

Sally Hemings, Thomas Jefferson, and the Authority of Science

Whether or not Hemings and Jefferson had children together, misreported DNA and misused statistics have skewed the paternity debate, discrediting science itself.

Steven T. Corneliussen

Before science transformed the Sally Hemings–Thomas Jefferson paternity debate, a cultural and historiographical discussion had begun about the debate itself. In that metadiscussion, critics charged that disrespect for historical evidence, founded mainly in disrespect for owned humans and their descendants, had hobbled the debate for decades. In 1997, Annette Gordon-Reed, a professor at New York Law School and a lifelong Jefferson student and scholar, engaged the disrespect in a book with a metadiscussion title, *Thomas Jefferson and Sally Hemings: An American Controversy*.

Then science entered the debate. Scientists in 1998 proved a male-line DNA marker to be held in common between a descendant of Hemings’s last child and descendants of Field Jefferson, Thomas’s paternal uncle. In 2000, a statistical study professed to prove causation in the apparent association between Jefferson’s sporadic stays at Monticello and estimated dates of Hemings’s conceptions.

In 2003, R. B. Bernstein’s *Thomas Jefferson* appeared. Gordon S. Wood, then at work on the 1789–1815 volume in *The Oxford History of the United States*, called it the best short Jefferson biography ever written. Bernstein echoed other scholars’ belief that solid proof of Hemings-Jefferson parenthood now exists, resting on one nonscientific and two scientific evidentiary “pillars”: historical, DNA, and statistical. But because the DNA analysis was grossly misreported—especially at the outset, spreading misunderstanding—and because statistical science was outright misused, the metadiscussion should also engage another kind of intellectual disrespect: abuse of the special authority of science.

In an earnestly science-respecting world, this science abuse matters intrinsically—whether or not the parenthood happened, and regardless of which evidentiary pillars may support a proof. Before science entered the paternity debate, Gordon-Reed’s preface declared: “The ultimate truth or falsity of the Jefferson-Hemings story would not change my view of the way some scholars and commentators have mishandled their consideration of it and mistreated black people in the process.”

Amen. Some in the paternity debate did indeed disrespect historical evidence and steal dignity from owned humans and their descendants. Those offenses might well outweigh the science abuse. But that abuse matters too.

Science stakes a general claim of unique, lofty authority. We nonscientists usually at least respect the claim, and often we completely accept it, even if some academics level the charge of scientism against excessive faith in it.

Consider a specific example of science’s general claim, unrelated to the Hemings-Jefferson

debate. *Nature*, the world's preeminent science publication, has praised the climatologists whose blog RealClimate.org seeks to improve civic discussion of climate change by correcting errors of scientific fact found in news and commentary. Concerning contrarian dissent from researchers' consensus that humans affect climate, these scientists have declared that although "giving equal coverage to two opposing sides may seem appropriate in political discourse, it is manifestly inappropriate in discussions of science, where objective truths exist."

Science, where objective truths exist. Obviously, no historian, journalist, writer, or humanities scholar would claim authority in this way for her own field. It should also be obvious, then, that in any important matter, abuse of science's commonly respected and regularly accepted special authority requires attention.

In fact, because countless public issues hinge on trust in science's authority, because the trust is fragile and the authority far from perfect, because science and the humanities often intersect but still have lots to learn from each other, and because scientists themselves have led the Hemings-Jefferson science abuse, it is no disrespect to Gordon-Reed or her causes to borrow her phrasing and recast it to say: The ultimate truth or falsity of the Hemings-Jefferson story would not change my view of the way some scientists, editors, scholars, commentators, and journalists have mishandled their consideration of it and mistreated science and history in the process.

Here are my two main Hemings-Jefferson science-abuse charges:

- In the DNA study, a valid scientific protocol led to valid but limited scientific findings that *Nature*'s editors and others misreported—thereby fostering misunderstanding among journalists, the public, and scientists, though less so among Jefferson scholars, who usually avoided misunderstanding by carefully considering the study itself.
- In the statistical study, an invalid scientific protocol led to invalid findings that credulous Jefferson scholars incorporated as a pillar of parenthood proof—thereby fostering misunderstanding among other historians, though less so among journalists and the public, who were often spared misunderstanding by being spared news about the statistical study, and not at all among scientists, who also seem unaware of it.

The second charge may at first appear the more serious, since it denies the very existence of any statistical pillar in the multipillared paternity argument. Nevertheless the misreporting about the perfectly valid DNA pillar matters too, because the paternity debate itself matters not just among those historians who avoided DNA misunderstanding. The debate also matters among journalists and others who did not avoid that misunderstanding, and whose faith in science likely pointed them toward firmer, or different, historical judgments than they might otherwise have reached.

Misunderstanding of the Hemings-Jefferson DNA analysis must surely have involved popular understanding of crime-scene DNA. By the time a scientific report on the analysis appeared in *Nature* in late 1998, O. J. Simpson's murder trial and President Clinton's impeachment had taken place. DNA had become a widely familiar and especially trusted kind of evidence, representing a

high level of science's authority. For many, DNA laboratory findings must have seemed absolutely able to identify not only criminals, but presidents who take advantage of young women.

Purely scientifically, however—that is, before being interpreted in the light of other evidence, as in the multipillared argument—the inherently limited DNA analysis in the Hemings-Jefferson matter could prove far less. Disrespect for this scientific limitation led to the DNA misreporting that abused the authority of science.

“DNA analysis of the Y chromosome,” the reporting scientists wrote in *Nature*, “can reveal whether or not individuals are likely to be male-line relatives.” And what those scientists called their “molecular findings” did reveal some important facts about male-line relationships, and the lack of them, among fourteen descendants of four men from past centuries. The four male ancestors included Field Jefferson, but not Thomas Jefferson, because the only son from Thomas Jefferson's only marriage didn't survive infancy.

Most importantly, the molecular findings established an interfamily DNA match. They proved that someone carrying a Jefferson family male-line DNA marker probably fathered Eston Hemings, the last of Sally Hemings's several children. The inherently limited DNA analysis could not, however, identify Eston's father from among about two dozen or more DNA-marker-carrying possibilities.

The two dozen or so known paternity candidates belonged to Jefferson's acknowledged extended family. But if the DNA marker had spread beyond the acknowledged family over the decades in Virginia's plantation culture, the circle of candidates could be larger. No one knows.

Nevertheless some paternity-debate participants believe that historical evidence cinches the sifting of the paternity candidates. They believe that only Thomas Jefferson's paternity of Eston can explain the molecular findings. Maybe they're right. But however simple other evidence may make, or seem to make, the sifting of paternity candidates, the DNA analysis itself can contribute to the sifting nothing whatsoever.

In *Nature*, the DNA scientists made plain this scientific limitation, even though they too found the sifting simple. In effect, they wrote both as DNA scientists and as historical observers, though they should have been clearer about their dual perspective. As scientists, they reported that someone carrying a Jefferson male-line DNA marker must have fathered Eston. As historical observers, they offered the opinion that “in the absence of historical evidence to the contrary,” the sifting of paternity candidates can lead only to Thomas Jefferson.

By definition, the reporting scientists' molecular findings can command the authority of science, but their historical opinion can't. So *Nature*'s editors abused science's authority when, without the DNA scientists' consent, they headlined the DNA report with an unambiguous declaration: “Jefferson fathered slave's last child.”

Now, maybe a combination of historical and other evidence proves conclusively, as many believe, that Eston Hemings's father was not just any Jefferson, as the DNA actually proved, but

Thomas Jefferson, as *Nature*'s headline proclaimed. In any case, the DNA analysis transformed the historical debate among serious participants, whether or not millions of onlookers worldwide mistook *Nature*'s headline for impeccable scientific authority. So how much did that headline abuse the authority of science?

I asked Catherine Westfall, a historian of science. She answered with an analogy involving a particle-physics breakthrough at Fermi National Accelerator Laboratory near Chicago. "I remember talking with then-Fermilab director John Peoples about the top-quark data," she recalled. "He said he was convinced they 'had the events,' but he insisted that neither of the two competing Fermilab groups report the 'likely' events before they had the flat, absolute statistical proof."

Though Westfall supposes that paternity believers have a strong case overall, she then observed that "[n]obody in science would dream of reporting data like 'shows that a Hemings descendant and Jefferson descendants have a common male relative' is the same as 'Jefferson was the common male relative.' ... It simply would not be allowed."

—

Nature's Hemings-Jefferson science abuse didn't end with the famously errant headline. The journal's editors have written that their "role is to publish the most innovative and influential papers that scientists can produce, and to present these results to the public." The issue of *Nature* containing the inaccurately headlined DNA report also contained, for the public, two more misinforming scientific declarations about the DNA. Both assertions were prominently displayed and well suited for fostering misunderstanding about what science had and had not proven in the paternity debate.

The DNA report itself appeared in *Nature*'s "Scientific Correspondence" section with three other scientific papers: "Coherent light scattering by blue feather barbs," "A lower jaw from a Cretaceous parrot," and "Hydrologic cycle explains the evaporation paradox." The extra pair of misinforming declarations appeared, however, in a commentary in a less densely technical section called "News and Views." There *Nature* serves nonspecialists and the public with articles discussing and explaining scientific reports.

Within that commentary's text, the scientist Eric S. Lander and the Jefferson biographer Joseph J. Ellis argued, in a carefully qualified way, that the proven interfamily DNA match "offers strong evidence that Jefferson fathered at least one" Hemings child, and that together with the historical evidence, the under-one-percent probability that mere chance explains the DNA match "seems to seal the case." Within their text, Lander and Ellis respected the distinction between what could and what could not wield science's special authority.

Above and alongside their text, however, no caution about conflating scientific and other evidence restrained either the article's prefacing thumbnail summary or the caption given for a Jefferson likeness. I don't know who inserted these misinforming declarations, but *Nature* in any case allowed their disregard for the commentary text's caution and their contradiction of the DNA scientists' report. The summary, in boldface, invoked the authority of science in

announcing unambiguously, and wrongly, that “DNA analysis confirms that Jefferson was indeed the father of at least one of Hemings’ children.” The caption invoked the authority of science by announcing unambiguously, and wrongly, that “DNA analysis ... shows” that Jefferson fathered at least one Hemings child.

In fact, however, DNA analysis could confirm or show nothing of the sort. What the scientists actually wrote about their DNA analysis bears repeating: “DNA analysis of the Y chromosome can reveal whether or not individuals are likely to be male-line relatives.” That’s the extent of DNA science’s evidentiary power in the Hemings-Jefferson debate. It’s important, but it’s not dispositive.

A week later, nevertheless, my local newspaper reprinted a *Los Angeles Times* article asserting unambiguously, and wrongly, that genetic tests “showed that Eston was sired by Jefferson” and that “DNA tests proved” the paternity. Similar misreporting abounded. Belief had spread through an earnestly science-respecting world that the Sally Hemings–Thomas Jefferson parenthood thesis had been proven directly and solely by DNA evidence, by Lofty Science Itself.

—

That kind of misinformation about Hemings-Jefferson DNA requires ignoring a crucial distinction: The laboratory results that the reporting scientists called their “molecular findings” truly represent scientific authority, but historical interpretations of those findings don’t. More often than being outright ignored, though, this crucial distinction is simply blurred.

Statements typifying the blurring appeared from Gore Vidal in late 2000, two years after the DNA report appeared in *Nature*. His statements provoked a brief letters-to-the-editor exchange in the *Times Literary Supplement* between Vidal and David Murray, a paternity skeptic then directing research at a Washington media-monitoring organization called the Statistical Assessment Service. In a *Washington Post* commentary two weeks after *Nature* published the DNA report, Murray had criticized those two weeks’ flood of misreporting. Now Murray protested Vidal’s unclarified scientific claim that the “conclusion of eight investigators (*Nature*, November 5, 1998) was that Jefferson was indeed father to Sally’s last son.”

As de facto historical observers, the DNA scientists, led by principal investigator Eugene A. Foster, did indeed espouse that historical interpretation of their molecular findings. But Vidal’s word *investigators*, coupled with his specifying an issue of the world’s preeminent science publication, plainly meant science, not just scientifically influenced historical interpretation. In reaching for the authority of science to prove rather than merely buttress a historical interpretation, Vidal blurred the crucial distinction between molecular findings and historical interpretation, as he seems later to have learned. In 2003 he wrote—with a simple accuracy that by then seemed a bit more common in public discussion—that DNA testing “establish[ed] consanguinity” between the Hemingses and Jefferson.

After Murray’s early protest of the flood of DNA misreporting, the *Washington Post*’s ombudsman, E. R. Shipp, listed and criticized contributions to that flood from the *Post*’s reporters, editors, and columnists. Though she was later to express her personal view that only

“contorted thinking” could deny Hemings-Jefferson parenthood, she disapproved of misreporting. Her May 30, 1999, ombudsman column concluded by citing the example of an especially ardent paternity disbeliever, Herbert Barger, who admitted that absolute certainty will likely always elude him. “Reporters and headline writers should be so honest,” Shipp wrote. Countless other media organizations could have used Shipp’s insistence on unblurred facts.

Some still could. The February 6, 2006, *Newsweek* feature “DNA Testing: In Our Blood,” for example, misreported that “genetic testing in 1998 revealed that Thomas Jefferson was most likely the father of at least one of Sally Hemings’s children.” Revealed? That’s the verb from the reporting scientists’ careful stipulation establishing what their DNA analysis could actually do: it could “reveal whether or not individuals are likely to be male-line relatives.” Given that the blurring of the crucial distinction between scientific findings and historical interpretation still crops up regularly in public discourse, it’s no surprise that *Newsweek* blurred it also in a sidebar and yet again in an on-line question-and-answer exchange.

Or consider a statement not blurred, but simply wrong, from the October 7, 2007, issue of the widely circulated Sunday magazine *Parade*. Lori Andrews, a law professor and bioethicist, had led the team that contributed to *Science* magazine the 2004 biohistory-ethics article discussed below. Among other things, that *Science* article had condemned overstatement of the Hemings-Jefferson DNA evidence. Yet in *Parade* in 2007, Professor Andrews asserted that “DNA analysis indicated that Thomas Jefferson fathered a child with his slave Sally Hemings.”

Historians too have occasionally advertised more from DNA analysis than it can deliver. In late 1998 Daniel P. Jordan was heading the Thomas Jefferson Foundation, which runs Monticello. From him the DNA news prompted a hasty statement announcing the foundation’s intention to evaluate the genetic study carefully, yet asserting—with no scientific basis—that “DNA evidence indicates a sexual relationship” between Hemings and Jefferson. Later, in *Sally Hemings and Thomas Jefferson: History, Memory, and Civic Culture*, historians’ essays built on the tenable premise that the parenthood is now proven. The book opens, however, with a genealogical diagram annotated with the blurred, misinforming scientific claim that according to the DNA analysis, “Jefferson’s paternity of Eston Hemings is almost certain.” When Alexander O. Boulton attacked paternity disbelief in the October 2001 *William and Mary Quarterly*, a leading journal for Jefferson studies, his review opened by asserting, wrongly, that “DNA analysis identified Thomas Jefferson as the most likely father” of Eston. And in BookTV’s April 2008 airing of a Jefferson discussion at the University of Virginia, Alan Pell Crawford said of the DNA, “the Carr boys—it cleared them,” though in fact the DNA excluded Jefferson’s Carr relatives from paternity of only one Hemings child, Eston. That scientific error seems common; in 2007, it had even appeared in Maura Singleton’s fine article summarizing the Hemings-Jefferson controversy in *The University of Virginia Magazine*.

—

Given that science itself has led the Hemings-Jefferson science abuse, maybe it’s no surprise that scientists have also occasionally scanted the crucial distinction between molecular findings and historical interpretation. A well-known chemist, for example, in an invited comment at Edge.org, claimed that DNA had confirmed Jefferson’s “affair” with Hemings. *The Scientist* declared that

“researchers used DNA testing to determine that Thomas Jefferson had fathered a child with Sally Hemings.” A March 28, 2007, University of Leicester scientific press release about new understanding of Jefferson’s ancestry opened by asserting that the 1998 DNA analysis “showed that Thomas Jefferson had fathered at least one of the sons of Sally Hemings.” The confusion even undermined two earnest attempts at the weekly *Science* to criticize misreporting about the DNA evidence.

To *Science*’s credit as a forum representing the American Association for the Advancement of Science, the world’s largest general scientific society, the publication did confront the DNA confusion. A January 9, 1999, *Science* news headline asserted the question that *Nature*’s DNA headline had skirted weeks earlier. It asked, “Which Jefferson Was the Father?” The article emphasized the crucial distinction between molecular findings and historical interpretation, and it quoted the principal DNA investigator’s observation that answering the headlined question is “a job for historians.” But after the opening paragraph brought up the “flood of news reports declaring that the third U.S. president, as rumored, fathered an illegitimate child” with Hemings, the article continued with the disheartening, because wrong, claim that “now”—please note that word—“the authors of the [DNA] report say the evidence for that is less than conclusive.”

Now? On that adverb *Science*, and more importantly small-*s* science itself, pivoted away from the best chance to establish clearly how science’s authority had been misused in the Hemings-Jefferson debate—so far, anyway, since this was a year before the statistical study appeared.

Nature and *Science* lead civic discussion within what was once called the republic of science but is now more likely to be called the international science enterprise. Sometimes this technocivic discussion involves science’s routine self-policing and self-correction, the system’s inherent counterbalancing of its individuals’ errors. If the Hemings-Jefferson matter is important, and if it’s true that science and the humanities will increasingly interpenetrate in what has been called consilience, it’s too bad that small-*s* science’s Hemings-Jefferson self-criticism has been stingy at *Nature*—and too bad that post-DNA confusion hobbled *Science*’s earnest attempts to establish what went wrong.

The DNA authors disclaim conclusiveness “now”? Six months before the DNA report appeared, the *Washington Post* had quoted the principal investigator about the impossibility of obtaining historical conclusiveness. Eugene Foster had explained that the DNA analysis could only “throw the probability” of Jefferson’s paternity “in one direction or another.” Accordingly, the DNA report itself opens by stating a limited scientific purpose: to “throw some scientific light on the dispute.” It’s true that in their report the DNA scientists only clumsily distinguished their reporting-scientists’ voice from their historical-observers’ voice. But even when they pressed their historical interpretation of their molecular findings, they stipulated that they “could not rule out other explanations.”

So that errant adverb shows that *Science* operated under the mistaken assumption that the DNA report, and its authors, required fundamental correcting for failing to recognize the crucial distinction between molecular findings and historical interpretation. That mistaken assumption meant that within science’s system of natural self-correction, *Science* missed holding *Nature*’s editors fully accountable for the misinforming headline, summary, and caption.

Science's hobbled awareness continued to matter when the journal published the April 9, 2004, "Policy Forum" article "Constructing Ethical Guidelines for Biohistory"—in my view not only an admirable, imaginative widening of science's routine self-criticism, but an important illustration of the consilience of science and the humanities. Because part of the article criticized the Hemings-Jefferson DNA study, the journal itself was rejoining the paternity-debate metadiscussion. But again, with this scientific historiography, with this second chance to correct misunderstanding, *Science* fell short.

The biohistory-ethics article posed and engaged important questions: "Should technologies of bioanalysis, such as DNA testing, be used to try to answer questions about historical figures? If so, what social, legal, and scientific standards should be used?" But despite the authors' own advocacy of scrupulous standards for accuracy, errors concerning the Hemings-Jefferson DNA analysis marred their essay. In a merely technical error, the authors twice referred to the acknowledged Jefferson family members whose DNA was tested as having descended from Thomas Jefferson himself, rather than from Field Jefferson. But in an error that actually blocked understanding, they condemned the DNA study's scientific protocol.

The biohistory-ethics authors noted that researchers must never ask genetic testing questions it can't answer. And so it's a big problem, they charged, that the Hemings-Jefferson DNA study "was intended to establish whether the president had fathered Sally Hemings' children" via a scientific protocol that was "inappropriate for determining the paternity." Of course, it was precisely because no DNA protocol could determine the paternity that the DNA scientists' report opened not by stating the intention that the biohistory authors mistakenly imputed, but by reiterating the longstanding purpose merely to "throw some scientific light on the dispute." The DNA scientists plainly fulfilled that limited purpose by establishing the interfamily DNA match and transforming the two-century-old debate.

The biohistory-ethics authors also omitted any mention of two other biohistorical purposes that the DNA analysis served quite well. The analysis established that some paternity disbelievers have probably been wrong, for two centuries, to include Eston Hemings when alleging that Jefferson's Carr family relatives fathered children by Sally Hemings. It also established that some paternity believers have probably been wrong, for two centuries, to allege that Thomas Woodson, said by many to have been Eston's brother, was Jefferson's son.

So in misreporting the DNA study's purpose and in mistakenly condemning its scientific protocol, *Science*'s biohistory-ethics authors had also committed a version of the error that *Science*'s news article had committed earlier. Uncorrected by any editor, fact checker, or peer reviewer at *Science*, they had in effect mistakenly blamed the DNA scientists for the confusion that followed *Nature*'s three misinforming declarations. In the process, they too missed the chance to hold *Nature*'s editors fully accountable for abusing the special authority of science.

Only *Science* compares to *Nature* for leading technocivic discussion in the republic of science. Yet both times *Science* tried to engage the Hemings-Jefferson science abuse in *Nature*'s three misinforming declarations, *Science* erred at a level that no student, in either science or history, could sustain in a seminar discussion. In my view this deficiency contributes to a supplemental Hemings-Jefferson science-abuse charge:

- The republic of science, despite its tradition of routine self-criticism, and despite having contributed two of the parenthood argument's three pillars, has fallen short in straightening the DNA pillar's use and in condemning the statistical pillar's very construction.

Neither *Science* nor *Nature* appears even to have known about the scientific report claiming to erect the statistical pillar, which is actually biostatistical in that it involves—or at any rate should involve—the variable length of human gestation. The report appeared in the *William and Mary Quarterly*, a journal far distant from science's usual forums. If small-*s* science had taken its self-criticism obligation more seriously, scientists would long since have noted the irony that this biostatistical study merits exactly the criticism that *Science*'s biohistory-ethics authors mistakenly leveled at the DNA study. It asserts a grand scientific proof, a paternity proof, that its scientific protocol inherently can't deliver.

Fraser Neiman, a Monticello and University of Virginia archeologist, conducted the study and presented it in “Coincidence or Causal Connection? The Relationship between Thomas Jefferson's Visits to Monticello and Sally Hemings's Conceptions.” This scientific report concludes the *William and Mary Quarterly*'s January 2000 collection of post-DNA Hemings-Jefferson essays by Gordon-Reed, Ellis, and other prominent Jefferson scholars.

In the report, Dr. Neiman claimed he could settle the paternity debate by mathematically evaluating the association, the correlation, between two sets of data. One set, involving important assumptions and an unjustifiable disregard for the variable length of human gestation, was his estimates of exact dates when Hemings conceived children during a selected portion of her life, a period beginning several years after she returned from Paris with the Jefferson family. The other data set, involving an important error of omitted fact, was the recorded dates of Jefferson's sporadic presences at Monticello during the selected period.

No one disputes that some sort of association exists between Jefferson's Monticello visits and Hemings's conceptions. No one disputes that, qualitatively, the coincidences are obvious and suggestive. Neiman, however, went much further. He claimed that he could quantify the probability that chance alone could explain the association, and that in doing so, he could quantify—and all but eliminate—uncertainty about the paternity question itself. Neiman sought to transform an intuitive, qualitative judgment into a quantified scientific pronouncement.

For this quantifying Neiman chose a two-step scientific protocol. First came computer simulations. He devised four general models of Hemings's reproductive life during the period he had selected. For each model, software randomly generated 100,000 specific possible scenarios. By comparing these with the dates in his two data sets, Neiman completed his first step. He calculated a probability of only about 1.2 percent that the “visit-conception relationship occurred in the absence of a causal connection.”

For his second step, Neiman applied to that calculated probability something called Bayesian analysis. By this procedure, controversial in the republic of science, he accounted for the strength of the Hemings-Jefferson parenthood belief that he and others had held before he even began his study. This accounting, he declared, yielded “a posterior probability of 99 percent—near certainty” that Sally Hemings and Thomas Jefferson had six children together. Neiman’s scientific report concludes with a sentence invoking the authority of science itself for his pre-existing historical belief. He wrote: “Serious doubt about the existence and duration of the relationship and about Jefferson’s paternity of Hemings’s six children can no longer be reasonably sustained.”

Among historians, Neiman’s grand scientific claim seems widely accepted. It was not only R. B. Bernstein, in that short Jefferson biography that Gordon S. Wood praised as the best ever written, who incorporated the study in a multipillared parenthood proof. Jan Lewis wrote the introduction to that *William and Mary Quarterly* post-DNA essay collection where Neiman’s report appeared. She called the study an “ingenious statistical evaluation” that “should quiet those who have resisted accepting Jefferson’s paternity.” She declared that when the study is seen together with the DNA and historical evidence, it justifies belief in Neiman’s grand scientific claim.

—

When I first read Neiman’s scientific report in 2000, I was skeptical that anyone using any array of assumptions could prove the paternity thesis by statistically evaluating the visits-conceptions association. But I’m a writer and media advisor working with physicists; I’m not a biostatistician or a scientist. So I wondered whether the study would draw informed comment from disinterested experts.

It’s true that probability theory, the report’s underlying science, evolved because the measuring of uncertainty can be useful. It’s true that uncertainty-measuring has become fundamental to efforts as disparate as drug development, industrial process control, and particle physics experiments. And it’s true that uncertainty-measuring can even yield useful understanding about distinctly human complexity, as in census-taking or demography. Nevertheless when it came to measuring uncertainty in the unique, unrepeatable, far from completely recorded circumstances of the two-century-old Hemings-Jefferson paternity mystery, I wondered whether the authority of statistical science could be wielded at all. It seemed to me that inherently unquantifiable uncertainties could well block the reliable compiling of data for computation. What about the fundamental statistical assumptions? Sometimes enslaved people were sent away from Monticello for this or that purpose. We know almost nothing about Sally Hemings. Is it certain she was always there? Neiman’s approach requires her kids to have had only one father. Can anyone measure that uncertainty?

Moreover, it was a humanities journal, not a scientific one, that published Neiman’s scientific study, and in both science and the humanities—C. P. Snow’s two cultures—people remember that 1996 hoax in *Social Text*, when the physicist Alan D. Sokal published the parody article “Transgressing the Boundaries: Toward a Transformative Hermeneutics of Quantum Gravity.” That scientific-sounding paper purported to present Truth embracing both cultures. In a parallel

Lingua Franca article, Sokal explained that he had wanted to see whether “a leading North American journal of cultural studies” would “publish an article liberally salted with nonsense.”

Obviously Fraser Neiman is no hoaxer. But no scholar’s or scientist’s academic stature proves anything about the adaptability of the *William and Mary Quarterly*’s peer-review process to the task of scientifically vetting a probabilistic evaluation, as Neiman called his study. Few statistical experts, presumably, read that journal. Yet Neiman’s report claims the authority of objective statistical science in support of an interpretation of historical circumstances. So even though a Neiman footnote acknowledges two distinguished demographers, my question remained: In accepting Neiman’s report, did the humanities journal apply more scientific rigor than *Social Text* applied in falling for Sokal’s hoax?

I didn’t know. So I tried to learn whether any expert scientific commentary about the report would appear. I queried the *William and Mary Quarterly* itself, the Thomas Jefferson Foundation, and David Murray’s Statistical Assessment Service. As far as I know, no technical review ever did appear, though Murray published a highly skeptical but mainly nontechnical critique in an ardently paternity-disbelieving anthology, *The Jefferson-Hemings Myth: An American Travesty*.

Then, four years later, a University of Virginia legal scholar asked me to criticize his Hemings-Jefferson manuscript. Robert F. Turner had chaired an independent, volunteer commission of thirteen academics from twelve institutions who studied the multipillared paternity argument at the request of the Thomas Jefferson Heritage Society, the citizens’ group that presented *The Jefferson-Hemings Myth*. With one mild dissent, Turner’s “Scholars Commission” had arrived at views ranging from acute skepticism about the parenthood thesis to a deep belief that it is almost certainly false. Turner himself had arrived at that deep belief. He sought my critique in part because I haven’t.

In the manuscript Turner conveys his personal views as a supplement to his colleagues’ report, all to appear in book form from Carolina Academic Press. When I reached his chapter about Neiman’s study, it seemed to me that he was asking astute questions, but that from any point of view—paternity-believing, paternity-disbelieving, paternity-agnostic—Neiman’s report still needed some kind of disinterested, technically authoritative reviewing. After four years, though, nothing like that seemed likely from the republic of science. After all, *Nature* and *Science*, though ready to represent small-s science in the controversy, had bungled the DNA and had never even mentioned the paternity argument’s statistical pillar.

So I did something that was within my own power. I asked two experts, William C. Blackwelder and David R. Douglas, to read and comment on Neiman’s scientific report.

Bill Blackwelder is a biostatistical consultant who completed a career at the National Institutes of Health. He was named a fellow of the American Statistical Association for “statistical leadership in major international vaccine efficacy trials; for contributions to design and analysis of trials of clinical equivalence; for improvements in the sampling and analysis of sibling data for genetic linkage analysis.”

Dave Douglas, physicist and fellow of the American Physical Society, is a senior scientist at Jefferson Lab—the Thomas Jefferson National Accelerator Facility in Newport News, Virginia. Statistics and probability theory pervade his specialty, which is the theory, design, and operation of particle accelerators and free-electron lasers. For over twenty years he has regularly run computer simulations akin to Neiman's.

It was as personal friends of mine and as interested private citizens that Bill and Dave accepted my request. For their efforts they earned only the satisfaction of engaging an important issue and the chance to sacrifice time. Neither had extensive general Hemings-Jefferson knowledge. Both say that media reports had made them suppose that paternity by Jefferson must be likely. Both are longtime, enthusiastic political liberals. Neither is guilty of any errors my essay might be conveying.

Bill began our informal e-mail discussion on fundamental scientific grounds. With Dave's emphatic agreement, he energetically condemned Neiman's grand scientific conclusion. According to Bill, a scientifically unsupportable leap from association to causation invalidates Neiman's claim that "[s]erious doubt about the existence and duration of the relationship and about Jefferson's paternity of Hemings's six children can no longer be reasonably sustained."

Bill framed and explained this condemnation in terms of the standard statistical practice of hypothesis testing. Neiman, he wrote, "seems to think he is testing one statistical hypothesis, whereas I believe he is actually testing a quite different one. A null hypothesis ... is one that is set up to be tested and (usually) is then either rejected or not rejected after examination of the relevant data." Neiman's conclusions, Bill continued,

seem to imply that he is testing the hypothesis that Jefferson was not the father of all six of the Hemings offspring considered in his analysis. He rejects this hypothesis and concludes that Jefferson was the father of all six. His assumptions are faulty, but neglecting that for the moment and assuming his analysis is perfect, the null hypothesis he is entitled to reject is that there is no association between the conception dates of Hemings' children and the dates of Jefferson's presence at Monticello. It is a basic tenet of studies of statistical association that association does not imply causation. Thus, Neiman is certainly not entitled to say, no matter how strong he believes the association to be, that he has proven Jefferson's paternity. He is not even entitled, based on the data presented, to say that he has shown that the probability is high that Jefferson was the father of any of the six children, much less all six. Neiman has shown, at most, that the association between the conception dates and dates of Jefferson's presence is greater than would be expected by chance alone.

It might be objected that this criticism addresses only Neiman's step 1, the computer simulations. It was only in step 2 of the scientific protocol, after all, that Neiman claimed to cinch the case with Bayesian analysis. But Bill and Dave condemned step 2 as well. Bill wrote:

After reporting the simulations, Neiman purports to apply Bayes' Theorem to derive a "posterior probability" that Jefferson was the father of all six children,

but it is not clear how he did so. The actual formula he used is not clear. In any case, Neiman commits a gross error in claiming that he is even considering the probability that Jefferson was the father of the six. As I tried to make clear in the beginning, Neiman is only investigating whether Jefferson was present at Monticello when Hemings conceived more often than would be expected by chance alone.

Concerning step 2, Bill and Dave also agree with David Murray that Neiman argues circularly, that he begs the question with a Bayesian procedure that somehow uses the paternity argument's nonstatistical threads to validate the simulation results. Concerning the study overall, Bill concluded: "To me this seems a clear case of 'guilt by association,' and the association is itself less strong than claimed. In my opinion, Neiman's conclusion ... is, on the basis of the data he presents, a gross misinterpretation."

—

Neiman's study has other problems too. Though Dave admired Neiman's ingenuity in attempting computer simulations, he puzzled over Neiman's approach to them. A related problem was the lack of what technical people call documentation. As Dave put it, "in a technical venue, you traditionally either give enough information to reproduce a computation or cite a reference giving that information—this paper is flawed in that it seems to do neither." In my view, when the *William and Mary Quarterly* made itself a venue for science, it became obligated to require science's common practices.

Nevertheless, setting aside the association-to-causation leap, other technical difficulties, and problematical assumptions, maybe some would argue that statistical science could still contribute to the paternity debate. Maybe some would argue that statistical science could yield a quantified statement of the likelihood that chance alone explains the visits-conceptions association. So even if Neiman's study can't prove its grand concluding claim, has it at least reliably contributed that quantified statement?

No. Among the study's deficiencies is one problem so fundamental that, just by itself, it cancels any possibility of the study's contributing usefully. By failing to apply an obviously necessary biostatistical technique, Neiman failed to account for the distinct statistical chance—in one case, the greater than fifty percent chance—that at the time of conception in four of the six cases Neiman designated for study, Jefferson could actually have been absent from Monticello.

The fundamental failure lay in the handling of conception dates. In five of the six cases, Neiman estimated them simply by counting back from each recorded or estimated birth date 267 days, a legitimate but misused number from the medical literature. In the sixth case, because Jefferson had arrived at Monticello after the conception date that Neiman computed, Neiman simply assumed a three-day fudge factor. According to Bill and Dave, however, all six cases required instead a probability distribution accounting for the variable length of human gestation.

What they meant is simple. Some babies do arrive precisely on schedule, but some arrive days early or late, and some arrive many days early or late. In a study that Neiman himself calls a

probabilistic evaluation, and that hinges on the dates of Jefferson's comings and goings, this uncertainty requires measuring too.

Dave and Bill—informal though expert scientific reviewers—didn't investigate every implication of Neiman's failure to use a probability distribution. Nor did they try to define a specific distribution that biostatisticians might use for the Virginia of two centuries ago—if such biohistorical knowledge is even obtainable, given the presumably incompletely known and measured living and health conditions and medical practices of Sally Hemings's time.

In what technical people call a back-of-the-envelope calculation—a simple, obvious estimate—Bill did note, though, one inevitable implication of any valid distribution. He wrote to me that it “would result in a probability less than 50% that Jefferson was present at Monticello at the time the second child was conceived, and thus an overall probability for his presence at all six conceptions of less than 50%.” Dave agreed about the back-of-the-envelope obviousness of this simple refutation of all of Neiman's computations. The two scientists were referring to Jefferson's arrival three days after the conception date that Neiman calculated for the case of Sally Hemings's son Beverly.

Although Cynthia H. Burton professes no statistical expertise, she too saw that any assessment of the visits-conceptions association—whether qualitative or quantitative—requires accounting for gestation's variable length. In 2005, she reported new Hemings-Jefferson research findings in *Jefferson Vindicated: Fallacies, Omissions, and Contradictions in the Hemings Genealogical Search*. Former Monticello director James A. Bear, Jr., wrote the foreword.

Burton consulted the *Merck Manual*, which gave her 266 days instead of 267, the also legitimate figure that Neiman misused in attempting to compute exact conception dates. The manual could not, of course, establish a standard for Sally Hemings's day and circumstances, but it also told Burton that about 90 percent of births occur within two weeks either way from on time. Burton calls this period the “four-weeks conception window.” Neiman failed to account probabilistically for variability within that window, some version of which must frame either the 267-day gestation-length standard he chose for Hemings or any other such standard. That's what Bill and Dave meant by calling for a probability distribution.

For assessing Neiman's computations' usefulness, anyone who consults both Burton's and Neiman's data can see that this failure matters fundamentally. Burton's data are needed because Neiman failed to record Jefferson's absence at the very end of the apparent four-week window for Eston Hemings's conception. Neiman's data are needed because Burton believes the unnamed third child never actually existed.

With the implications accepted for a four-week window framing a 266- or 267-day standard, the combined data show Jefferson absent for a day or so at the window's very end for both Eston and his brother Madison, absent for a week from the unnamed third child's window, and—as Bill and Dave emphasized on the back of that envelope—absent for more than half the window for Beverly. That longest absence, as Bill said, just by itself means that the overall probability for Jefferson's presence at all six conceptions is under fifty percent.

Concerning the crucial case of Eston, however, there's more. Neiman himself cites an article called "The Length of Uncomplicated Human Gestation" from the June 1990 issue of *Obstetrics & Gynecology*. That article mentions evidence that modern-era black mothers might have comparatively short pregnancies. So another look at the combined data reveals another complex question that Neiman didn't engage, and that maybe can't be answered. This time the question is about gestation lengths in the era of the partly African-American Sally Hemings. Might Jefferson's absence from the end of Eston's conception window have lasted not just a day or so, but longer, since the window itself might actually be shifted closer to the date of birth? Such a shift would increase the already appreciable chance that Thomas Jefferson was not present at Monticello at the time of conception of the only Hemings child linked to the Jefferson DNA marker.

But disregarding that biohistorically unspecified distribution, and sticking to the also problematical and unspecified one that would exist around Burton's and Neiman's nearly equivalent 266- and 267-day figures, the combined data show Jefferson absent for appreciable periods in fully four of the six cases that Neiman designated for study, including the sole case involving a DNA link. Neiman's probabilistic evaluation failed to account probabilistically, in other words, for two minor absences, a large one, and an enormous one.

So as a matter of credible biostatistical analysis, and also as a matter of common sense, the conventional wisdom that Jefferson was invariably, absolutely present at the times of Sally Hemings's conceptions is itself notably uncertain. Neiman's probabilistic evaluation fails to measure or account for that uncertainty. The study cannot serve as a scientific pillar of parenthood proof.

—

Maybe the Neiman study's deficiencies have finally begun to be recognized, at least obliquely. As of March 16, 2006, the on-line, evolving encyclopedia Wikipedia reported that the "*William and Mary Quarterly* published a probabilistic analysis of the timing of Jefferson's visits to Monticello and Hemings's pregnancies which concluded that it was highly unlikely the two series of events were unrelated." Unlikely they were unrelated? That's all? Though technically accurate, Wikipedia's understatement misleads almost comically about what Neiman in fact concluded, but in fact could not conclude scientifically. The grand scientific conclusion that Neiman actually claimed bears one last repetition: "Serious doubt about the existence and duration of the relationship and about Jefferson's paternity of Hemings's six children can no longer be reasonably sustained." Wikipedia's timid statement obscured the study's obvious fundamental problems.

Despite that obviousness, almost no one from science has criticized the study, even though historians assert it as a pillar of Hemings-Jefferson parenthood proof. To those historians' error in judgment R. B. Bernstein, introducing the pillar in his short Jefferson biography, adds another dimension. A "statistician," he reports concerning the archeologist Neiman, "analyzed the odds that anyone other than Jefferson was the father of Eston Hemings." Only Eston? From start to finish, Neiman's study actually seeks to determine the paternity of six children. Maybe that's why Bernstein then describes the visits-conceptions pattern. Next he asserts that Neiman

“concluded that the odds against anyone but Jefferson being the father were ten thousand to one.” The father of Eston? The father of six? Never mind; Neiman’s probability statements were actually lower by many orders of magnitude anyway. Both landed in the neighborhood of 1 in 100.

At what point does carelessness with scientific arguments become disrespect? At what point does carelessness with the authority of science become science abuse? Before science entered the Hemings-Jefferson paternity debate, Annette Gordon-Reed criticized a pervasive disrespect for historical evidence that had originated mainly in disrespect for owned humans and their descendants. Now that science has entered the debate, what explains the ways in which some scientists, editors, scholars, commentators, and journalists have mishandled science’s part in the debate, in the process disrespecting both science and history?