PHI DELTA KAPPAN

JANUARY 1998

The Social Consequences Of Bad Research

Daniel Tanner

Illustration by Susan Todd

The Social Consequences Of Bad Research

By Daniel Tanner

In many cases it would seem that school practitioners can test research findings for practicability and generalizability more effectively than some of the editors of the leading educational research journals, Mr. Tanner contends.

THERE IS a sacred taboo in social research against addressing the problem of research bias. Now the word *taboo* comes from Tongan, the Polynesian language of the Tongans. As the anthropologists tell us, a taboo is the sacred and primitive tribal prohibition of an action or word — under pain of supernatural punishment. Hence I realize that I may be putting myself in mortal peril — if not from supernatural forces, then from the natural forces of professional tribal politics. However, if you'll pardon the pun, I may be saved by the "bell curve."

We are all too familiar with the relentless attacks on the public schools over the past half century by national commissions, by the mass media, by special-interest groups, and by individual school blamers. These attacks have come from within the profession as well as from outside it and from all parts of the political spectrum. Traditionally, the school blamers required no hard data to support their allegations and indictments, and politicians from the White House to the state house could blame the public schools for any and every shortcoming, even as they cut the allocations for elementary and secondary education.

Since the time of the so-called Coleman Report of 1966, educational researchers have been busy generating from their burgeoning data banks a deluge of increasingly elaborate empirical studies that examine the limitations of schooling rather than its potential for furthering social progress.[1] Gunnar Myrdal pointed out that the direction and findings of social research "normally come from the political interests that dominate the society in which we live."[2] Hence we find educational researchers at times moving as a flock in generating statistical data that portray the public schools as either utterly anemic[3] or beset by problems of such crippling proportions "that our entire public education system is nearing collapse."[4]

To many researchers and education policy makers the data would seem to indicate that the public schools are virtually brain dead. Myron Lieberman has gone a step further in the title of his book published by Harvard University Press: *Public Education: An Autopsy*.[5] With regard to the diagnoses that find the public schools moribund, as a longtime movie buff I cannot help but recall the scene from *Horse Feathers* in which Groucho Marx as Dr. Hackenbush is taking someone's pulse and exclaims, "Either this man is dead or my watch has stopped."

Groucho was indeed using the right instrument to measure the patient's pulse, but merely using the right instrument is no guarantee of valid conclusions. To put the matter more soberly, as Myrdal did, the cultural and political setting will often tempt researchers "to aim opportunistically for conclusions that fit prejudices" and will shape "the approaches we choose in research, . . . the concepts, models, and theories we use, and the way in which we select and arrange our observations and present the results of our research."[6] The consequence, Myrdal argued, is systematic bias in our work — even in our manmade programming of computing machines.[7]

In recent years, as reflected by the journal *Educational Researcher*, a number of members of the American Educational Research Association have built their careers on the debate between qualitative research and quantitative research.[8] I had believed that John Dewey had dumped that issue into its final resting place back in 1929 when he wrote in *The Sources of a Science of Education* that all good research must be built on powerful qualitative ideas, that elaborate statistical measurements will have little scientific value when the problems under investigation lack generalized significance, and, of course, that the same is true for nonstatistical research.[9]

Abandoning the Great Tradition

Regardless of the false dualism that divides social research into quantitative and qualitative camps, social researchers of either ilk appear to have abandoned what Myrdal called "the great tradition in social science" — namely, "for the social scientists to take a direct as well as indirect responsibility for popular education."[10]

Not only have social scientists largely abandoned their responsibility for popular education, but they have focused primarily on the limitations of public elemen-

DANIEL TANNER is a professor in the Graduate School of Education, Rutgers University, New Brunswick, N.J.

MANY RESEARCHERS CHOOSE THE INSTRUMENT AND SET THE CONDITIONS IN AN EFFORT TO PROVE THEIR PREMISES, OR THEY SET THE STAGE FOR MEASURING THE EASILY MEASURED.

tary and secondary education. One has only to look at the research generated over a period extending from Christopher Jencks' *Inequality* in 1972, which launched many careers on "no school effects" research, to the *Bell Curve* of 1994.[11] In effect, whether opportunistically or unwittingly, many of the most vocal and visible members of the social science research community have fallen in league with the political Right on issues of education policy.

With the findings of the Jencks study widely endorsed as valid by leading social science researchers, President Nixon's Commission on School Finance used Jencks' work to argue that schooling is inconsequential in the lives of other people's children — namely, the disadvantaged. Daniel Patrick Moynihan, then a professor of education and urban politics at the Kennedy School of Government at Harvard University and a consultant to the President, cited the report of the commission in holding that "with respect to school finance there is the strong possibility that we may be already spending too much."[12] Moynihan then concluded that "a final fact is that at a point, school expenditure does not seem to have any notable influence on school achievement," and "this discovery was one of the major events in large-scale social science."[13]

Can anyone imagine our professoriate in the social sciences generating research showing that a college education makes little difference in people's lives and that we may well be allocating too much public funding for higher education? On the contrary, the research findings on the effects of a college education on people's lives are invariably positive and judged to warrant increased expenditures.[14]

Seeming to anticipate the "no school effects" research, Gunnar Myrdal commented in 1968 that "there is truth in the biblical saying that, 'He that seeketh, findeth'; if a scientist seeks what isn't there, he will find it." He will find it, Myrdal continued, "as long as empirical data are scanty and he allows his logic to be twisted."[15] On the research problem of finding what isn't there, the humorist Robert Benchley once cautioned, "You can't prove that platypuses don't lay eggs by photographing platypuses *not* laying eggs." Hence, even though more than two-thirds of the variance in the research on school effects was unaccounted for, the researchers proceeded to treat this as a negative finding on school effects. In essence, the researchers prove the case by *not* finding something.

The lead article in the April 1994 issue of *Educational Researcher* bears the title "Does Money Matter?"[16] Just ask your spouse, neighbor, banker, dean, or university president — not to mention the Pentagon. I recall a TV interview some years ago in which a U.S. senator offered this candid comment: "There are two things that are important in politics. The first is money, and I can't remember what the other one is." Incredibly, a quarter of a century after Jencks, we are still debating whether money matters in public elementary and secondary education. Somewhat belatedly the meta-analysis in the April 1994 issue of *Educational Researcher* found that, contrary to most of the earlier studies, the relationship between resource inputs and school outcomes is systematically positive — something that every teacher and parent has always known.[17] However, when the prevailing political policy is one of social and educational retrenchment, opportunity strikes for researchers to garner grants for themselves and to build their reputations by shattering the conventional wisdom.

According to Myrdal, our elaborate statistical techniques for generating and interpreting data often make our social research even more susceptible to bias. He argues that we need to put our value premises up front and put our research to the test of relevance and practical significance to our democratic social ideals.[18] As I noted above, many of those who posture as advocates of the disadvantaged have been quick to endorse the kind of research and public policy that begins with the limitations of schooling as opposed to the potentials.[19]

Returning to our good friend Groucho, Dr. Hackenbush may indeed have been using the right instrument and technique for taking the measurement, but his faulty premises guaranteed false findings. However, many researchers choose the instrument and set the conditions in an effort to prove their premises, or they set the stage for measuring that which is most easily measured. Albert Einstein put the matter quite clearly when he commented, "I have little patience with scientists who take a board of wood, look for the thinnest part, and drill a great number of holes where drilling is easy."[20]

Finding no connections is far easier than finding connections. The use of elaborate statistical techniques and hyper-abstract theoretical models to create the impression of scientific objectivity and to mask judgments and questionable premises is coupled with a cascade of unnecessarily technical and esoteric jargon that serves only to obscure issues, to mask the poverty of ideas underlying the research, and to isolate the communication within the cocoons of narrow academic specialties. In Myrdal's words, "While a great tradition in social science was to express reasoning as clearly and succinctly as possible, the tendency in recent decades has been for social scientists to close themselves off by means of unnecessarily elaborate and strange terminology, often to the point of impairing their ability to understand one another — perhaps occasionally even themselves."[21] This is no less the case with much so-called qualitative research.

The situation is such, Myrdal noted,

that obscure technical language has become the instrument for spawning exercises in hyper-abstraction that spiral ever upward. One who is able to express ideas directly and with extraordinary clarity, such as John Kenneth Galbraith, observed Myrdal, is eyed with suspicion by his more pedestrian colleagues who refuse to recognize the contributions to knowledge that he has made.[22]

Politics of Standards

Recently, I was looking at an oversized, multicolored foldout from a publication issued in 1991 by the Council for Basic Education (CBE). It identified standards by grade level in each of the academic subjects. According to the CBE, the standards were gleaned from those developed by the national professional associations in the various academic disciplines and from state curricular frameworks, and, in the words of the CBE, they "represent the best wisdom currently available from the field."[23]

Now, get a load of this standard in science, cited by the CBE as formulated at the English Coalition Conference and issued by the National Council of Teachers of English and the Modern Language Association. It is to be attained by fourth-graders in the year 2000: fourth-grade students will "write and speak eloquently about observations and experiments." (I can only assume that the underlying wisdom of those who formulated this standard was that the fourth-graders of 1991 would still be in the fourth grade in the year 2000, struggling to meet this standard.) I ask you, how many academicians are able to "write and speak eloquently about observations and experiments"?

The current national standards movement can be traced to the failure of the nation's political leadership to meet the national goals that grew out of the Education Summit of governors, which was convened by President Bush in 1989. As a result of that summit, the Bush Administration produced its education strategy, called America 2000. The top goal was that, by the year 2000, "All children will start school ready to learn."[24] Back in 1989 the year 2000 seemed far off in the future — certainly beyond the governors' terms of office. So the current crop of governors and President Clinton shifted conveniently to rigorous and narrowly defined national standards for student testing. In effect, they sidestepped what was once the number-one national goal (which would have required massive social reconstruction on a scale comparable to the New Deal) and transferred the responsibility and accountability to "Mrs. Jones and her kindergartners."

On 8 September 1993 U.S. Secretary of Education Richard Riley held a press conference to release *Adult Literacy in America*, a study conducted by the Educational Testing Service (ETS) for the National Center for Education Statistics.[25] All three major network news programs that evening opened with an alarming story taken from Secretary Riley's press release declaring that most U.S. adults can't read. The following day, the front page of the *New York Times* carried a story headlined "Study Says Half of Adults in U.S. Can't Read or Handle Arithmetic."[26]

It doesn't take much intelligence to eye such sweeping pronouncements with suspicion. Yet the mass media immediately seized upon a report that made sensational copy. At the same time, the research methodology and findings in the report were accepted without challenge in academe. Here was yet another national report to use in cannonading our schools — a cannonade issued almost exactly a decade after *A Nation at Risk*.[27]

Even a cursory review of the premises and research methodology of the ETS national literacy study reveals gaping flaws that should have raised questions about the underlying motives of the research sponsors. First, the definition of literacy formulated for the study was utterly different from any that had been used previously by the U.S. Bureau of the Census or by any other national or international agency. For the ETS study, literacy was defined as "using printed and written information to function in society, to achieve one's goals, and to develop one's knowledge potential."[28] Aside from the ambiguity of the definition, it is doubtful that very many people in our society would claim that they have achieved their goals and have developed their knowledge potential, let alone voluntarily agree to submit to a lengthy written test to demonstrate that proficiency or to risk revealing their level of incompetence.

Adult Literacy in America claimed that it was based on a nationally representative sample of 13,000 adults (16 years of age or older) who submitted to a written test, coupled with interviews and tests conducted in 27,000 households in the U.S. In making a few simple calculations of my own, I find that the proportion of immigrants in the national sample far exceeded the proportional representation of immigrants in the national population. What's more, most of these immigrants had never attended U.S. schools. Also grossly overrepresented were disadvantaged minorities whose years of schooling, test scores, and economic conditions are below those of the national population. Furthermore, almost one in 10 males in the study were inmates of federal and state penitentiaries — hardly in a position "to function in society, to achieve one's goals, and to develop one's knowledge potential." The proportion of males incarcerated in federal and state prisons is actually less than 0.8%. Penitentiary inmates were overrepresented in this "scientific" sampling by almost 1,200%. The proportion of female inmates was also grossly overrepresented in the sample.

As expected, the test scores were positively and strongly correlated with years of schooling in the U.S. Of the 25% of residents in the 27,000 households who tested at the lowest level, almost one in five suffered from visual difficulties that impaired their reading of ordinary print materials under ordinary lighting conditions, let alone under the conditions of taking a timed test under the supervision of a stranger — the stranger being an ETS examiner in their home. In other words, they met the legal definition for being blind. More than one in four had physical, mental, or health difficulties that prevented them from participating in regular work, school attendance, housework, or other activities.

Furthermore, the report failed to indicate the percentages of people who declined to submit to the tests and to the questionnaires and interviews that required more than an hour of time under direct supervision. One might reasonably infer that busy people are less likely to make themselves available for such a test. Aside from the inherent biases in the research design, this factor may help account for the disproportionate representation of the incarcerated, the homebound, the debilitated, and the unemployed among subjects in the study.

It must also be acknowledged that one surely needs to know how to take a multiple-choice test to get through school or college, but one hardly encounters such

tests in real life. Hence we must question whether such a test is a fair simulation of life conditions.

The validity of many of the test items must be regarded as suspect to say the least. In TV interviews, school critics made much of the finding that most adults in the study had difficulty deciphering a bus schedule. The schedule contained numerous distractors and conditional information in a hypothetical situation.[29] The school critics proceeded to declare that 80% of adults are unable to figure out which bus will get them home. To use Groucho's diagnosis, 80% of American adults are not finding their way home each day. Common sense should recoil at the way so-called authorities fashion their factoids in support of their biases — never submitting themselves to the test of the method of intelligence.

I am often amused when I should perhaps be appalled at seeing William Bennett or Gore Vidal, interviewed on TV as authors of national best-selling books, attacking the public schools for the allegedly endemic illiteracy of the adult American population. If people can't read, who is buying Bennett's *Book of Virtues* or Vidal's latest best seller? The fact is that, despite the concern that television and computer screens are distracting the population from old-fashioned books, the number of books sold annually per capita has not undergone a decline as predicted by many pundits, and Americans continue to spend more on books than on almost any other medium.[30]

We are all familiar with the impact of revisionism on historical and social science research, including education. It has become increasingly fashionable in some circles of the social sciences to build reputations and to convey the impression of scientific inquiry by generating hard data so as to overthrow conventional wisdom. Indeed, those researchers who have made use of elaborate mathematical and statistical techniques and computer-generated models of analysis have intimidated many of their colleagues, while others have taken the research seriously.

A notable example is the two-volume study by two noted economists, Robert Fogel of the University of Chicago and Stanley Engerman of the University of Rochester, published in 1974 under the title *Time on the Cross*. Fogel and Engerman set out to examine the economics of American slavery through advanced statistical techniques used by those who call themselves "econometric historians" and "cliometricians." Fogel and Engerman amassed data proving that "the slave diet was not only adequate, it actually exceeded modern recommended daily levels of the chief nutrients."[31] Among their other findings was that "the slave mortality rate in childbearing was lower than the maternal death rate experienced by southern white women."[32] Fogel and Engerman presented data to support the finding that "the average daily diet of slaves was quite substantial" and that "the energy value of their diet exceeded that of free men in 1879 by more than 10%."[33] Their reasoning was based on a comparison of the nutritive value of sweet potatoes, a staple of the slave diet, against that of white potatoes, a staple of the diet of the white population. Fogel and Engerman concluded further that "the material conditions of the lives of slaves compared favorably with those of free industrial workers" and that "over the course of his lifetime, the typical slave field hand received about 90% of the income he produced."[34]

Following the perverse premises of Fogel and Engerman and employing their analytical techniques, one can easily amass data to prove that a herd of milk cows on modern dairy farms enjoys a far better level of care and nutrition and a lower mortality rate in calf-bearing than does the general human population in child-bearing and that the cows receive, in the care given them, the equivalent of 90% of the income they produce. This allows the farmer a net profit of 10%. Of course, this defies all sense and sensibility, but it serves to illustrate how statistical data can be used to validate research premises that are dead wrong to begin with.

The Fogelman/Engerman study is another example of generating elaborate quantitative measures to support a perverse qualitative idea. One only has to read the eloquent indictment of slavery in Darwin's *Voyage of the Beagle* to realize that no amount of statistical data can convey the realities of slavery more scientifically than the impassioned words of Darwin.

Practicability and Generalizability

In an age of mounting specialism in research, it has been said that we are learning more and more about less and less. This has commonly been called the "knowledge explosion." Consequently, we have neglected the synthesis of knowledge that would allow us to generalize and use our knowledge to help solve practical problems of social significance. The nation — and indeed the world — suffers not from a knowledge explosion but from a problems explosion.

Professional schools, including schools of education, have a mandate to engage in research and to advance professional practice with the mission of improving society. This mission is not necessarily shared by all academic disciplines, and for this reason universities have schools of medicine, public health, social work, architecture, urban planning, engineering, agriculture, education, and so on. Unfortunately, much that is taken for social research serves no social purpose other than to embellish reputations in the citadels of academe and sometimes even to undermine the democratic public interest.

It is my contention that school practitioners and researchers should be able to evaluate the efficacy of educational research and to guide their own practice through a commitment to the best available evidence. In many cases it would seem that school practitioners can test research findings for practicability and generalizability more effectively than some of the editors of the leading educational research journals. A review of these journals reveals that many of the articles fail to meet the tests of practicability and generalizability.

For example, consider a not at all atypical lead article in the spring 1992 issue of the *American Educational Research Journal*, which addressed the problem of giving insufficient emphasis to higher-order thinking and problem solving in the classroom. Now this is clearly one of our most significant and pervasive curriculum problems. However, the author of the article, a Stanford professor and theorist in organizational behavior in education, proceeded to attack the problem from the vantage point of behavioral decision theory and the economics of organizations, rather than to treat it as a curriculum problem. As a consequence, the organizational theorist came up with a bizarre solution — a solution that nevertheless warranted its choice by the editors as the lead article. The organizational theorist proposed to have two types of teachers in the school: one type would specialize in teaching the basic skills to ensure that students at least develop these lower-order skills; the second type would specialize in higher-order

thinking and problem solving. "We argue," she wrote, "that such a division would result not only in higher-quality education, but also in a more equitable delivery of services from classroom to classroom."[35]

Just imagine being a member of a school faculty divided evenly between those designated teachers of lower-order thinking and those annointed teachers of higher-order thinking and problem solving! Aside from the disastrous impact on faculty morale, there are the predictable perverse effects on the curriculum and learning — predictable because there is a vast body of curriculum research revealing conclusively that such a dualism between skills and thinking is counterproductive. From the time of the Winnetka Plan of the early 1920s, when such a division was actually instituted schoolwide,[36] to contemporary work in areas such as "writing across the curriculum," failure is predictable when skills are severed from the ideas that govern thinking. Unfortunately, the organizational theorist formulated a solution without having reviewed the pertinent curriculum research literature.[37] Had she done so, she would have found a powerful and consistent pattern of evidence amassed over a period spanning three-quarters of a century that clearly invalidates her line of inquiry.[38] Actually, ordinary school experience and common sense should have been sufficient for the organizational theorist to realize that her line of inquiry was invalid from the beginning.

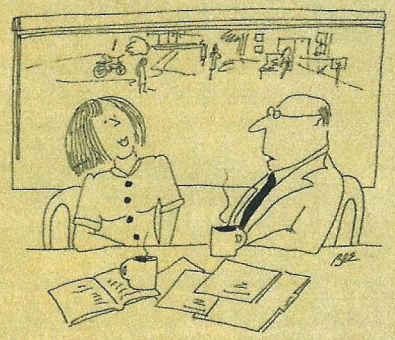
Unfortunately, in the social sciences and education, research reports in our journals rarely go back more than five years in reviewing the pertinent literature. J. Robert Oppenheimer commented, "The openness of this world derives its character from the irreversibility of learning; what is once learned is part of human life. We cannot close our minds to discovery."[39]

In science, once something has been discovered, it might later be disproved, but it does not need to be discovered anew. In education, we have a vast and rich knowledge base on which to build. If we do not build on that base, our research and school practices will shift unwittingly with whatever sociopolitical tide is dominant. The capacity to build on and draw from the knowledge base requires that our theory be tested continually for its power for generalizability and practicability in a wide range of situations. In social research, theory must have the generative power for revealing useful pathways to solutions of social significance.

Early in this century, John Dewey warned that educational *practices* must be the source of the ultimate problems to be investigated if we are to build a science of education.[40] We may draw from the behavioral sciences, but the behavioral sciences do not define the educational problems. The faculties of the professional schools draw on the basic sciences and behavioral sciences, but their mandate is mission-oriented, not discipline-centered. Hence they must, in Dewey's words, "operate through their own ideas, plannings, observations, judgments."[41] To do otherwise in education "is to surrender the education cause."[42]

---

1. James S. Coleman et al., *Equality of Educational Opportunity* (Washington, D.C.: Office of Education, U.S. Department of Health, Education, and Welfare, 1966); and Christopher Jencks et al., *Inequality: A Reassessment of the Effect of Family and Schooling in America* (New York: Basic Books, 1972.)
2. Gunnar Myrdal, *Objectivity in Social Research* (New York: Pantheon, 1969), p. 48.
3. Jencks et al., op. cit.
4. John I. Goodlad, *A Place Called School* (New York: McGraw-Hill, 1984), p. 1.
5. Myron Lieberman, *Public Education: An Autopsy* (Cambridge, Mass.: Harvard University Press, 1993).
6. Myrdal, p. 49.
7. Ibid., p. 44.
8. Larry V. Hedges, Richard D. Laine, and Rob Greenwald, "Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes," *Educational Researcher*, April 1994, pp. 5-14; and Eric A. Hanushek, "Money Might Matter Somewhere: A Response to Hedges, Laine, and Greenwald," *Educational Researcher*, May 1994, pp. 5-8.
9. John Dewey, *The Sources of a Science of Education* (New York: Liveright, 1929).
10. Myrdal, p. 41.
11. Jencks, op. cit.; and Richard J. Herrnstein and Charles A. Murray, *The Bell Curve: Intelligence and Class Structure in American Life* (New York: Free Press, 1994).
12. Daniel P. Moynihan, "Equalizing Education: In Whose Benefit?," *The Public Interest*, Fall 1972, p. 71.
13. Ibid., p. 73.
14. Alexander W. Astin, *What Matters Most in College* (San Francisco: Jossey-Bass, 1993); and Ernest T. Pascarella and Patrick T. Terenzini, *How College Affects Students* (San Francisco: Jossey-Bass, 1991).
15. Myrdal, p. 41.
16. Hedges et al., op. cit.
17. Ibid.
18. Myrdal, pp. 59, 63.
19. In addition to Moynihan and Jencks, I would also cite Samuel Bowles and Herbert H. Gintis, *Schooling in Capitalist America* (New York: Basic Books, 1976).
20. Albert Einstein, *Out of My Later Years* (New York: Philosophical Library, 1950), p. 63.
21. Myrdal, p. 42.
22. Gunnar Myrdal, "How Scientific Are the Social Sciences?," *Bulletin of the Atomic Scientists*, January 1973, pp. 31-37.
23. Council for Basic Education, "Standards: A Vision for Learning," *Perspective*, 1991. Table insert, n.p.
24. Ibid., p. 3.
25. Irwin S. Kirsch, Louis J. Jungeblut, and Andrew Kolstad, *Adult Literacy in America* (Washington, D.C.: U.S. Department of Education, 1993).
26. William Celis, "Study Says Half of Adults in U.S. Can't Read or Handle Arithmetic," *New York Times*, 9 September 1993, p. A-1.
27. National Commission on Excellence in Education, *A Nation at Risk: The Imperative for Educational Reform* (Washington, D.C.: U.S. Department of Education, 1983).
28. Kirsch, Jungeblut, and Kolstad, p. 2.
29. Ibid., p. 91.
30. *Book Industry Trends, 1996* (New York: Book Industry Study Group, 1996).
31. Robert W. Fogel and Stanley L. Engerman, *Time on the Cross: The Economics of Negro Slavery* (Boston: Little, Brown, 1974), p. 115.
32. Ibid., p. 123.
33. Ibid., p. 113.
34. Ibid., pp. 5-6.
35. Jane Hannaway, "Higher-Order Skills, Job Design, and Incentives: An Analysis and Proposal," *American Educational Research Journal*, Spring 1992, p. 20.
36. Carleton W. Washburne and Sidney P. Marland, Jr., *Winnetka: The History and Significance of an Educational Experiment* (Englewood Cliffs, N.J.: Prentice-Hall, 1963).
37. John I. Goodlad and Zhixin Su, "Organization of the Curriculum," in Philip W. Jackson, ed., *Handbook of Research on Curriculum* (New York: Macmillan, 1992), pp. 327-44; and Wilford M. Aikin, *The Story of the Eight-Year Study* (New York: Harper, 1942).
38. Daniel Tanner and Laurel Tanner, *History of the School Curriculum* (New York: Macmillan, 1990).
39. J. Robert Oppenheimer, "Prospects in the Arts and Sciences," in *Man's Right to Knowledge* (New York: Columbia University Press, 1955), p. 113.
40. Dewey, p. 33.
41. Ibid., p. 76.
42. Ibid., p. 74.



"We're a prestigious department, yet we have only three active 'prima donnas.'"