The Sciences (May/June 1987).

THE NINE LIVES OF DISCREDITED DATA

Old Textbooks Never Die—They Just Get Paraphrased

by DIANE B. PAUL

ONE OF THE LESSONS today's undergraduate science majors may glean from their genetics textbooks is that differences in intelligence, as measured by IQ, are due primarily to differences in genes. Students consulting the latest edition of John B. Jenkins's *Human Genetics* (1983), a current best-seller, will learn that "the genotype has a greater influence on IQ than do environmental factors." And those studying H. Eldon Sutton's *Introduction to Human Genetics* (1980) will be given to understand that IQ variations are "largely genetic in whites."

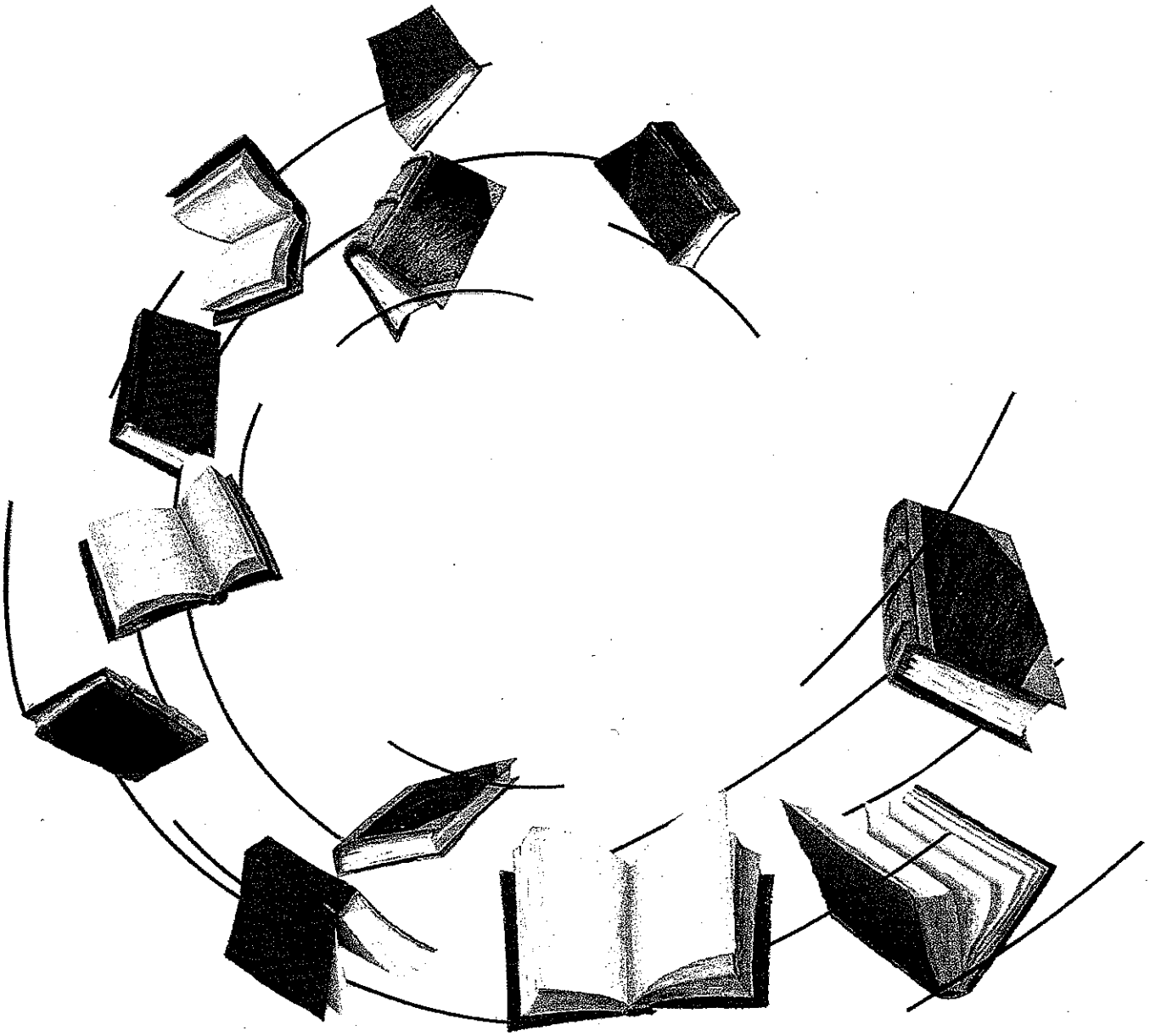
Fifteen years ago, few geneticists would have argued with such assertions. That IQ is seventy to eighty percent heritable seemed indisputable in the light of experiments performed by the English psychometrician Cyril Burt on groups of identical twins. In five studies published between 1955 and 1966, Burt and his collaborators reported that the IQ scores of identical twins were always closely matched, whether the twins had grown up together or apart. His results were authoritative—no other investigator had claimed such success at tracking down twins who had been separated at birth and reared in different environments—and they seemed conclusive. There was just one problem: Burt's impressive findings were fraudulent.

Suspicions were aroused in 1972 (a year after Burt's death), when Leon Kamin, a Princeton psychologist, noted that Burt's IQ correlations (0.771 for twins raised separately, 0.944 for those raised together) had remained constant throughout various studies involving different numbers of subjects—nothing short of a statistical miracle. Then, in 1976, Oliver Gillie, a medical correspondent for *The Sunday Times* of London, reported evidence that Burt had invented his ostensible research collaborators, "J. Conway and M. Howard," and had fabricated some of his data. These charges were not proved until 1979, when Burt's biographer, Leslie Hearnshaw, confirmed by examining personal diaries that Burt had never conducted many of the studies he reported. But no one had doubted since the mid-seventies that, whether through incompetence or chicanery, Burt's work was tainted.

The Cyril Burt scandal was widely publicized, and by the time it was over it had focused attention not only on Burt's methods but also on those of other investigators. Scientists differed, often bitterly, on whether methodological difficulties doomed any and all attempts to measure the heritability of human mental traits; critics insisted—and still do—that even if genes affect variation in IQ, the "genetic" and "environmental" contributions to such variation are too intricately entangled to separate and measure with precision. But even the advocates of such research came generally to agree that many of the early experiments had been flawed, and a new generation of studies was launched. Today, some scholars maintain that the new studies have confirmed a link between genetic and intellectual variation, though a weaker link than the early research suggested. Others remain skeptical. In any case, responsible scholars are no longer basing their claims solely on the early research.

Authors of genetics textbooks responded to these events in a curious way: they stopped citing Burt as an authority, but many continued to cite his results. In a study of twenty-eight texts published between 1978 and 1984, I found that most of the nineteen discussing the heritability of IQ assert that it is high. As evidence, eleven of these texts cite a review article, published in the journal *Science* in 1963, in which L. Erlenmeyer-Kimling and Lissy F. Jarvik incorporated results from fifty-two early studies into a figure indicating a strong inverse correlation between IQ variations and degrees of genetic relatedness. Had the authors of these textbooks read the review article closely, they would have noticed that it included Cyril Burt's results. Yet most (eight out of the eleven) went so far as to reproduce the figure that accompanied it.

How could so many authors be so thoroughly out of touch? The answer lies in the dramatic changes that have taken place over the past quarter-century in the way



Julian Opie, Incident in the Library I, *1983*

textbooks are published. The repetition in text after text of discredited data is part of a larger trend—a trend toward greater emphasis on packaging and less concern with content. Today's textbooks are thicker, slicker, more elaborate, and more expensive than they used to be. They are also more alike. Indeed, many are virtual clones, both stylistic and substantive, of a market leader. These trends are not unique to genetics texts: in fact, cribbing—authors' borrowing liberally from other textbooks—is widespread. And as the bizarre durability of Cyril Burt's data makes clear, the practice can have pernicious, if unintended, consequences.

As recently as the 1960s, textbooks tended to be idiosyncratic, reflecting the author's own approach in both style and substance. Three introductory genetics texts that led the field at the end of the decade—Adrian Srb, Ray Owen, and Robert Edgar's *General Genetics,* Monroe Strickberger's *Genetics,* and Eldon Gardner's *Principles of Genetics*—varied considerably in organization, emphasis, and tone. The same could be said of texts in other fields, including such classics as Robert Winthrop White's *Abnormal Personality,* P. A. M. Dirac's *Principles of Quantum Mechanics,* Eugene P. Odum's *Fundamentals of Ecology,* and Linus Pauling's *Nature of the Chemical Bond.* Their singularity was not surprising, since authors wrote texts mainly to impress their stamp on a field. "When I first came into the [textbook industry]," David P. Amerman, a marketing director at Prentice-Hall, recalled in a 1977 interview in *The Chronicle of Higher Education,* "the way you published a book was to find an academic with a reputation and hope he could write." If he couldn't, editors were inclined to preserve the author's voice, even at the expense of readability.

The trend toward homogenization began with the enrollment surge of the sixties. During that decade, the

number of undergraduates in U.S. colleges more than doubled. The most rapid rise occurred in state schools, particularly in two-year community colleges, in which nationwide enrollment rose from fewer than half a million in 1960 to more than two million in 1970. The expansion opened a whole new market, which textbook publishers moved aggressively to exploit. Two-year schools became a mainstay of the industry, and remain so today, enrolling more than forty percent of all undergraduates.

But community colleges demanded a new sort of textbook. In many ways, these institutions were more like high schools than like traditional four-year colleges. Faculty members were not expected to do research and so were given heavy teaching loads: four, five, even six courses a semester, sometimes covering every subfield of a discipline. Since instructors were not well equipped to handle such a wide range of subjects (few had Ph.D.'s, and many were part-time), they looked for texts that came with teaching manuals and ready-made tests. Indeed, some community college instructors were former high school teachers who had come to expect such satellite materials.

Instructors also demanded simpler texts, because their students had poorer reading skills, on average, than students at four-year schools. Some community colleges even required that books be written at a tenth-grade reading level, as defined by such standard tests as SMOG, Flesch, or the Frye Graph (which measure number of syllables, length of sentences, and familiarity of words). Since most publishers yearned to capture as wide a market as possible, they adjusted the reading levels of their texts—and the nature of supplemental materials—to community college standards.

Meanwhile, the changing demands and increasing volume of the college textbook market attracted a new kind of publisher—one with a heightened concern for the bottom line. A number of conglomerates entered the market, including ITT (which acquired G. K. Hall & Company and Bobbs-Merrill), IBM (which bought Science Research Associates), CBS (which acquired Holt, Rinehart & Winston and others), RCA (onetime owner of Random House), Raytheon Company (which purchased D. C. Heath & Company), Bell & Howell (which bought Charles E. Merrill), and Xerox Corporation (which acquired Ginn & Company, then sold it to Gulf+Western, owner of Prentice-Hall and Allyn & Bacon).

The new players were prepared to invest huge sums in texts, and this had the effect of reducing competition by raising the costs of production and driving smaller presses into specialized niches or out of the market altogether. Publishers dressed up their books with photographs and full-color figures; packaged them with such accessories as instructors' manuals, slides (with accompanying lecture notes), and tutorial programs on floppy disks; and even offered subsidies for the purchase of educational films. Large banks of test questions, sold with the texts, were offered in a variety of formats: on floppy disks, formatted for the instructor's personal computer; on magnetic tapes, for use on the campus mainframe; or as separately bound booklets. With these test banks, instructors could generate tests on specific chapters or topics or to fit particular course objectives, which some publishers offered to print. The preface to *Psychology: An Introduction* (a current best-seller) assures teachers that "test preparation and typing can be obtained within 24 hours through the Prentice-Hall phone-in testing service."

Such "bells and whistles," as one textbook editor describes them, are expensive. Professionals in test construction, for instance, charge three to five dollars for each of the one thousand to two thousand questions in a typical test bank. Technical illustrators may charge as much as four hundred dollars for a single drawing. And quarter-page photographs, of which there are often hundreds in a basic text, cost as much as two hundred and fifty dollars each, just for permission to reproduce. As a result, publishers came to spend increasing amounts of time and money on packaging. Today, the prevailing belief is that a basic science, social science, or business text that does not include the standard satellite materials will fail—regardless of its other virtues—since many instructors look first at the supplements and only later at the text itself.

As the market grew and textbooks changed, some publishers started looking for a different kind of author. They became less interested in a writer's scientific expertise and more concerned with his ability to reach a mass audience. Hence, many publishers stopped recruiting authors from prestigious universities—where professors may not have taught introductory courses in years and were more prone to write for their peers than for students—and began to look for successful teachers of large classes at state schools. But, in the end, this development was probably less significant than changes in the authors' own motivations.

Some textbooks are, of course, still written out of authors' beliefs that they have something important to offer, and these authors have no incentive to copy other texts; indeed, it would defeat their purpose. But some editors say that a new kind of writer has emerged: one motivated more by potential profit than by the desire to leave an intellectual legacy. Writing textbooks had never conferred great prestige, but as enrollment rose during the sixties, it suddenly became a plausible route to wealth. As a result, it began to attract authors who have little emotional involvement with the text and few ideas of their own—authors who draw inspiration from editors and, especially, from other textbooks.

The incentive to borrow from other texts is heightened by the need to cover an expanding number of topics. Since the mid-sixties, biology textbooks have increased in length by about two-thirds (most are now between eight hundred and twelve hundred pages long), and the average length of psychology textbooks has grown from fewer than five hundred pages to more than seven hundred. This is partly due to the expansion of knowledge—many of the topics in current texts, such as genetic engineering or sociobiology, scarcely existed twenty years ago—but it is also partly the result of marketing considerations. Publishers trying to capture the largest possible market are loathe to omit anyone's pet topic. Professors asked to review manuscripts often agree that the text is too long but may not agree on what should be cut. Hence, the safest policy is to leave everything in, and textbooks grow without evidence that students are actually reading more pages.

Of the multitude of topics covered in contemporary

textbooks, the author is likely to have expertise in only a few. One way to master the unfamiliar topics, of course, is to read the professional literature—to comb through specialized monographs and journals. Some authors do this. But it is much easier to borrow predigested material from other textbooks. And with so many textbooks currently in print—more than a hundred in introductory psychology alone—authors who crib can feel secure that their sources will not be easily identified.

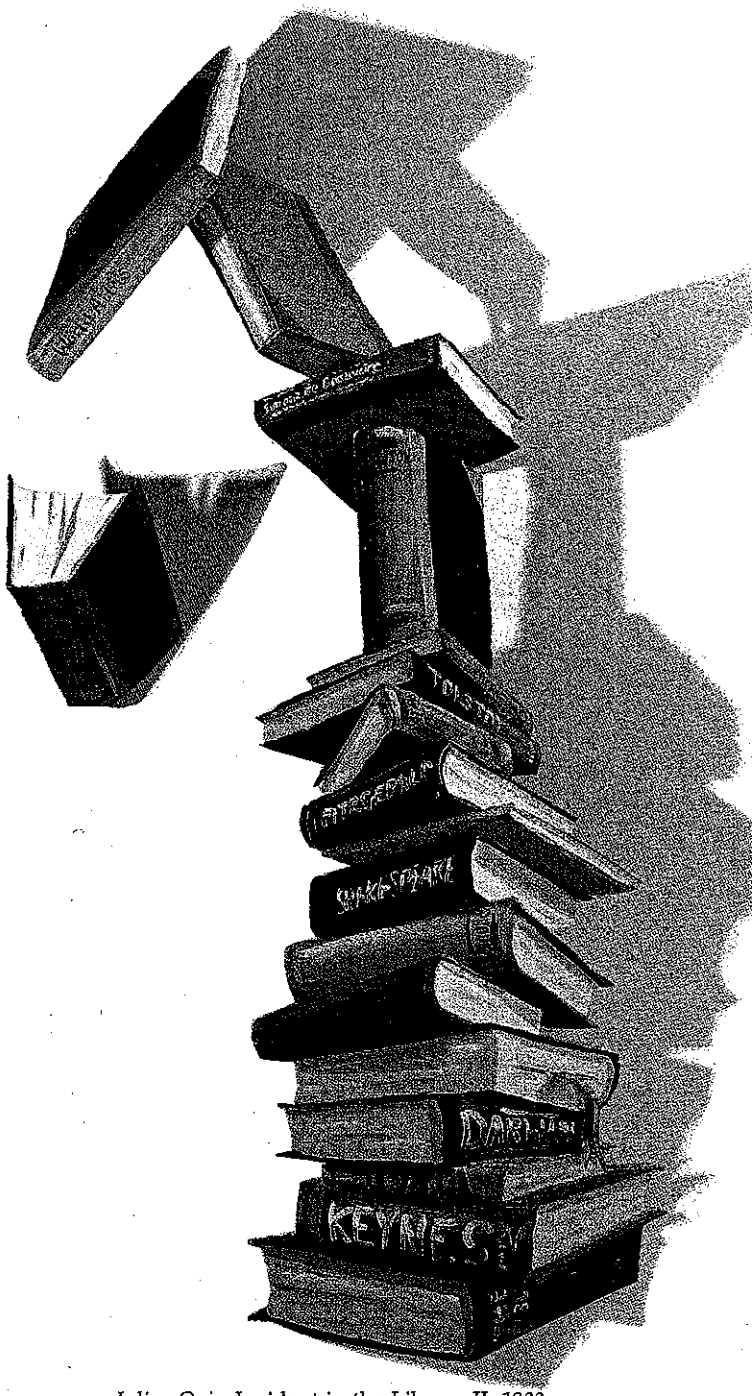
Even when authors want to be original, publishers may pressure them to conform. At times this pressure is quite overt; the publisher explicitly sets out to mimic the style and content of the most successful text in the field. In a 1974 lawsuit, Harper & Row charged the Meredith Corporation with plagiarizing its developmental-psychology text *Child Development and Personality,* which at the time was enjoying approximately a thirty-percent market share. The suit unearthed internal memorandums indicating that Meredith had hired free-lance writers, many having no background in psychology, and had provided them with detailed chapter outlines of the Harper & Row text, on which they were to base their drafts. (These draft chapters were to be edited by a well-known psychologist, the official "author" of the text.) One memorandum even warned writers to "resist the temptation to impose your own view of the subject matter; the model [the Harper & Row text] and the marketing report are the arbiters combined with your own common sense."

Such extensive copying of a single text is unusual. What is not unusual, however, is a fear of deviating from the mainstream—from textbook formulas that have already proved successful. It is not uncommon today for a press to invest as much as half a million dollars in a single text. To protect that investment, the publisher relies heavily on the results of market research and manuscript reviews to ensure that the product is salable. These results, as it turns out, almost inevitably prod the publisher to produce a textbook that resembles all others in the field.

VIRTUALLY ALL PUBLISHERS use the same forms of market research—principally questionnaires that ask potential adopters of a text how much emphasis various topics should receive, in what sequence they should appear, and how the book in question compares with others. The research typically indicates that most college teachers will resist any change in a textbook that necessitates revising their lecture notes. In a 1971 lawsuit involving the alleged plagiarism of Campbell R. McConnell's *Economics: Principles, Problems and Policies,* the judge summarized one of the defendant's successful arguments as follows:

> Economics professors, who shape the market, desire texts to which their own class notes can be adapted. Their notes, in turn, are the products of long familiarity with what might be described as "Samuelson methodology." [Paul A. Samuelson wrote the first edition of his classic textbook, *Economics,* in the 1940s; it is now in its twelfth edition.] These professors are presumptively unwilling to effect a reorganization of their own notes merely to satisfy the whim of a new textbook writer.

If the guidelines an author receives for writing his book are based largely on market research, the editing of the



Julian Opie, Incident in the Library II, *1983*

text depends largely on manuscript reviews. Two types of reviewers are used: experts, chosen for their ability to judge the accuracy of the text, and "market knowledgeable" reviewers, selected not for their expertise but for their preferences as consumers.

The larger and more competitive the market for a text, the greater the dependence on market-knowledgeable reviewers; for an introductory text in a field such as psychology, market-knowledgeable reviewers often outnumber experts by two to one. Such reviewers naturally reflect the market's conservatism, and when they dislike what is original in a new book's organization or approach, the editor often responds by encouraging the author to "study" other texts. Thus, the whole process of textbook development conspires to wash out any substantive inno-

vation, even in books that were originally attractive because they appeared to offer something new. Of course, publishers must do *something* to distinguish their texts from the dozens of others on the market, so while meaningful innovations are eliminated, novelty is introduced in the externals—the color illustrations, teaching manuals, lecture slides, and test banks. Originality is thus restricted to areas in which it is trivial, and it becomes little more than a strategy for marketing the same old book under a new author's name.

The pressures that have produced so many meaningless variations on standard textbooks are, if anything, increasing with hard times in the industry. College enrollment stabilized around 1981 and is expected to decline by about ten percent by the end of the decade, and a growing used-book industry has added to the strains on publishing houses. The number of hardbound texts sold declined by three and a half percent in 1985 and by another three percent last year. The conglomerates that bought out so many textbook publishers twenty years ago, with visions of virtually risk-free profit, have now begun to sell them. Textbook publishing, in short, has become an intensely competitive business.

This competition might have inspired greater innovation in the writing of texts. Instead, it has created a situation in which textbooks are being produced and sold like toothpaste. In *The Book Publishing Annual* of 1984, industry analyst Thomas W. Gornick summed up the new ethic with his prediction that future textbooks will have "more elaborate designs and greater use of color. . . . The ancillary packages will become more comprehensive, resembling the elementary–high school materials, and more costly. . . . New, more aggressive marketing plans will be needed just to maintain a company's position. The quality of marketing will make the difference."

ONE COULD ARGUE that these developments are really no cause for alarm. After all, not every textbook published before 1970 was a model of wit, clarity, and scholarship. Some of the old, idiosyncratic texts were genuinely inspiring to students, but others were simply exercises in self-indulgence: poorly written, lightly edited, and unintelligible to anyone but a specialist. The prose in today's homogenized primers may be bland, but in most cases it is clear. And there is no denying that the lavish use of photographs, figures, and illustrations has made textbooks more engaging. Nor is their substantive similarity a bad thing, *per se*. The purpose of an introductory text is to summarize the central facts and theories of a discipline, not to break new ground or convey novel insights. Books covering the same material are bound to be similar. So what is the problem?

If the leading texts were ideal, there would be no problem. But when the models are flawed, imitating them stifles development of better ones. And to the extent that imitation consists of cribbing information or insights, it guarantees that textbooks will become less reliable as a field advances. An author working from the professional literature is not likely to fill a text with dated ideas and discredited data. But an author drawing from existing textbooks, even good ones, has no way of knowing whether he is describing the current state of a discipline. Rather than discard worthless remnants from the past, such as Cyril Burt's studies of IQ heritability, he gives them a new air of authority.

It is doubtful that the authors still publishing such data are trying to mislead their readers; more likely, they are simply playing by the industry's new rules—modeling their textbooks on others and ignoring the literature they *claim* to be summarizing. In the genetics texts I surveyed, this is often obvious from the manner in which the literature is treated. The most frequently cited evidence for the influence of genes on intellectual performance (many of the texts cite no sources at all, even when reproducing charts and graphs from other works) is the figure that accompanied the 1963 review article by Erlenmeyer-Kimling and Jarvik. Anyone who took a close look at that article, or read the numerous critiques of it, in *Science* and elsewhere, would see that (among other shortcomings) it incorporated Burt's phony data. But authors relying on other textbooks are not privy to such insights. Some, because they are familiar with the Cyril Burt scandal, end up actually denouncing the same data they are reporting. Jenkins's *Human Genetics*, for example—after saying the figure accompanying the 1963 article "points out clearly the strength of the genetic component of IQ"—notes that Burt's data were "manipulated," resulting in their "exclusion from current reviews." Similarly, Robert H. Tamarin's *Principles of Genetics* (1982)—after citing the 1963 figure as evidence that "the measured heritability of IQ is relatively high"—spends two pages detailing "the case against Sir Cyril Burt." Still another text, Anna Pai and Helen Marcus-Roberts's *Genetics: Its Concepts and Implications* (1981), documents the link between intellectual and genetic variation with a diagram showing test scores of one hundred and twenty-two twin pairs, fifty-three of which are necessarily Burt's. Yet it, too, recounts the Cyril Burt scandal, and even refers readers to studies by Kamin and others who helped discredit Burt's data. Students exploring these suggested readings would be hopelessly baffled by the contradictions between the text and its ostensible sources. But no one expects students to be so enterprising. References are essentially decorative; indeed, one editor at a major publishing house calls them "window dressing."

It would be a mistake to presume that genetics texts alone are propagating this sort of nonsense; the practices that generate it are evident throughout the industry. But reliable textbooks are especially important, and shoddy ones particularly invidious, in the sciences. For whereas humanities professors often assemble reading lists from current paperbacks, a textbook is still the typical gateway to biology or chemistry or physics. As the sciences explode into subfields—making it less likely that any given professor will be expert in all the subjects he must teach—reliable textbooks become all the more important. In short, circumstances are forcing us to place ever greater faith in science texts, and fewer and fewer seem to warrant it. ●

DIANE B. PAUL is an associate professor of political science at the University of Massachusetts, in Boston, and a visiting scholar at the Museum of Comparative Zoology at Harvard University. Her study "Textbook Treatments of the Genetics of Intelligence" appeared in the September 1985 issue of THE QUARTERLY REVIEW OF BIOLOGY.