *The myth of measurability*. New York: Hart Publishing.
- Curtin, P. D. (1964). *The image of Africa: British ideas and action, 1780–1850* (Vols. 1 & 2). Madison: University of Wisconsin Press.
- De La Luz Reyes, M., & Halcon, J. J. (1988). Racism in academia: The old wolf revisited. *Harvard Educational Review*, 58(3), 299–314.
- Diop, C. A. (1974). *The African origin of civilization: Myth or reality*. Westport, CT: Lawrence Hill. [See, especially, chap. 3: "Modern Falsification of History," pp. 43–84.]
- DuBois, W.E.B. (1973). *Black reconstruction in America: An essay toward a history of the part which black folk played in the attempt to reconstruct democracy in America 1860–1880*. New York: Antheneum. [See, especially, chap. 17: "The Propaganda of History," pp. 711–730.]
- Fairchild, H. H. (1991). Scientific racism: The cloak of objectivity. *Journal of Social Issues*, 47(3), 101–115.
- Forbes, J. D. (1977). *Racism, scholarship and cultural pluralism in higher education*. Davis: University of California, Native American Studies Tecumseh Center.
- Frederickson, G. M. (1971). *The Black image in the White mind: The debate on Afro-American character and destiny, 1817–1914*. New York: Harper Torchbooks.
- Gossett, T. F. (1973). *Race: The history of an idea in America*. New York: Schocken.
- Gould, S. J. (1981). *The mismeasure of man*. New York: Norton.
- Guthrie, R. (1976). *Even the rat was white*. New York: Harper & Row.
- Hodge, J. L., Struckmann, D. K., & Troust, L. D. (1975). *Cultural basis of racism and group oppression: An examination of traditional "Western" concepts, values, and institutional structures which support racism, sexism and elitism*. Berkeley, CA: Two Riders Press.
- Hofstadter, R. (1955). *Social Darwinism in American thought*. Boston: Beacon Press.
- Jones, S. M. (1972). The Schlesingers on Black history. *Phylon, the Atlanta University Review of Race and Culture*, 33(2), 104–111.
- Joseph, G. G., Reddy, V., & Searel-Chattejee, M. (1990). Eurocentrism in the social sciences. *Race and Class*, 31(4), 1–26.
- Kamin, L. (1974). *The science and politics of IQ*. New York: John Wiley & Sons.
- Ladner, J. (Ed.). (1973). *The death of White sociology*. New York: Vintage.

The Scholar as Racist/White Supremacist

- Alan, C. (1992, January 5). Gray matter, Black-and-White controversy. *Insight—Washington Times*, pp. 4–9, 24–28.
- Ani, M. (1994). *Yurugu: An African centered critique of European thought and behavior*. Trenton, NJ: Africa World Press.
- Baratz, J., & Baratz, S. (1970). Early childhood intervention: The social science base of institutional racism. *Harvard Education Review*, 40, 29–50.
- Benedict, R. (1959). *Race: Science and politics*. New York: Viking.
- Biddis, M. D. (1970). *Father of racist ideology: The social and political thought of Count Gobineau*. New York: Weinright & Talley.
- (1987). Historians blamed for perpetuating bias, outgoing president challenges colleagues. *Black Issues in Higher Education*, 4(4), 2.
- Carruthers, J. (1983). *Science and oppression*. Chicago: Center for Innercity Studies.
- Chase, A. (1977). *The legacy of Malthus: The social cost of scientific racism*. New York: Alfred.

- Lewis, D. (1973). Anthropology and colonialism. *Current Anthropology*, 14(5), 581–602.
- Montague, A. (1968). *The concept of the primitive*. New York: Free Press.
- Murray, C., & Herrnstein, R. (1994). *The Bell Curve*. New York: Free Press.
- Pearce, R. H. (1965). *Savagism and civilization: A study of the Indian and the American mind*. Baltimore: Johns Hopkins Press.
- Peukert, D.J.K. (1982). *Inside Nazi Germany: Conformity, opposition, and racism in everyday life*. New Haven, CT: Yale University Press.
- Pine, J., & Hilliard, A. G., III. (1990). RX for racism: Imperatives for America's schools. *Phi Delta Kappa*, 71(8), 593–600.
- P'Bitek, O. (1990). *African religions in European scholarship*. New York: ECA Associates. (original work published 1970)
- Poliakov, L. (1974). *The Aryan myth: A history of racist and nationalist ideas in Europe*. New York: Meridian.
- Quigley, C. (1981). *The Anglo-American establishment: From Rhodes to Cliveden*. New York: Books in Focus.
- Schwartz, B. N., & Disch, R. (1970). *White racism: Its history, pathology, and practice*. New York: Dell.
- Snyderman, M., & Rothman, S. (1990). *The IQ controversy: The media and public policy*. New Brunswick, NJ: Transaction.
- Stanton, W. (1960). *The leopard's spots: Scientific attitudes toward race in America, 1915–1959*. Chicago: University of Chicago Press.
- Thomas, A., & Sillen, S. (1972). *Racism and psychiatry*. New York: Brunner/Mazel.
- Walker, S. S. (1981). *Images of Africa in the Oakland Public Schools: Assessment of educational media resources pertaining to Africa*. Berkeley, CA: School of Education, Professional Development and Applied Research Center.
- Wa Thiong'o, N. (1987). *Decolonizing the mind: The politics of language in African literature*. London: James Currey.
- Weinreich, M. (1946). *Hitler's professors: The part of scholarship in Germany's crimes against the Jewish people*. New York: Yiddish Scientific Institute–YIVO.
- Wilson, C. C., II, & Gutiérrez, F. (1987). *Minorities and media: Diversity and the end of mass communication*. Beverly Hills, CA: Sage.