ETHICS

The Purloined Letters

NIH scientists thought they could prove whether a biographer accused of plagiarism was guilty or not. They used a computer and raised more questions than answers.

By PAUL GRAY

NEARLY EVERYONE WHO WRITES for publication has had the Nightmare. Not the one about appearing naked in a public place, or the other one about being forced to take a final exam without having attended any of the classes. These are nasty but tolerable, given the alternative. An author's worst dream is to be accused of plagiarism, of stealing ideas and language from someone else and parading them as original. This charge is a lightning bolt to the bole of a writer's reason for being—the task of adding to, as opposed to filching from, the sum total of human wisdom, knowledge or expressiveness. It has the additional disadvantage of being monstrously hard to refute, even when it is false.

But when is that, exactly?

Nothing better illustrates the difficulties behind this question than a bizarre case that has been simmering in academia and that last week bubbled up in some unexpected places in Washington. The plot involves a prominent historical biographer, two U.S. government scientists who specialize in tracking down fraudulent research and a computer program that they developed, known as the plagiarism machine.

In the beginning (a borrowing, let it be stated in all candor, from the first sentence of the *Book of Genesis*), Stephen B. Oates, an author and a history professor at the University of Massachusetts, Amherst, wrote *With Malice Toward None: The Life of Abraham Lincoln*. Published in 1977, this biography received a number of positive notices. The New York *Times Book Review* predicted that Oates' book "is very probably going to replace Thomas' book as the standard one-volume life of Lincoln," which, abetted by a paperback reprint the following year, is essentially what happened. The "Thomas" the *Times* reviewer cited was Benjamin P. Thomas, author of *Abraham Lincoln: A Biography* (1952).

After the popular success of his Lincoln book, Oates went on to write a great deal more, including biographies of Martin Luther King Jr. (1982) and William Faulkner (1987). His travail began in 1990, when an American literature professor named Robert Bray delivered a paper at an Illinois historical conference that pointed out some close similarities between passages in Oates' and Thomas' Lincoln biographies. Some other scholars jumped at this scent and began combing through Oates' writings, looking for evidence of unacknowledged borrowing from other sources. A year later, complaints of plagiarism against Oates were brought to the American Historical Association.

Oates, who is not a member of the AHA, vigorously and angrily denied all the accusations. He argued that any resemblances between his book and Thomas' were due simply to a reliance on the same historical documents or to an inevitable and entirely innocent overlap between separate descriptions of the same scene or event. He received some impressive support when 23 Lincoln or Civil War historians, including C. Vann Woodward of Yale and James M. McPherson of Princeton, signed a public statement claiming the plagiarism accusations against Oates "are totally unfounded."

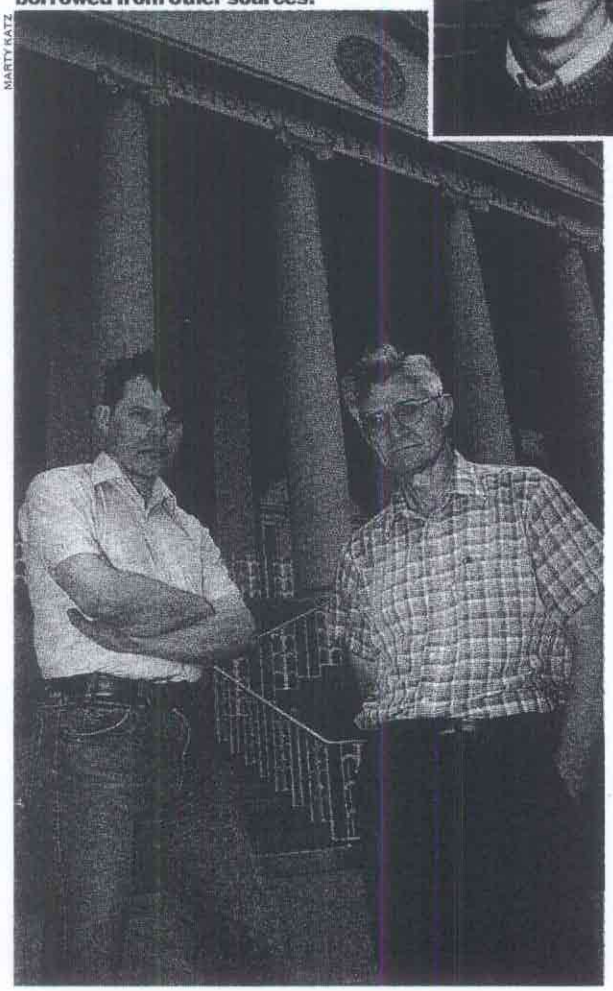
After investigating the complaints for about a year, the AHA last May quietly rendered its verdict in a letter to the principals: ". . . the American Historical Association finds that Stephen Oates' account of Lincoln's early years in *With Malice Toward None* is derivative to a degree requiring greater acknowledgment of Benjamin Thomas' earlier biography of Lincoln. The Association recognizes Mr. Oates' original contribution and style but concludes that he failed to give Mr. Thomas sufficient attribution for the material he used."

This judgment did not include the dreaded *P* word, a detail that puzzled some of the historians who read it; the failure to give sufficient attribution, after all, is a pretty good working definition of plagiarism. Still, both Oates and his adversaries could—and did—claim vindication, and there all the sound and fury (Shakespeare, *Macbeth*) should have ended.

It did not. One of the original accusers, professor Michael Burlingame of Connecticut College, remained convinced that Oates had got off too lightly. He took this belief to research physicist Walter Stewart and cell biologist Ned Feder, colleagues at the National Institutes of Health in Bethesda, Maryland, who had become specialists in the investigation of sci-

Twice-Told Tales?

Stewart, left, and Feder copy-catalogued the copycat writing of historian Oates (inset). His biography of Lincoln, they charge, has 175 passages borrowed from other sources.



MARTY KATZ

entific fraud and misconduct. They had also developed a computer program that could, they believed, make the detection of plagiarism easier and more precise. Both welcomed the chance to try out their machine on the evidence the AHA had considered in making its judgment on Oates. Says Stewart: "We saw a unique opportunity in the sense that here was a smallish body of material that could be handled quickly, where scholars familiar with the field had gone over the thing extensively."

The two scientists plunged in, first turning the pages of Oates' books and his suspected sources into computer-readable texts. They then fed the documents into the machine, and the program compared all 30-character strings in Oates' work with all 30-character strings in the others. This length was chosen, Feder explains, because shorter strings would turn up too many meaningless matches (e.g., the United States of America). If it were keyed to match longer strings—say, 60 characters—the program would then ignore shorter stretches of duplications.

Warming to their task, Stewart and Feder eventually reached well beyond the evidence assembled for the AHA investigation and ran some 60 books in their entirety through the program. What they decided they had discovered astonished them: 175 instances of plagiarism in the Lincoln biography by Oates, taken from Thomas and other sources, and 340 more in Oates' biographies of King and Faulkner.

To go public with this information, and to lodge another formal plagiarism complaint against Oates, Stewart and Feder put together a 1,400-page document, which they sent by messenger in late February to AHA headquarters in Washington, and an abbreviated version that they sent by Federal Express to a number of historians, including those who had earlier signed the statement defending Oates.

The scientists had matched up a sizable number of brief, identical passages, but had they identified the more comprehensive appropriation of another writer's work that is plagiarism? "The computer is not very smart," Feder says. "There's no simple formula, certainly none the computer could use, for making that decision. This is judgment." Some of the examples of alleged plagiarism cited in their complaint to the AHA are clearly laughable. The plagiarism machine nails Oates for stealing the phrase "balcony of the Tremont House" from Thomas; a little human judgment might have intervened here, since the ways of denoting the balcony of the Tremont House (Thomas, *Abraham Lincoln*) seem fairly limited, at least in English. Other examples (see the box) will strike different observers as accidents or cause for suspicion.

Another question raised by the Stewart-Feder report is one the compilers might have asked themselves before sending it off: Had not technological zeal got the better of prudence and common sense? Why were two scientists, paid by the government to look out for misconduct mostly in federally funded research projects, using their office hours and expensive computer time to investigate a historian who was receiving no federal subsidies? Oates asked this question in his public response to the charges made by Stewart and Feder. He says that last month he and his wife filed a number of complaints against the scientists to the Department of Health and Human Services, which oversees the NIH, as well as to a number of Congressmen, charging the pair with, among other things, misconduct.

Is This Plagiarism?

Stephen B. Oates *With Malice Toward None (1977)*	**Benjamin P. Thomas** *Abraham Lincoln (1952)*
With them came Dennis Hanks, illegitimate son of another of Nancy's aunts, a congenial, semi-literate youth of nineteen.	With them came Dennis Hanks, an illegitimate son of another of Nancy's aunts, a cheerful and energetic waif of nineteen,...
Thanks to a combination of inefficiency and fraud, the War Department had purchased huge quantities of rotten blankets, tainted pork...	The government had bought vast quantities of rotten blankets, tainted pork, disintegrating shoes.
Around midnight Major Eckert of the telegraph office brought supper into the room, and Lincoln awkwardly dished out fried oysters to everyone.	Toward midnight Major Thomas T. Eckert provided supper. Lincoln awkwardly dished out the oysters.
A nervous, windy fellow, Herndon stepped about in fancy clothes, a big silk hat, kid gloves, and patent leather shoes.	Herndon was something of a dandy in his younger years, affecting a tall silk hat, kid gloves, and patent-leather shoes.

Something must have worked. Two weeks ago, Feder and Stewart got letters from their supervisor informing them that their project was being abolished and that they would be reassigned elsewhere effective May 1. Oates takes the credit: "All this publicity and criticism from Capitol Hill and the inquiries resulted in Stewart and Feder being shut down, their plagiarism machine unplugged." For their part, the scientists remain convinced that Oates is a plagiarist and are appealing the NIH decision to halt their work.

In the end, this imbroglio solved nothing. Oates' reputation has been damaged, again, at least in the minds of those disposed to think the worst. He is currently writing a biographical study of Clara Barton's Civil War nursing career: "I'm just trying to be very careful. I always thought I was, but especially right now." He continues to deplore what he believes is the mechanistic description of plagiarism adopted by Stewart and Feder: "No writing could ever be done by their definition. If you can't say Lincoln was born in 1809 because the first biographer said that, we're getting down to some ridiculous stuff."

The dispute also opened—and left open—the question of just how helpful computers and their programs can be in providing evidence of something as shady and nebulous as plagiarism. Thanks to their computational speed and power, computers can riffle through reams of data and pinpoint patterns of repetition the naked eye might never notice. But what do these patterns signify? Intentional theft? Random clusters of words attracted to each other by grammar or syntax? Something in between? Interestingly enough, some historians who received the Stewart-Feder report decided it exonerated Oates of any suspicion of plagiarism, since the examples showed how much the majority of his writing differed from his presumed sources.

On the brighter side, the spectacle of so many intelligent people becoming exercised over the possibility of plagiarism serves as a reminder that the subject is of more than academic concern. The theft of ideas or expressions degrades the currency of information exchanges. True wit, after all, is Nature to advantage dressed; what oft was thought but ne'er so well expressed (Pope, *Essay on Criticism*). —**Reported by Elizabeth Rudulph/New York**