

When You Have a Hammer . . .

*The Misuse of Statistical Races*¹

Kenneth Prewitt

Carnegie Professor of Public Affairs, School of International and Public Affairs, Columbia University

Abstract

Race statistics and race policy have been intertwined in American history since its founding, starting with the infamous three-fifths clause, continuing with policies based on nineteenth-century race science, the restrictionist immigration at the turn of the century, the Jim Crow regime, and carrying into the civil rights era through such policy concepts as institutional racism, statistical proportionality, disparate impact, and affirmative action. Across this history, the policies and the statistics were about “race,” whether they punished or benefited, were racist or antiracist. But can there be policy that misuses race statistics, that is presented as about race when it should not be? Race statistics are a powerful policy hammer in American history, but not everything is, in fact, a nail. Today the census undercount is argued over as if it is about race; it isn’t really. Posing far greater danger, census race categories have worked their way into genomic medicine. The nineteenth-century belief that “race is biological” lingers in the American mind. The use of census categories in genomic medicine risks re-biologizing race. Maybe we should not leave the hammer lying around.

Keywords: Census, Racial Classification, Genomic Medicine, Race Policy, Biological Races, Census Undercount

INTRODUCTION

There is a long, tortuous interaction between America’s race policies and the “statistical races” produced by the U.S. Census and related government statistical programs. It starts with the 1790 census used to implement the infamous three-fifths clause (counting slaves as 0.6 of a person) that secured the hold of slave-owning states over many policies directly affecting their interests for the first half of the nineteenth century. Race science added other chapters to the story; census results in 1850, for instance, purported to prove that freeing Blacks from slavery would drive them insane, and were cited by Senator John C. Calhoun to justify extending slavery to the new state of Texas. This is probably the first instance of a broad sociological

Du Bois Review, 9:2 (2012) 281–301.

© 2012 W. E. B. Du Bois Institute for African and African American Research 1742-058X/12 \$15.00
doi:10.1017/S1742058X1200029X

finding, wrong that it was, being used in “evidence-based policy.” Other nineteenth-century chapters included how to count and treat the Chinese and then the Japanese, and of course the remarkable effort to use the census to keep the “wrong kind of Whites” from adding to their numbers at the turn of the twentieth century. The same census race classification that penalized minorities from 1890 until the mid-twentieth century made a 180 degree turn in the 1960s, when the Civil Rights Act ushered in statistically based policy concepts—institutional racism, disparate impact, statistical proportionality, and, of course, affirmative action—that attempted to undo a racist history. Multiculturalism, identity politics, interracial marriages, resurgent immigration, postracial talk, and a Supreme Court bringing color-blind considerations back into legal play are, today, raising fresh doubts about the government’s race classification—a classification that, remarkably, gives official status to the five race groups famously introduced in 1776 by Johan Blumenbach: Black, Brown, Red, Yellow, and White. Today the government tacks on the cumbersome ethnic distinction of Hispanic and Non-Hispanic, as well as the awkward “mark-one-or-more” option to allow the mixed race person to find a home on the census form.

Whether used to punish or benefit racial minorities, to reflect or help create racial identities, to advance progressive or conservative political agendas, or to broaden or restrict immigration, the long march of statistically based race policy across American history *almost* always made sense in one critical way: the statistics and the policy were, in fact, about race.

This article looks at the *almost*. It considers two cases where racial statistics are being used when they should not be—when a policy issue being racialized would be better were it not about race at all. Race statistics are one very big hammer in American history. Given the easy availability of this hammer, are we surprised to find policy issues that look like a nail—but aren’t? The first case is taken from the census itself, here labeled the Fair Census Controversy; the second considers the Human Genome Project.

The Fair Census Controversy is about a technical issue—the census undercount—that had the statistical races at its center. On close inspection, however, we learn that it need not have been—and *would have been better had it not been*—about race at all. There is a danger to policy making posed by the ubiquity of race statistics. When race statistics are available for policy making, policy can be inadvertently or inappropriately racialized. Color-blind advocates warn of this; they have a point. It is part of their critique of affirmative action, which they insist should not be about race or ethnicity but about social and economic disadvantage.

The Human Genome Project considers an issue altogether more consequential than the census undercount: *race-targeted medicine*. It hovers in the background of genomic medicine. If it comes to pass, it is odds-on that it will use the race classification at hand. It will recall the eighteenth-century assumption that the five Blumenbachian races are biologically real, and can be put to work on a biochemistry project: different medical treatments for different races.

THE FAIR CENSUS CONTROVERSY

Whether, at the margins, a particular statistical method can improve the census is not a trivial issue. It is, though, small scale in comparison to the three-fifths clause, the claim that freedom leads to insanity among Blacks, quota-based immigration policy, or affirmative action. Small scale though it was, the issue did—if briefly—fully and sharply engage every branch and level of government, and was caught up in an

intense partisan battle that spanned two decades. It soured statistical sampling if linked to the decennial census, and as a result the Census Bureau has largely discontinued technical work on a statistical methodology successfully used in other countries to improve census coverage.²

No census is a complete count of the population. President George Washington (1939) complained about the nation's first census: "One thing is certain," he insisted, "our real numbers will exceed, greatly, the official returns of them" (p. 329). Every president since Washington could offer the same complaint. Censuses are plagued by an undercount. Some people refuse to cooperate, asserting that their right to privacy trumps the government's need to know. Others don't want the bother and use unlisted phones, gated communities, or the protective wall of guards and doormen to shield them. Negligence on the part of census takers plays a role, though today quality-control procedures identify and correct such problems quickly. People who live alone and who have few community ties are missed, as are those who travel a lot. For those in the country illegally, fear of deportation motivates census avoidance (the fear is groundless; the census does not ask about legal status). Around the world, census offices typically design special efforts to reach the hard to count.³

Throughout American census history an undercount was routinely assumed and anecdotally discussed. But statistical theory was insufficiently advanced to provide reliable estimates of its magnitude and distribution. There were no good answers to seemingly straightforward questions: How many people are missed; what are their characteristics; where do they live?⁴ A moment's reflection suggests why. There is only one census. To what, then, can it be compared to determine how complete it is? To assess census coverage reliably would require two population counts so the results of one could be used to check the other. Historically this would have been prohibitively expensive.

Then, in the 1940s, there was an unexpected opportunity to compare the census with an independent and unusually accurate count of a major population group—men between the ages of twenty-one and thirty-five. These draft-eligible Americans were quickly registered following the declaration of war after the Japanese attack on Pearl Harbor. Though hardly its intent, the universal military registration gave demographers two independent estimates of young men—the first based on the 1940 census, the second from the compulsory registration. Comparing these two numbers provided the first reliable measure of how many people, at least among young men, were missed in the census.

From this starting point a method to estimate the census undercount for both genders and all age groups was developed. Called demographic analysis, it is in theory quite simple. It starts with a basic population number and then, using vital statistics and other administrative records, updates it by adding every birth and every arriving immigrant and subtracting every death and every person who moves out of the country.⁵

The Census Bureau went to work on reducing the undercount, and was remarkably successful (see Table 1).

It is the Differential Undercount that Matters

Most of the benefits of a census are allocated not in terms of absolute numbers but as proportionate shares of the total count. The most notable example is each state's representation in the House of Representatives. The 435 positions are apportioned among the fifty states in relation to what share each state has of the nation's total population.⁶ Federal grants-in-aid in education, transportation, and health are also allocated on a proportionate-share basis.

Table 1. Percent Net Census Undercount: 1940–1990*

Census Year	Percent Net Undercount
1940	5.4
1950	4.1
1960	3.1
1970	2.7
1980	1.2
1990	1.8

*Based on a comparison of census results with a population estimate from demographic analysis. The net undercount is calculated by subtracting the estimated number of duplicates in the census from the estimated number of persons missed. One source of the comparatively low net undercount of 1980 is the large number of suspected duplicates in that census, which has the result of increasing the overcount and thus depressing the net undercount. Census-to-census comparisons are difficult because procedures, including how duplicates are removed, vary from one census to the next.

This makes the differential undercount politically important. A differential undercount occurs when the census misses a higher percent of the population from one state or region compared to another. If the census misses 3% in every state, each state's share of the total population would be identical to its share if there were no undercount. But what if the percentage of persons missed varies from one state to the next. The state with a higher percentage missed would get less than its fair share of federal funds and could even unfairly lose a seat in Congress.⁷

The differential undercount can apply to demographic groups as well as geographic regions. The census could miss a higher proportion of the elderly than the middle-aged, or of women than men. More consequentially for the claim made in this article, the census could miss a higher proportion of racial minorities than it does of Whites. That, in fact, seemed to be the case. We saw above that young males in the country, as counted by the mandatory military registration in 1941, were more numerous than the 1940 census had counted. The country in the 1940s was of course racially segregated, and when young men showed up at the draft board they were assigned a race, mainly White or Black.

The population statisticians who compared the census with the draft records took note that a proportionately higher share of young Black males had been missed in the 1940 census than had young White males. The key notion of differential undercount entered census lore as a differential *racial* undercount. Table 2 shows that the magnitude of this differential undercount persisted despite the Census Bureau's success at reducing the undercount overall.

The Differential Undercount and the Politics of Racial Justice

The visionary social scientist, Daniel Patrick Moynihan, then a Harvard professor, knew that the Census Bureau did a better job at counting Blacks in 1960 than it had in 1950 or 1940. But he would not have missed the troubling fact that the Black/White gap had grown over the same period. In 1967, Moynihan organized a confer-

Table 2. Percent Census Undercount by Black/Non-Black: 1940–1990*

	1940	1950	1960	1970	1980	1990
Black	8.4	7.5	6.6	6.5	4.5	5.7
Non-Black	5.0	3.8	2.7	2.2	0.8	1.3
Difference	3.4	3.6	3.9	4.3	3.7	4.4

*Based on demographic analysis.

ence on social statistics and the city. The published volume from that conference (Heer 1968) drew the obvious conclusion:

Where a group defined by racial or ethnic terms and concentrated in special political jurisdictions is significantly undercounted in relation to other groups, then individual members of that group are thereby deprived of the constitutional right to equal representation in the House of Representatives and, by inference, in other legislative bodies. They are also deprived of their entitlement to partake in federal and other programs designed for areas and populations with their characteristics. In other words, miscounting the population could unconstitutionally deny minority political representation or protection under the Voting Rights Act. It could also deny local jurisdictions grant funds from federal programs (p. 11).

Thus a technical problem in census taking, the undercount, moved quickly into the fractious politics of racial justice. The Census Bureau came under enormous pressure to fix the differential undercount.

Although demographic analysis could document the inequity in census numbers, it was not a method that could fix it. Demographic analysis provides a national estimate of the undercount, but the inequity occurs at local levels—the very place where congressional districts are drawn and federal funds are allocated. The census needed a more sophisticated statistical method if it was to fix the undercount problem.

Dual system estimation (borrowed from wildlife research) became the favored method: first take the census and then immediately follow it with a large sample survey. If correctly done, the sample survey offers a second estimate of the population that, when compared to the census, allows adjustment—block by block—that aligns the census with the “true” number of people in the country.⁸ Starting in 1950, the method was steadily improved from one census to the next, though not until 1990 was the bureau confident that it could be applied to adjust the undercount.

Whether to adjust the undercount had by 1990 become the focus of an intense, nasty partisan battle. The initial salvo was legal action by Democratic leaders in 1980 trying to force the Census Bureau to adjust that decennial census, even though the bureau’s statisticians felt that the adjustment methodology was not yet robust enough to use. The court ruled in favor of the Census Bureau but instructed the bureau to continue working on the methodology in expectation that it would be used in 1990 (Anderson and Fienberg, 1999; Darga 1999).

The politics did not let up. Because the electoral strength of the Democratic and Republican parties was closely balanced and because the party margins in the Congress were close, small changes in census numbers used to reapportion and draw electoral boundaries could determine party control of Congress. There was political pressure to squeeze every possible advantage in the redrawing of congressional districts after

the decennial census. Increasingly sophisticated computer-assisted methods allowed the two parties to make fine-grained assessments down to the block level. The Democratic Party, responsive to pressure from civil rights groups, believed that census adjustment of the undercount might produce more congressional districts favorable to Democratic candidates. For this reason, Republican Party strategists strongly resisted the adjustment method. Republicans demanded a traditional census, which works well in the White suburbs, where much of their support is concentrated.

In 1990, the Census Bureau was sufficiently confident of its methodology that it proposed to adjust the census apportionment counts. It was overruled by the Republican Secretary of Commerce (the Bureau is a unit of the Commerce Department).⁹ Census method had collided with partisan interests. Both sides continued to dress their arguments in high-minded language. Democrats spoke of fairness, not partisan advantage. They insisted that the Census Bureau be allowed to apply whatever scientific methods it believed would improve census accuracy. Republicans cited the Constitutional provision that an actual enumeration be taken as reason to reject any plan using sampling. Both sides found support among reputable statisticians.¹⁰

The technical challenge of producing an accurate count did not go away, and the Census Bureau continued to improve dual system estimation. After the election of Democratic President Bill Clinton in 1992, the Bureau's technical work had more political protection, but partisan polarization reached new highs when the 1994 congressional elections brought to Congress a number of conservative Republicans deeply mistrustful of Clinton. The congressional election gave Republicans control of Congress, which they held during the entire period in which the 2000 census was being planned.¹¹

Congressional Republicans who reviewed the census plans and appropriated funds were told that allowing the census to be adjusted would have a "negative effect in the partisan makeup of 24 Congressional seats, [and at the state level] 113 State Senate seats and 297 State House seats nationwide . . ." (Nicholson 1997, p. 1).¹² Charges of census tampering were in the air. The Congressional Black Caucus—all Democrats—took up the census as a leading civil rights issue; they were often joined by Hispanic and Asian members of Congress. The Leadership Conference on Civil and Human Rights (2000), a leading civil rights coalition with 200 member organizations, framed the issue:

Because the accuracy of the census directly affects our nation's ability to ensure equal representation and equal access to important governmental resources for all Americans, ensuring a fair and accurate census must be regarded as one of the most significant civil rights and equal rights issues facing the country today (p. 2).

When Democrats played the race card by accusing Republicans of derailing a fair count of minorities, Republicans answered by voting funds for efforts targeted to reducing the minority undercount: more money for community partnerships and promotional campaigns in minority neighborhoods and higher salaries for enumerators working in difficult-to-count areas. The Republican Congress, the Democratic President, numerous advocacy groups, and widespread media coverage turned the 2000 census into a debate about racial fairness.

But Was Race the Issue?

Are households missed in the census because they are Black or included because they are White? No. They are missed when the Census Bureau's address file has errors,

when the household is made up of unrelated persons, when household members are seldom at home, when there is a low sense of civic responsibility or perhaps an active distrust of the government, when occupants have lived at the current address a short time and will move again soon, when English is not spoken, when community ties are not strong. Race was shorthand for many household characteristics that themselves were not racial. The long, nasty partisan debate over census methodology need not have been about race at all. It happened only because the U.S. government counts by race.

The first reasonable estimate of the differential undercount had a racial dimension *only* because the draft system set up in 1941 recorded the race of every registrant. Birth and death records also record race. So when statisticians used these vital statistics in demographic analysis, they had a racial breakdown easily at hand. From that time on, the undercount was always discussed as if it were about race. Table 2, documenting the undercount by race, could not be constructed using, for example, a measure of social isolation or of linguistic barriers to completing a census form. Those measures are not available in the vital statistics on which demographic analysis is based.

Shifting to the statistical method of dual system estimation continued the practice, again for reasons that were a by-product of other considerations. Dual-system estimation relies on census counts at the block level; these are available only from questions that are asked on the census short form.¹³ Very few questions are asked on the short form; these few include age, race, and Hispanic ethnicity. This information is required by the Voting Rights Act of 1965, which is administered using block-level counts of the voting-age population broken down by age, race, and ethnicity.

In 1990 and 2000, however, the short form also asked whether the home was rented or owned. The Voting Rights Act does not use this question. Why was it necessary to have block level data on rental status, but not, for example on marital status or veteran status or any of the other fifty or so variables on the census long form? Because the Census Bureau had determined that home ownership is a useful predictor of census coverage. The rental/ownership item correlates with census cooperation, and it was on the short form so that it could be used in dual-system estimation and census-adjustment calculations. It is a far-fetched hypothetical, but if every young male registered for the draft had been asked not his race but if he came from a household that he rented or owned, and if administrative records on rental and ownership status were accurate in 1941, the entire history of the census undercount could have turned on that variable rather than race.

In fact there are many household characteristics that in principle might offer much better predictors of census coverage than does race; these include language skills of household members, length of time in the country, education and income levels, marital status, and whether there are other family members living nearby.¹⁴ These characteristics may correlate with race, but that doesn't make race the predictor variable.

The U.S. Bureau of the Census (2010) itself has now recognized this. In preparation for the 2010 census, the bureau identified the especially "hard-to-count" census tracts. Six "person characteristics" and six "household characteristics" were used.¹⁵ Race was *not* among the characteristics, though when the bureau reported to the public on the 2010 undercount it did so on the basis of the rate at which racial groups were missed—still more than 2% for Blacks and nearly 5% for Native Americans. Although the Census Bureau had concluded that characteristics other than race were more likely to predict census coverage, what was available—the statistical races—once again became the lead story.

I use the census-adjustment issue to emphasize that, in the policy and political worlds, what is easily available is what is used—even if on close inspection it is poorly matched to the task at hand. In the 1960s, what was available to the government—how many Blacks there were in the labor pool and how many Blacks got jobs in the construction industry—became the foundation of affirmative action policy. This made sense. Affirmative action was a policy designed to end racial discrimination and improve employment of racial minorities. It was about race, as is the Voting Rights Act that uses block-level race data to challenge election boundaries that weaken minority-voting power.

In the 1960s, what was available to the Census Bureau—how many minorities were recorded in the nation's vital statistics and how many were counted in the census—became the foundation of a statistical method to improve the decennial census. But, unlike affirmative action or voting rights, adjusting the census is not *per se* a race issue. If knowing that someone who rents or is unmarried or has no telephone is a better predictor of whether that person is counted, then that characteristic, not race, should be at the center of a statistical procedure to improve the census.¹⁶

Racializing the undercount was unfortunate, as was the partisan debate over sampling to improve the census more generally. It would have been better for the decennial census and for the quality of political debate in the nation had it not happened. But it did, and today it hampers progress toward census improvement. If the hammer-nail truism were limited to the example just summarized—a statistical method in census taking—it could be brushed aside as only marginally important in the big sweep of American politics. If, however, the truism is relevant to an issue of much broader importance to America's views about racial differences, we should have second thoughts about leaving the hammer lying around.

THE HUMAN GENOME PROJECT AND RACE

Small genetic differences between different population groups can be traced to the ancestors of modern humans migrating from Africa, to where the migratory streams settled, and for how long and how completely they were isolated from each other (Oppenheimer and Bradshaw Foundation, 2003). The earliest archaeological evidence of our genetic ancestors dates to about 160 thousand years ago, and it clusters in a small area of Eastern Africa's Rift Valley. Over the next twenty-five thousand years humans spread to southern and western Africa, then gradually down the Nile. Not until ninety thousand years ago was there a permanent settlement along the southern coast of the Arabian Peninsula. All non-African people are descended from this group.

Over the next ten to fifteen thousand years, humans made their way along the Indian Ocean coast, settling as far south as today's Indonesia and as far north as lower China. In today's terms, three race groups—Africans, Asians, and Pacific Islanders—were then identifiable. Following an ice age that dramatically lowered the world's population and temporarily disrupted migration, humans spread to Australia and up the western coast of today's China. Not until about fifty thousand years ago was the Fertile Crescent settled; this also was the time frame in which humans first crossed the Bosphorus into today's Europe, creating the fourth of Blumenbach's (and the Census Bureau's) five primary races.¹⁷ Another twenty-five thousand years passed before humans reached northeast Eurasia, the departure point for crossing the Bering land bridge connecting Siberia to Alaska. The last Ice Age again slowed migration, leaving a few isolated groups in North America. Then, about fifteen

thousand years ago, humans reached South America, and gradually spread along the Pacific coast line and into the interior of Northern and Southern America, adding the fifth race to the world's population.

Blumenbach's climate theory of race differences is based on the five continental descent groups that were formed in the 150 thousand years it took for *Homo Sapiens* to spread across the planet from their origins in Africa's Rift Valley. The basic conclusion of modern genetic science is, of course, that it is nonsense to think that there are character traits such as inventiveness or indolence that can be traced to the small amount of genetic variation from one continental descent group to another (Holden 2003).

Evolutionary biologists and geneticists concur that the variations in humans that correlate with conventionally defined races or geographic descent groups are comparatively small, and in any case the within-group variation is much greater than the between-group variation. Contemporary genomic medical science has not yet reached consensus on whether paying attention to even these small between-group differences can advance medical treatment. This issue comes up because certain diseases are more often found in particular race or ethnic groups. Examples frequently cited include the prevalence of sickle-cell anemia among Africans and Tay-Sachs among Jews from northern Europe.

The founder effect in evolutionary biology explains why descent groups might vary in their susceptibility to particular diseases. Although the migratory flows sketched above started from a common gene pool, migration to East Asia led to a people isolated by culture and geography from those who migrated to Europe, and so forth. Because of inbreeding, random changes from one generation to the next (known as genetic drift) could produce genetic traits common in one region and absent in others. When a mutant gene led to a rare disease, the group reproducing in isolation from the rest of the human species would, over multiple generations, share that disease susceptibility more frequently than unexposed groups. For example, because of a sickle-cell mutation, equatorial Africans and their descendants today have comparatively higher resistance to *falciparum malaria*, the type most likely to prove fatal and fairly common in tropical Africa. This particular mutation has a down side, however. The unusual formation of the red blood cells relevant to malaria resistance makes it harder for them to absorb oxygen, and this has some predictable negative effects. There are dozens of such examples, as Orr (2006) makes clear.

Genetic clustering by descent group means that the human species "possesses what population geneticists call 'population structure': that is, certain DNA sequences are found at slightly higher frequencies in Africa, others in Asia, and so on. . . . These findings also mean that different groups of people might differ in subtle aspects of their biochemistry" (Orr 2006, p. 20). Given what is known about population genetics, population structure on a small fraction of the human genome is not surprising.

Geneticists have used computer programs to identify population clusters based only on genetic information. Although research on population clustering is in its infancy, based only on subsets of genes, and early findings will no doubt be modified, one research team found four clusters that map to the continental descent groups, or what today are taken to be four of the world's major ethnoracial groups. In this study, research subjects, not knowing what descent group their genes had assigned them to, were asked their race using the standard racial classification categories. Overwhelmingly, the research subjects' racial self-identification matched the population cluster they had been assigned by the computer program using only genetic information (Risch 2005).¹⁸

With findings such as this it is understandable that the ability of science to map the human genome has accelerated the search for associations between diseases and gene expression. When those associations appear in one ancestry (race) group more

frequently than another, medical science has reason to argue that knowing a patient's ancestry is diagnostically useful. For many health researchers and medical practitioners, not to use that diagnostic tool would be medically irresponsible. It could retard health improvements for the group in question. In this reasoning, if those population groups are conventionally labeled by race terms—Europeans, Africans, and so on—this does not lessen the importance and usefulness of the diagnosis. However, not all scientists agree with this reasoning.

Those disagreeing do not question that genes are powerfully predictive of susceptibility to various diseases or that pharmacogenomics has great promise in the treatment of illness and improvement of health. However, argue many scientists, it is the genes that matter and *not* the race of the person carrying those genes. Race categories found in the nation's statistical system were created for political and policy reasons that have nothing to do with medicine. In fact, this argument continues, genetic medicine loses efficiency by using race as shorthand. Even if certain race groups are more prone to particular diseases, not everyone of that race is susceptible and not everyone susceptible belongs to that race. Race-targeted drugs and medical treatment will necessarily and often miss the real target.

While Waiting for Personalized Medicine, An Interim Strategy?

The response says, "You are right, but you aren't being practical." When geneticists have analyzed every person's individual genome and doctors can tailor person-specific treatment, the great promise of genomic-based personalized medicine will (perhaps) be realized. But the cost is enormous, even for the richest nations of the world. Personalized medicine is not around the corner.¹⁹ Race-targeted diagnostics and therapeutic intervention are accepted as a reasonable first step, a placeholder while medicine works toward personalized medicine, according to this argument. As geneticist David Goldstein remarked, if individually tailored medical treatment is a distant promise, "the big question is the *interim strategy*: how to use ancestry now" (Holden 2003, p. 596; italics added for emphasis).

It is this interim strategy that has many worried. Under the sponsorship of the Wellcome Trust, a private foundation, scientists from the United Kingdom offer a thoughtful reflection and warning on using ancestry (race) as an interim strategy as we await the promise of genomic-based personalized medicine (Martin et al., 2007). Their argument merits summary (and attention by scientists in the United States, where no comparable review exists).

The Wellcome Trust report starts with a familiar fact: results of medical studies routinely report differences across racial groups. This comes as no surprise; nor does the fact that they are attributed to everything from prenatal care to lifestyle choices such as smoking, diet, and exercise, which vary from one group to the next. At issue is not whether differences in disease and health vary by race group, or whether environmental factors specific to experiences of race groups have explanatory power. At issue is whether the definition of race found in official statistics is a suitable variable when attention shifts from environmental to genomic explanations of disease susceptibility and optimal therapeutic treatment. It was this issue that motivated the U.K. investigation of how contemporary genetic and biomedical research makes use of race and ethnic census categories.

Biomedical scientists in the United Kingdom routinely apply a race classification in their research. The classification is not, however, constructed scientifically. It is borrowed from the U.K. census. Why? Because the biomedical scientists report that these race categories are:

. . . felt to have proven practicability and portability—i.e., they had political legitimacy; they were acceptable to the public; they were easy to use; they permitted comparisons between studies; and they facilitated the translation of research findings into clinical practice (Martin et al., 2007, p. 6).

The scientists who offered “political legitimacy” and “easy to use” as reasons for borrowing such a key variable as race from political sources are not altogether comfortable about it. It may be safe, they reason, to use the official categories for drawing samples or as descriptive variables. However, it:

. . . seems unlikely that “official” socio-political categories such as these [from the U.K. census] will be useful for identifying the underlying causes of any observed differences in disease susceptibility or therapeutic efficacy, because these categories can only ever offer crude proxies for the complex interplay of structural, socio-cultural and genetic factors involved (Martin et al., 2007, p. 6).

Moreover, points out the report, the official categories developed for use in the U.K. censuses:

. . . arose out of complex social and political debates about immigration, representation and equality, as well as from external and internal claims for separate group identities that continue to this day. As such, these categories are not fixed but have changed over time and are likely to continue to do so in response to shifts in emphasis within social and political arenas (Martin et al., 2007, p. 4).

Medical scientists unfamiliar with this history might adopt the categories as analytic variables, which can “prove problematic . . . due to their socio-political origins and inherent flexibility over time and place” (Martin et al., 2007, p. 5). Census categories operate over a time frame measured in decades; genetic science operates on a time scale measured over millennia.

If the problems inherent in applying census categories in biomedical research were only scientific in nature, scientists, it could be assumed, would identify and correct these problems. But there is a much more urgent alarm to sound, and the thoughtful report does so in its title: *Reviving “Racial Medicine?”* Could the use of census race categories in genetic research, because they are practical and because they have political legitimacy, bring a return of eighteenth-century racial medicine?

The authors of *Reviving “Racial Medicine?”* are worried.²⁰ Applying the official categories in genetic research “runs the risk of ‘geneticising’ or ‘biologizing’ such categories by treating them *as if* they reflected a reliable, valid and natural classification of discrete and biologically homogenous racial/ethnic groups” (Martin et al., 2007 p. 5; italics added for emphasis). It is this *as if* that poses the moral challenge to genetic medicine. If scientists can fall into the *as if* trap—can treat a census classification as if it is a truth of nature—is not the general public likely to come to the same conclusion? If you have a hammer, does not everything look like a nail?

The *As If* Risk in the United States

Not surprisingly, U.S. scientists have engaged issues similar to those debated in the United Kingdom. The match between the five races in the U.S. census with population clusters identified by evolutionary biologists and geneticists make it inevitable that there are worries about treating census races as genetically bounded population

groups. According to H. Allen Orr (2006) the fact that certain DNA structures appear at higher frequencies in different ancestry groups does “*not* mean that human races necessarily or even probably differ in profound ways” (p. 20). It does mean that subtle differences in biochemistry from one group to the next could, at the margins, be relevant to disease susceptibility or responsiveness to pharmacological treatments.

In a widely cited paper, the Stanford University geneticist Neil Risch and colleagues argue this point. They claim that “from both an objective and scientific (genetic and epidemiologic) perspective there is great validity in racial/ethnic *self-categorizations*, both from the research and public policy points of view” (Risch et al., 2002; italics added for emphasis) and that people categorize themselves racially in genetically meaningful ways. After reviewing the major efforts by population geneticists to identify clusters, they reach this conclusion:

Effectively, these population genetic studies have recapitulated the classical definition of races based on continental ancestry—namely African, Caucasian (Europe and Middle East), Asian, Pacific Islander (for example, Australian, New Guinean and Melanesian), and Native American (Risch et al., 2002).

If these differences are cosmetic only, they will have no consequences for medical research. But they are not only cosmetic; they are genetic, suggest Risch et al. (2002), if that term:

. . . is defined by susceptibility to, and natural history of, a chronic disease, then again numerous studies over past decades have documented biological differences among the races. In this context, it is difficult to imagine that such differences are not meaningful.

It is, therefore, irresponsible to ignore these differences in medical research.

A “race-neutral” or “color-blind” approach to biomedical research is neither equitable nor advantageous, and would not lead to a reduction of disparities in disease risk or treatment efficacy between groups. Whether African Americans, Hispanics, Native Americans, Pacific Islanders, or Asians respond equally to a particular drug is an empirical question that can only be addressed by studying these groups individually. Differences in treatment response or disease prevalence between racial/ethnic groups need to be studied carefully; naive inferences about genetic causation without evidence should be avoided. At the same time, gratuitous dismissal of a genetic interpretation without evidence for doing so is also unjustified.

Elsewhere, Risch is quoted as saying that ignoring race in gene studies will “lead to the disservice of those who are in the minority” (Wade 2002, p. 2). Many in the biomedical community echo his arguments. The Nobel Laureate Francis Collins (2004), appointed Director of the National Institutes of Health by President Obama, is quick to say that race and ethnicity “are poorly defined terms that serve as flawed surrogates for multiple environmental and genetic factors in disease causation, including ancestral geographic origins, socioeconomic status, education and access to health care” (p. 13). Flawed though they may be, Collins is nevertheless endorsing the statistical races as surrogates for genetically based diseases. “It would be incorrect to say that genetics never has a role in health disparities,” he continues (p. 13). The bottom line: “We must continue to support efforts to define the nature of human variation across the world, focused primarily on medical goals” (p. 14).

If many scientists are in agreement with Risch and Collins, many are not. Among the naysayers are most social scientists, but they are joined by major biologists and

geneticists. Richard Lewontin of Harvard University, author of a classic 1972 paper which argues that races as conventionally defined are not biologically significant, has since written extensively on evolution and human variation. About the genetic studies cited by Risch, Lewontin (2006) points out that the only way in which to cluster populations that correspond to the classical continental races is to use:

a special class of non-functional DNA microsatellites. By selecting among microsatellites, it is possible to find a set that will cluster together African populations, European populations, and Asian populations, etc. These selected microsatellite DNA markers are not typical of genes, however, but have been chosen precisely because they are “maximally informative” about group differences. Thus they tell us what we already knew about the differences between populations of the classical “races” from skin color, face shape, and hair form (p. 2).

If Lewontin is correct, it is unlikely that the clusters produced by microsatellites will lead to major advances in medical science. The population geneticist, David Goldstein of University College, London, makes this point. He believes that the best science occurs when patients are assigned to different groups based only on their DNA, and not in groups based on racial self-categorization. To understand the geographic pattern of human genetic variation, notes Goldstein, “You want the best representation you can find, and it is a technical question as to whether explicit genetic representation or racial labels are better. That’s an argument we will have in the scientific literature, and Neil [Risch] will lose” (Wade 2002, p. 2).

Other scholars, including Duster (2005), warn against the fallacy of misplaced concreteness, the assumption that what is labeled a race “coincides with the obdurate character of the empirical world” (p. 1050). An example of misplaced concreteness is Risch’s assumption that the number of Americans who self-identified as being of more than one race in the 2000 census coincides with the true number of multirace Americans. Risch et al. (2002) write about the “mark-one-or-more option” introduced in the 2000 census:

According to these numbers, if mating were at random with respect to these racial categories, 42% of individuals would result from “mixed” matings and hence derive from more than one race, as opposed to the 2.4% reported. These figures highlight the strong deviation from random mating in the U.S.²¹

Mating in the United States is of course not random. But whatever the true number of Americans with DNA traces from more than one race group, it is not 2.4%, no matter what is put down on a census form. Risch is extremely careful in how he handles genomic data in his writings. He is careless in his treatment of demographic data, especially self-reported race in the U.S. census. The mark-one-or-more responses to the 2000 census are scientifically unreliable. Of those who gave multiple-race responses in 2000, 40% answered differently to a follow-up quality survey conducted by the Census Bureau a year later. As Bennett (2003) writes, of those who did say that they were multiple-race in the follow-up survey, nearly half (45%) had declared they were in one race group a year earlier.²² A genetic experiment based on data that unreliable would be discarded.

Of course the census race categories can find their way into genomic medicine without relying on the self-reported data collected by the census itself. DNA is inherited; this implies ancestors and descendants. Descent groups have a “genomic geography.” This phrase “refers to how, through the tools and practices of human

genetics, bits of genomic sequence become associated with specific geographic locations, posited as the place of origin of people who possess these bits" (Fujimura and Rajagopalan, 2011, p. 7).

Joan Fujimura and Ramya Rajagopalan (2011) offer a sobering account of the slippery slope from "genomic geography" to "biological races." Genomic geography, they observe, "is one thread between population, race, and 'genetic ancestry' that renders the three concepts difficult to untangle" (p. 3). Based on ethnographic research in the laboratories and professional meetings of genetic scientists, they conclude that the boundary between genetic ancestry and the government's "non-scientific" race classification is easily and frequently blurred.

Scientists using an analytic technique specifically designed to avoid race categories are vulnerable to downstream translation by colleagues who confuse "clusters of genetically similar samples" with "samples with similar 'genetic history' or 'shared ancestry,'" and then take the latter to imply race or ethnicity (Fujimura and Rajagopalan, 2011, p. 15). They are vulnerable even to the public relations efforts of their own universities, which "often use simplifying, provocative terms in press releases about their research results" (p. 16).

It is beyond the scope of this paper (and this author's competencies) to offer an informed view on whether statistical races, based on self-categorization or on technologies that create clusters based only on DNA, can be of worth in medical research. Genomic research is in its infancy; there is great uncertainty over the pace, scope, and validity of "race-targeted medicine." One thing we do know: in June 2005, the Food and Drug Administration (FDA) licensed a heart drug, under the label BiDel, exclusively for use by Black patients of African descent (Kahn 2007). This generated an intense scientific and political debate; it is not yet clear how effective this drug is, or whether non-Black patients have been put at risk by targeting it as suitable only for Black patients. Seven years into the very well funded and extremely active field of genomic medicine, it remains the only race-targeted medicine ever marketed.²³

Our focus, however, is not the future of race-targeted medicine; it is the risk associated with the blurry boundary between genomic medicine and the government-sanctioned statistical races. As Fujimura and Rajagopalan (2011) report, the National Institutes of Health (NIH) encourage "the inclusion of women and racial minorities in clinical research and mandates that practitioners in clinical and basic biomedical research receiving federal funding should report on the diversity of their research subjects," using the Office of Management and Budget (OMB) race and ethnic categories (p. 22). The NIH, however, do not require that the race categories be used in analysis, and Fujimura and Rajagopalan find that geneticists they interviewed do not interpret the NIH reporting requirement as an edict to use race in their research.

Other government initiatives, however, have had the effect of promoting the OMB categories in genomic research. A large international collaboration, launched in 2002 as the HapMap project, brought together two prior efforts: the government-sponsored Human Genome Project, with laboratories from around the world, and an international consortium composed of large pharmaceutical companies. The HapMap Project is tasked with cataloging common genetic variants in human beings, with the goal of locating patterns that frequently occur within persons having particular health risks. Four population groups were sampled for the HapMap project: Americans of European ancestry, Nigerians, Japanese, and Chinese. A plan to include a sample of Native American Indians met resistance; this project, then, represents only three race groups: European, African, Asian (Bullard and Dudoit, 2008).

The international collaborators involved in HapMap include countries with widely different census racial classification systems, none of which matched the

American five races/one ethnicity structure. In setting up HapMap there was intense argument in the scientific community about the proper sample design. Some scientists promoted global grid sampling, which would ignore previous notions of who was related to whom and would not take into account national or continental boundaries as they exist today. The alternative was to recognize predefined ethnic, linguistic, and racial populations as a guide to reproductive isolation (Bliss 2009; Roberts 1992). This strategy is adopted in the HapMap sample.

Decisions about how to sample human populations for genomic research were at various critical steps influenced by the American government, especially by the NIH and the FDA. These agencies have different constituencies, medical and health researchers in the former case and drug companies conducting clinical trials in the latter. Both, however, applied the OMB-recommended race classification that, of course, was constructed for the U.S. census and related population surveys. America's statistical races have now migrated into international collaborations mapping the genome of the world's population.

Science is only a few decades into genomic medical research, and I reemphasize that much will shift as new discoveries are made. Yet, also worth reemphasis, category systems have a lot of staying power in research and in policy.²⁴ That worldwide genomic research practices and clinical trial protocols are incorporating a basic racial classification system from the American census increases the stakes for the *as if* risk articulated in the U.K. Report. There is the risk that census statistical races adapted in genetic research will be treated *as if* they were discrete, homogenous groups found in nature.

Why the *As If* Risk Matters

Despite the triumphs of the civil rights era, the public mind has not completely dismissed the thought that persistent racial differences in educational attainment or rates of criminal activity signal differences in aptitude and attitude that, perhaps, are biological. Herrnstein and Murray's (1994) *The Bell Curve: Intelligence and Class Structure in American Life*, a dense academic book suggesting that enduring racial differences in genetically inheritable IQ help account for who fails and who succeeds in American education and employment markets, reached the best-seller list in 1994. One of its authors, a political scientist, has also offered an evolutionary and genetic explanation for elevated Jewish intelligence (Murray 2007). Although genetic explanations for IQ are strongly denounced by other scholars, they retain a measure of academic respectability.

More generally, Ann Morning (2011) has shown that racial essentialism remains pronounced in what students are taught in high school and college. Textbooks, for example, routinely present genetic explanations of diseases in terms of the familiar census race categories. The assumption that "race is biological" is not only true across the biological sciences, it is found in textbooks used in physical anthropology, world history, and often in the social sciences—with the only exception being social anthropology, where the emphasis is on races as socially constructed.

In this context, it seems plausible that race-targeted medicine could legitimate other biologically based arguments about racial differences, perhaps that the brain evolved faster in some races or that there are genetic predispositions to violence. Medical science does not take place in a social and political vacuum. When the FDA licenses a heart failure drug it sends a message to an audience much broader than merely the doctor who prescribes it and the patient who takes it.²⁵ It is that broader audience and what it is prepared to hear, perhaps even wants to hear, that raise nonmedical questions about medical practices.

And the message does not come only from medicine. In the lead article in a prestigious political-science journal, Charney and English (2012) report “a significant association” between genetic traits and voting behavior for African Americans, but not for Asians, Native Americans, and non-White Hispanics (p. 6). The authors offer a complicated genetic-based explanation, one that cautions against the conclusion that differential voter turnout across the race groups has a genetic base. The explanation does, however, include this passage: “. . . a gene that is in linkage disequilibrium with a polymorphism of MAOA and occurs at a higher rate among African Americans predicts voter turnout” (p. 6). There will be more such articles, always cautious, but still leaving an impression that racial genetic differences have predictive power for behaviors—voting in this case—well beyond health issues.

“Race in America has always explained too much and too little,” writes the historian of science Evelyn Hammonds (2006; p. 2). She continues, “Yet, Americans are deeply attracted to and readily accept racial narratives—especially when they are produced by biology” (p. 2). The arrival of genetic medicine can easily deepen and strengthen this narrative, as race-as-biology moves from disease incidence to intelligence to criminology to social violence to voter turnout. Genetic science might stop at disease incidence. This will not prevent the political resurgence of a racial narrative that appropriates a scientific metaphor for seemingly analogous public policy questions.

CONCLUSION: BIOLOGY’S CENTURY

We might not worry about this so much were it not for the powerful way in which biology is reaching into our lives. It is now commonplace to say that what physics was to the twentieth century, biology will be to the twenty-first.

As Freeman Dyson (2007) writes, “biology is now bigger than physics, as measured by the size of budgets, by the size of the workforce, or by the output of major discoveries; biology is likely to remain the biggest part of science through the twenty-first century. Biology is also more important than physics, as measured by its economic consequences, by its ethical implications, or by its effects on human welfare” (p. 4).

The above prediction leads Dyson (2007) to an even more audacious claim: “The domestication of biotechnology will dominate our lives during the next fifty years at least as much as the domestication of computers has dominated our lives during the previous fifty years” (p. 4). By domestication he means, for example, the biological equivalent of a computer in every household, down to video and computer games for children, games that will be played with actual eggs and seeds. The winner of a science fair might “be the kid whose egg hatches the cutest dinosaur” (p. 4).²⁶ Even if this image is far-fetched (and it might not be), it is certain that we enter an age of greatly increased attention to biology and genes.

It is no small matter, then, to ask: Is this not an exceptionally risky time to adopt race as an interim strategy on the road to personalized genomic medicine, an interim strategy that *only* makes sense if race is treated as biologically real? Perhaps medicine will reach its goal of individually tailored diagnostics and health treatments, and will then set aside its interim strategy. Even if this occurs within several decades (very unlikely), having used biological race as an organizing principle in the interim will leave a sizeable imprint on public consciousness (Varmus 2009). It is a judgment call whether the expected benefits of reporting research and, if indicated, targeting drugs to race groups, outweigh the danger that political interests and social processes

unrelated to medicine will latch onto the “new proof” that race is biological. I believe that danger is real, despite reassurances by the biomedical community.

Neil Risch and colleagues (2002), among many others, have provided a carefully argued scientific rationale for “identifying genetic differences between races and ethnic groups,” emphasizing what can be learned about disease susceptibility or variation in drug response. Risch et al. recognize that, “Great abuse has occurred in the past with notions of ‘genetic superiority’ of one particular group over another.” Reflecting the view held by the vast majority of scientists engaged in genomic research, Risch et al. insist that the “notion of superiority is not scientific, only political, and can only be used for political purposes.”

It is of course these *political purposes* that are worrying. In the 1930s, reputable geneticists were tragically naïve in failing to understand the political purposes to which eugenics was being applied in the laboratories of Hitler’s Germany (Barkan 1992; Black 2003). I intend no direct analogy to the current adoption of the race classification by modern genomic science, but history confirms time and again that science has few defenses against the political appropriation of its results. Among the physicists who developed the atom bomb were many who expected that it would never be dropped on a civilian population. It was. Among the eugenicists were many who expected that their research and theories would never be used to justify genocide. They were.

Reputable geneticists today can insist that their science does not imply or impute racial inferiority, but they are mistaken if they think that others will forego the racist ammunition they are providing. Hanna Arendt taught us that the Enlightenment did not banish radical evil. Concentration camps, she wrote, must cause “social scientists and historical scholars to reconsider their hitherto unquestioned fundamental preconceptions regarding the course of the world and human behavior” (Katznelson 2005, p. 72). A political rebirth of scientific racism is not out of the question. We hear echoes in how the French speak of the Roma, the Dutch speak of Muslims, and the Americans speak of undocumented Mexicans—all countries proud of their racial tolerance, but in some quarters now given to xenophobic and racist assertions about the foreign born.

I am not predicting a resurgence of racist political doctrine. I am claiming that it is naïve to assume its impossibility. If it does reemerge, in whatever guise, race-targeted medicine will not be an innocent bystander. It certainly will not be so if racial terms are superimposed on continentally mapped genetic variation, and this becomes a commonly accepted global vocabulary via genomic research. In this context, what are we to make of the government, in the agency of the OMB telling the country that its official census categories “should not be interpreted as having anthropological or scientific origins;” and that they are a “product of U.S. political and social history” (Wallman et al., 2000, p. 1707). The status of this social constructionist assumption is not so clear if the government, in its National Institutes of Health research or its protocols for drug trials, is also saying that the census classification is being used by genomic researchers because they have “documented biological differences among the races” (Risch et al., 2002, p. 4).

The voices insisting that race is socially constructed will find themselves defending what will be dismissed as “political correctness” against arguments from the “real science of biomedicine.” If it comes to this, if it is “my science is better because I give African Americans a drug that will protect them against heart attacks,” OMB saying that the census statistical races have no scientific origin gets lost in the noise.

I conclude this article by returning to its title. The census undercount never was a nail, but it was treated *as if* it were. That’s the problem when a hammer is the tool at hand. Real damage was done, though at a scale that can be shrugged off.

If the hammer-nail analogy is even slightly applicable in the case of genomic medicine, the damage will not be easily shrugged off.²⁷

Corresponding author: Kenneth Prewitt, Carnegie Professor of Public Affairs, School of International and Public Affairs, Columbia University, 1314a International Affairs Building, New York, New York, 10027. E-mail: kp2058@columbia.edu

NOTES

1. This article is based on a chapter in Kenneth Prewitt's *What Is Your Race? The Census and the Flawed Effort to Classify Americans*, scheduled for publication by Princeton University Press in May, 2013. The term "statistical races" is defined in more detail in the book, but essentially means the races resulting from government-adopted racial categories for use in the census, related statistical programs, and administrative records. It is these races that find their way into public policies, whether or not they match lived races, socially constructed races, identity races, biological races, or any other race categories established by social practices and attitudes.
2. The discussion of the census undercount draws heavily from and in places directly quotes a more extended treatment in Kenneth Prewitt's 2003 *Politics and Science in Census Taking*.
3. The efforts in the American 2000 census are documented in Hillygus et al. (2006).
4. Censuses can also overcount if duplicate forms are submitted; historically this was assumed to be a much smaller number than the undercount. As census practice has improved, the Census Bureau has found ways to find duplicates in the census record and to lower the overcount.
5. The actual calculation is more complicated. It has to estimate the completeness of birth and death records by age, gender, and race. Allowance also has to be made for imprecise estimates of immigration, especially the probable number of undocumented residents, and the estimates of out-migration, on which records are incomplete.
6. Except that no state, however small its share of the total national population, can have fewer than one seat in the House, the result being that the least populated state, Wyoming, gets a seat even though its population in 2000 was approximately 150,000 fewer people than that of congressional districts elsewhere in the country.
7. In the 2000 reapportionment, the final congressional seat went to North Carolina. The other state in contention for the 435th seat was Utah, whose population was only 857 persons fewer than North Carolina. It sued to have its Mormon missionaries temporarily stationed overseas included in the census count. Had it won that case, the 435th seat would have gone to Utah instead of North Carolina.
8. For an excellent nontechnical overview of dual system estimation, see Wight and Hogan (1999). For a more technical treatment, see Note 11 citations.
9. Submitting results of a census to political office-holders for preapproval was unprecedented. With even less justification this was repeated after the 2000 census. A statistical decision became a political decision, the most serious violation of the basic principle of census independence since the professionalization of census taking. For an account, see Kenneth Prewitt's (2010) article, "What is Political Interference in Federal Statistics?"
10. A high-level panel of the Committee on National Statistics in the National Academies generally supported dual system estimation. For its final report, see Citron et al. (2004). A comprehensive summary of the major criticisms of the method is Brown et al. (1999).
11. From late 1999 through January 20, 2001, I was Director of the Census Bureau and called to testify before a congressional committee more than two dozen times—an indication of the intensity of congressional attention to the design and conduct of the 2000 census.
12. This assertion of how many seats would be "lost" was never documented. Given the numbers involved, it is somewhere between highly implausible and statistically impossible for census adjustment to move this many seats, and certainly impossible a priori to calculate partisan shifts in legislatures resulting from a decennial census.
13. The census short form questions are asked of every household in the country. What in 2000 (and in censuses of the second half of the twentieth century, which is the period on which this discussion focuses) was called the long form, went to approximately one-sixth

of households. In addition to the standard short-form questions, the long form asked about education, health, income, marital status, housing, and other topics. Data from the long form is relevant to analyzing census coverage, but could not be used in dual-system estimation because estimates from long-form data are not statistically reliable at the block level. In the decennial of 2010, the long-form questions appear on the American Community Survey, and the decennial census itself was just the short form delivered to every household.

14. A thorough study of census cooperation in 2000 found that marital status and whether the respondent had family members living in the community—an indicator of strong community ties—did predict cooperation after controlling for race (Hillygus et al. 2006).
15. “Person characteristics” provided by U.S. Bureau of the Census (2010) were: language isolation, unemployment, mobility, below poverty level, receiving public assistance, and no high school diploma. “Housing characteristics” were: crowded housing, multiunit buildings, lack of telephone in home, vacancy rate, renter occupied, and complex households.
16. If the census included questions on the hard-to-count characteristics but not race, dual-system estimation would still be possible. A case can be made that it would be statistically superior to the heavy reliance on the race variable. The task here, however, is not to redesign the census, let alone take up issues of dual-system estimation that are statistically much more complicated than the treatment I have offered.
17. Of course, it was Blumenbach’s influential predecessor, Carolus Linnaeus, who in 1738 published his pamphlet *Systema Naturae*. The 12th edition (1776) includes a four-category typology of human races—African, American, Asian, and European—and described their respective characteristics, for example:

Africanus: black, phlegmatic, relaxed; hair black, frizzled; skin silky; nose flat; lips tumid; women without shame, they lactate profusely; crafty, indolent, negligent; anoints himself with grease; governed by caprice; *Europeaeus*: white, sanguine, muscular; hair long, flowing; eyes blue; gentle, acute, inventive; covers himself with close vestments; governed by laws.

Blumenbach separated the Brown race of Pacific Islanders from Linnaeus’s single category of Asians, resulting in the five centennial race groups in frequent use today.

18. Only five of the 3,636 subjects had DNA that matched a group different from the race box they had checked at the beginning of the study.
19. The costs and a time frame for bringing personalized medicine to the nearly seven billion people of the world have, in the material I have reviewed, not even been estimated.
20. I have no information on how widely the perception of risks put forward in the report are shared across the scientific community in the United Kingdom.
21. The “mark-one-or-more option” is described in detail in my book *What is Your Race?*.
22. Bennett was an analyst in the Racial Statistics Branch of the Population Division, U.S. Census Bureau.
23. For reservations about the promises of genetic medicine, see Evans et al. (2011).
24. Fujimura and Rajagopalan (2011) conclude, “Although we have identified new actors—both human and technical—working to avoid the use of race categories in biomedical genetics research, it is not clear how much they can change . . . institutionalized and historical practices” (p. 22).
25. As noted above, the drug, under the trade name BiDil, was approved the Food and Drug Administration in June, 2005. It was the first race-based prescription in the United States. It was not a market success, and is not currently being sold.
26. Attaching his quotation to the “winner of science fair” is not in the text, but is consistent with his example.
27. I take this concern into account when, in my book *What Is Your Race?* I propose a major change in how the U.S. government should measure race and ethnicity.

REFERENCES

- American Anthropological Association (2010). *Race: Are We So Different?* (<http://www.understandingrace.org/home.html/>) (accessed August 15, 2012).
- Anderson, Margo J. and Stephen E. Fienberg (1999). *Who Counts? The Politics of Census-Taking in Contemporary America*. New York: The Russell Sage Foundation.

- Barkan, Elazar (1992). *The Retreat of Scientific Racism*. Cambridge, UK: Cambridge University Press.
- Bennett, Claudette (2003). Exploring the Consistency of Race Reporting in Census 2000 and the Census Quality Survey. Paper Presented at the Joint Meetings of the American Statistical Association, San Francisco, CA, August 3–7.
- Black, Edwin (2003). *War Against The Weak: Eugenics and America's Campaign to Create a Master Race*. New York: Four Walls Eight Windows.
- Bliss, Catherine (2009). Genome Sampling and the Biopolitics of Race. Unpublished Manuscript, Department of Sociology, Brown University.
- Brown, Lawrence D., Morris L. Eaton, David A. Freedman, Stephen P. Klein, Richard A. Olshen, Kenneth W. Wachter, Martin T. Wells, and Donald Ylvisaker (1999). Statistical Controversies in Census 2000. *Jurimetrics*, 39: 347–375.
- Bullard, James H. and Sandrine Dudoit (2008). R/Bioconductor: A Short Course. (<http://wiki.biostat.berkeley.edu/~bullard/courses/T-mexico-08/lectures/hapmap/slides-2x2.pdf>) (accessed August 15, 2011).
- Charney, Evan and William English (2012). Candidate Genes and Political Behavior. *American Political Science Review*, 106(1): 1–34.
- Citro, Constance F., Daniel L. Cork, and Janet L. Norwood (Eds.) (2004). *The 2000 Census: Counting Under Adversity*. Washington, DC: The National Academies Press.
- Collins, Francis (2004). What We Do and Don't Know about "Race," "Ethnicity," Genetics and Health at the Dawn of the Genome Era. *Nature Genetics*, 36: S13–S15.
- Darga, Kenneth (1999). *Sampling and the Census: A Case Against the Proposed Adjustments for Undercount*. Washington, DC: The American Enterprise Institute.
- Duster, Troy (2005). Race and Reification in Science. *Science*, 307(5712): 1050–1051.
- Dyson, Freeman (2007). Our Biotech Future. *The New York Review of Books*, July 19, 4.
- Evans, James P., Eric M. Meslin, Theresa M. Marteau, and Timothy Caulfield (2011). Deflating the Genomic Bubble. *Science*, 331(6019): 861–862.
- Food and Drug Administration (1998). Investigational New Drug Applications and New Drug Applications. (<http://www.fda.gov/oashi/patrep/demo.html>) (accessed August 15, 2011).
- Food and Drug Administration (2005). Guidance for Industry: Collection of Race and Ethnicity Data in Clinical Trials. (<http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm126396.pdf>) (accessed August 10, 2011).
- Fujimura, Joan H. and Ramya Rajagopalan (2011). Different Differences: The Use of "Genetic Ancestry" Versus Race in Biomedical Human Genetic Research. *Social Studies of Science*, 41(1): 5–30.
- Hammonds, Evelyn M. (2006). Straw Men and Their Followers: The Return of Biological Race. Social Science Research Council Web Forum, *Is Race "Real?"* (<http://raceandgenomics.ssrc.org/Hammonds/>) (accessed September 3, 2012).
- Heer, David (Ed.) (1968). *Social Statistics and the City*. Cambridge, MA: Joint Center for Urban Studies.
- Herrnstein, Richard J. and Charles Murray (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. New York: The Free Press.
- Hillygus, D. Sunshine, Norman Nie, Kenneth Prewitt, and Heili Pals (2006). *The Hard Count: The Political and Social Challenges of Census Mobilization*. New York: The Russell Sage Foundation.
- Holden, Constance (2003). Race and Medicine. *Science*, 302(5645): 594–596.
- Kahn, Jonathan (2007). Race in a Bottle. *Scientific American*, 297(2): 40–45.
- Katznelson, Ira (2005). *When Affirmative Action Was White*. New York: W.W. Norton & Co.
- Leadership Conference on Civil and Human Rights (2000). *The Census 2000 Education Kit/ Census 2000 Everyone Counts!* Washington, DC: Leadership Conference Education Fund.
- Lewontin, R. C. (2006). Confusions about Human Races. Social Science Research Council Web Forum, *Is Race "Real?"* June 7. (<http://raceandgenomics.ssrc.org/Lewontin>) (accessed August 28, 2012).
- Linnaeus, Carolus ([1735] 1766). *Systema Naturae*. 12th edition.
- Martin, Paul, Richard Ashcroft, George T. H. Ellison, Andrew David Smart, and Richard Tutton (2007). Reviving "Racial Medicine"? The Use of Race/Ethnicity in Genetics and Biomedical Research, and the Implications for Science And Healthcare. London: Faculty of Health and Social Care.
- Morning, Ann (2011). *The Nature of Race*. Berkeley, CA: University of California Press.
- Murray, Charles (2007). Jewish Genius. *Commentary*, 123(4): 29–35.

- Nicholson, Jim (1997). Memo from the Chairman of the Republican National Committee, May 20.
- National Institutes of Health (1994). NIH Guidelines on the Inclusion of Women and Minorities as Subjects in Clinical Research. *NIH Guide 23.11*. Washington, DC: NIH.
- Oppenheimer, Stephen and the Bradshaw Foundation (2003). *Journey of Mankind: The Peopling of the World*. (<http://www.bradshawfoundation.com/journey>) (accessed August 15, 2012).
- Orr, H. Allen (2006). Talking Genes. *New York Review of Books*, 53(14): 20–26.
- Peters, Mary E. (2003). Determination of Reasonable Rates and Terms for the Digital Performance of Sound Recordings by Preexisting Subscription Services. *Federal Register*, 68(117): 36,469–36,470.
- Prewitt, Kenneth (2003). *Politics and Science in Census Taking*. New York and Washington, DC: Russell Sage Foundation and Population Reference Bureau.
- Prewitt, Kenneth (2010). What Is Political Interference in Federal Statistics? *The ANNALS of the American Academy of Political and Social Science*, 631(1): 225–238.
- Prewitt, Kenneth (Forthcoming). *What's Your Race? The Census and the Flawed Effort to Classify Americans*. Princeton, NJ: Princeton University Press.
- Risch, Neil (2005). Genetic Structure, Self-Identified Race/Ethnicity, and Confounding in Case-Control Association Studies. *American Journal of Human Genetics*, 76(2): 268–275.
- Risch, Neil, Esteban Burchard, Elad Ziv, and Hua Tang (2002). Categorization of Humans in Biomedical Research: Genes, Race and Disease. *Genome Biology*, July 1. (<http://genomebiology.com/2002/3/7/comment/2007>) (accessed August 26, 2012).
- Roberts, Leslie (1992). How to Sample the World's Genetic Diversity. *Science*, 257(5074): 1204–1205.
- U.S. Bureau of the Census (2010). (<http://2010.census.gov/partners/research/>) (accessed August 12, 2011).
- Varmus, Harold (2009). *The Art and Politics of Science*. New York: Norton.
- Wade, Nicholas (2002). Race is Seen as Real Guide to Track Roots of Disease. *New York Times*, July 30. (<http://www.nytimes.com/2002/07/30/science/race-is-seen-as-real-guide-to-track-roots-of-disease.html>) (accessed August 1, 2011).
- Wade, Nicholas (2006). *Before the Dawn: Recovering the Lost History of Our Ancestors*. New York: Penguin.
- Wallman, K. K., S. Evinger, and S. Schechter (2000). Measuring Our Nation's Diversity: Developing a Common Language for Data on Race/Ethnicity. *The American Journal of Public Health*, 90(11): 1704–1708.
- Washington, George ([1889] 1939). *The Writings of George Washington* (Vol. 31). Washington, DC: Government Printing Office.
- Wight, Tommy and Howard Hogan (1999). Census 2000: Evolution of the Revised Plan. *CHANCE: A Magazine of the American Statistical Association*, 12(4): 11–19.