International Journal of Communication 2 (2008), 1168-1194 1932-8036/20081168

The Informationalization of Race:

Communication Technologies and the Human Genome in the Digital Age

PETER A. CHOW-WHITE
Simon Fraser University

This paper suggests that a new form of racialization is being produced in the information age through developments and innovations in communication technologies. Increasingly, racial knowledge is being constructed from seemingly neutral and unrelated pieces of information, which are collected, sorted, and analyzed through two key technologies: databases and the Internet. I call this interaction between technology and identity the "informationalization of race." As a mode of representation, a structuring device, and as a biological category, race is undergoing a significant transformation in the digital age. I ground this concept in a case study of the next Human Genome Project — the HapMap Project — to understand how technologies are being shaped in a specific institutional setting. Advances in human genomics have recently re-invigorated scientific research into the relationship between race and biology. Where the HGP concluded that humanity is similar at the genetic level, the HapMap Project began by looking for differences between white, African, and Asian groups. It's anticipated that promising findings from the HapMap project will be of help in developing pharmaceuticals that can target common diseases, such as cancer. However, this development also opens the door to old biological conceptions of race and a new phase of the biopolitics of the human body.

During the past three decades, social scientists have been discussing and researching the perils and possibilities of the information age. Many have highlighted the positive impacts of enhanced information and communication technologies as well as their potential for democracy. Utopian visions of the Internet in the 1990s anticipated that equality of access, coupled with a cyberspace where race and gender are rendered irrelevant, would transform social relations. A number of survey studies in the late 1990s and early 2000s, such as the PEW Internet & American Life Project, began to shift from outright praise to measured description by collecting systematic data on who was using the Internet and for what purposes. Then came the end of the dot-com bubble and 9/11. The view of technology as a panacea for

---

Peter A. Chow-White: **petercw@sfu.ca**
Date submitted: 2007-10-05

societal ills was muted and racism more pervasive. The Internet had not turned into the killer application for social equality many had envisioned. Scholars and NGOs published more nuanced and theoretically informed studies about the shaping and meaning of cyberspace. For example, Mack (2001) showed how the uneven diffusion of information and communication technologies created a racial digital divide in the United States. Others examined racial representation in cyberspace and found that race mattered as much online as in real life (Kolko et al., 2000; Lee & Wong, 2003). While the Internet allowed for fluid performances of gender and race, usually in the form of identity tourism where users could "act" a different gendered or racialized identity, participation often occurred through stereotypes (Nakamura, 2002). Analogue racial codes were being "cybertyped" into digital code, resulting in new and old identity formations. However, what remained in question is whether innovations in information and communication technologies would be in the employment of greater social inclusion or of more social control (Gray, 2005; Lyon, 2003).

The transcoding of race into cyberculture was not always in the service of dominant discourses. Contrasting the "use" surveys, a collection of scholars, in one of the first anthologies on race and technology in the Internet age, observed a disjuncture between, on the one hand, utopian government and commercial representations of the information society that promised an integrated, universally accessible technological system and, on the other, the conditions of work, cultural production, and people of color's everyday experiences using and appropriating new technologies (Nelson et al., 2003). The Web and related information technologies were also becoming new media spaces and significant points of participation where people from minority groups could challenge dominant ideologies of race and ethnicity (Leung, 2005). This could be accomplished directly by deploying counterhegemonic images or indirectly through community building at the local or transnational levels (Ignacio, 2005; Landzelius, 2006; Nakamura, 2008). After a century of struggle over inclusion in the largely one-way media of radio, film, and television, the Internet and information technologies seemed to hold new possibilities for participation, appropriation, and representation.

As the Web moved from text-based 1.0 to Web 2.0, it has become multimedia, image-based, participatory, and commercially driven. Technology's positive implications with regard to racism and societal ills are being touted again, and U.S. election pundits envisage a new, "post-racial" America. Within popular and commercial discourses of globalization and neoliberalism, there is a sense that race and racism are diminishing in the information age. However, while conceptions of race are changing due to transformations in culture, technology, and science, these processes are not necessarily countering racial hierarchies that have been developing for the past 400 years. Research into race and digital culture has shown how new technologies reinforce old forms of racial identity in new ways that are less obvious than in previous media forms (Nakamura, 2008). The conditions for identity building have shifted as new social, political, and economic formations emerge in and through new media- and techno-scapes, shaped most importantly by the Internet. This article examines how the convergence of changing concepts of race with information technologies has produced a new paradigm for race — what I refer to as the informationalization of race. This paradigm, while it emerges in discourses of the information age that present themselves as color-blind, is, I argue, in fact, a product of new regimes of racial knowledge that are oriented around the digital cultures of communication technologies, especially the Internet and databases, both in everyday practices and social institutions.[1]

Human genomics is an important area to investigate how race is being recoded via culture and technology. This burgeoning field has been referred to as an information science, since communication technologies, such as computers, the Internet, and digital databases, are essential for sequencing DNA. Without them, genome science would be virtually impossible. It is the meeting ground for genetic, computer, and racial codes and is both challenging preconceived notions of racial meaning and reproducing racial classification. For example, where Craig Venter and Francis Collins — the leaders of the private Celera Geonomics project and the public Human Genome Project (HGP) — concluded that we are all 99.9% the same at the genetic level; whereas the next HGP, the International HapMap Project, is mapping differences between African, Asian, and European groups. Drawing on interviews with members of the HapMap Project (see Appendix) and textual data from biomedical and scientific journals, this study examines technological innovation and cultural and scientific discourses in genomics. Utopian visions of this emerging biotechnology suggested that scientists would be able to see into the human body in new ways and begin to map and manipulate the very building blocks of life to improve health. However, genome research into group differences is also a crucial site of negotiation in the relationship between race and technology.

The Informationalization of Race

Critical scholars have delineated the process of racialization into two interdependent ontological paradigms: race as biology and race as culture. In the race as biology paradigm, racial groups are identified according to a collection of physical phenotypes, such as skin color, hair texture, and shape of eyes and nose, and tied to a hierarchical order of racial groups, with whites at the top and ranging to blacks at the bottom (Banton, 1998; Jordan, 1974; Miles, 1989). The formative process of racialization has been linked to the rise of enlightenment thought (Eze, 1997), the spread of European colonialism (Césaire, 1972; Fanon, 1963, 1967; Memmi, 1965), slavery (Jordan, 1974), and scientific racism (Barkan, 1992; Gould 1996; Mosse, 1978; Stepan, 1982). In the post-civil rights era, biological classification shifted to cultural codes, a development that has been referred to as the culturalization of race (Razack, 1998) or the new racism (Barker 1982; Collins 2004). Group characteristics continue to be ascribed based on a symbolics of the body, yet the real differences between groups are no longer biological, but rather ethnic or cultural. Cultural notions of race are deployed in a similar manner to biological ones as homogenous, dislocated from historical context, static, and containing an implicit association to group position in the social order: "a fixed property of social groups, [rather than] something intrinsically fluid, changing, unstable, and dynamic" (Gilroy, 2000, p. 266). Racial signification is rearticulated in cultural terms where everyday talk and political statements about the cultural characteristics of a particular group

---

[1] When I refer to the term "communication technologies," I am following Castells and others to include the "*converging set* of technologies in micro-electronics, computing (machines and software), telecommunications/broadcasting, and opto-electronics" as well as Internet applications and DNA technologies (Castells, 2000, p. 29).

often are, fundamentally, thinly disguised racial claims. Kim argues that the move to culturally coded racial discourse has, in fact, stabilized white privilege in the post-civil rights era.

> It is precisely because it has been revamped in nonracial language that the field of racial positions functions so effectively to reinforce white privilege today. Representing a cultural explanation for group inequalities, the field of racial positions implies that American society is substantially color-blind and that the American Dream is still viable. (Kim, 1999, p. 117)

Scholars refer to the dominant racial ideology that is characterized by coded discourse about race as color-blind racism (Bonilla-Silva, 2003) and color-blindness (Brown et al., 2003; Wellman, 2003). Bonilla-Silva (2001) suggests that the civil rights movement marked a change in the racialized social system from Jim Crow racism to color-blind racism. In the former phase, blacks and other minorities were considered inferior to whites because of their biological and moral inferiority. One of the most significant changes in the post-civil right era is in the use of nonracial and cultural terms to articulate racial differences in the public sphere. Bonilla-Silva (2001) describes race talk during the Jim Crow era as "direct and blunt" (p. 68). Racial claims now tend to be made indirectly. Hockey, basketball, special interests, inner city, thug, welfare, and sickle cell disease are racially coded in a color-blind manner. Instead of the terms of inclusion and exclusion — superiority and inferiority being based in genetics — a group's success in society is reliant on their cultural competency. Many even speak of a post-racial society. In spite of the continued existence of racial inequality in terms of social, economic, and political power — what Lipsitz (1998) refers to as the *possessive investment of whiteness* — the turn to color-blind racism attributes differential standings of minorities to market forces, naturally occurring phenomena (i.e., segregation is natural because people want to live with others who are like them, a phenomenon that Taguieff (1990) refers to as the *biologization of culture*), and cultural variations between groups (Bonilla-Silva, 2003, p. 2).

A new regime of racial signification, characterized by the creation and deployment of information, is emerging through the digital space of communication networks, computer codes, computational algorithms, and databases. New communication technologies increasingly make up the central media systems in which racial meanings are created, transformed, and destroyed, to borrow from Omi and Winant (1994). This networked system is not simply a delivery tool for ideas and meanings; its very structure and scope, both hypertext and globally linked, is productive of new mechanisms of racialization. Where conventional conceptions of race have been articulated in terms of culture or phenotype, in the digital age, information is the material by which we work on racial meaning.. However, race as information does not replace the dependency of racialization on ethnicity or skin color. Rather, as the paradigm of race as culture emerged from the paradigm of race as biology, I would argue that the paradigm of race as information has emerged from both to create a new racial formation — the informationalization of race.

What distinguishes race as information from other modes of racialization is the transformation of society due to globalization, the new economy, and communication technologies. Racial identity, meanings, and structures are being created through the use and shaping of new media and communication technologies. This is significantly different from previous forms of the social construction of race as the body becomes "posthuman" (Hayles, 1999). That is, the "meat" has been left behind and cultural signification has become embedded in computer programs and complex algorithms, hidden from the front end, user interfaces. In an early example of this process, Hammonds (1997) describes a computer program that could morph people into different races and create entirely new people from an amalgam of a number of facial features of different racialized groups. Rather than deconstructing biological notions of race, Hammond argues, they reinforce centuries-old stereotypes of racial difference and cultural anxieties of miscegenation. The computational routines that produce racial representations are "new technologies of race," translating ideologies of bio-race into the seeming neutrality of digital space. As cultural and phenotypic signification meet computer code, the analog systems of racial signification enter the digital world, producing new modalities as well as reproducing old ones. Traditional media analyses of the mechanisms of racialization tend to focus on representation in television, film, and news. The concept of the informationalization of race directs attention to the algorithms, data, and discourses produced in the hypertext communication environments of computers, databases, data mining, and the Internet.

Unlike the more visible forms of new media technologies (mobile phone, laptop computer, and the iPod) or 'invisible' delivery systems (Wi-Fi or the Web), the database is a central innovation in the information economy (Elmer, 2004; Cubitt, 2000; Loro, 1995; Manovich, 1999; Poster, 1991). Similar to old media forms, such as the novel, film, and television, the database is a new media technology that structures our personal and institutional experiences both symbolically and materially. Differing from largely one-way mediums, databases enable a "networked multilogue" (Loro, 1995, p. 55) between producers and consumers through the process of sorting and storing data, networking information, and constructing knowledge. Data mining (DM) is a technique for searching and creating knowledge out of digital databases. Derived from the computer sciences, data mining is a step in discovering knowledge in data. Within specified parameters, computer programs search databases, using mathematical algorithms to find significant patterns. Data mining techniques are made up of sophisticated algorithms, neural networks, and artificial intelligence. They can work from pre-determined sets of categorical variables or they can go beyond what a user knows to request and "discover" unseen patterns, facts, and relationships between the data (Danna & Gandy, 2002; Zarsky, 2003).[2] As the Internet has been employed by government agencies and marketing firms to gather information on citizens and customers, the goals for data mining have increasingly moved from description to prediction (Gandy, 2006). Companies across all business sectors, from grocery shopping to Amazon.com to genetic services, employ databases and DM to find out characteristics of their clientele. In order to provide personalized services, firms collect demographic and behavioral data on individuals and groups to understand who they are, what they like, and their prospective future buying potential. In the case of genomic services, companies provide information on genetic potential for diseases as well as ancestry and racial identity. While databases and the Internet have opened up new possibilities for consumer participation, alternative media production, and the expression of individual and group identities, many have been warning about the increased surveillance capabilities of new technologies.

---

[2] The technique of data mining was initially developed in the late 1980s by Usama Fayyad, a graduate student at the University of Michigan, for sorting through the increasing mess of digital and analog records gathered by GM Motors. According to Fayyad, "there were hundreds of millions of records, no human being could go through it all" (Waldrop, 2001, online; see also Fayyad, 1991).

One of the central purposes of information infrastructures is for surveillance (Lyon, 2002, 2003). Surveillance can include government records of citizens, DNA databanks, customer information collected and shared by companies, and employers monitoring employees. Different technologies are utilized to make up the information infrastructures such as CCTV, biometrics, and DNA. Foucault (1977) has shown how surveillance was a central technique in the management of populations. Surveillance systems that depend on digitized information infrastructures can operate efficiently from afar and thus exploit the advantage of remaining largely invisible. While advocates for data mining and Internet surveillance argue that these technologies are good for the consumer for fighting crime and 'terror' as well as personalized health, others question these claims and suggest that, in fact, that these new technologies of social sorting may be inherently discriminatory or, at best, are deployed in social systems that are structured in dominance (Haggerty & Ericson, 2006; Lyon, 2003).

The term "informationalization" builds on the insights of the literature on surveillance and the information society that examines the shift in societies from industrial economies to services economies due to the restructuring of capitalism, the new technological paradigm, and globalization (Webster, 2002). Informationalism is a "specific form of social organization in which information generation, processing, and transmission become the fundamental resources of productivity and power" (Castells, 2000, p. 21). Traditionally, the concept of information denotes a neutral set of facts, data, or observations. When information is networked and takes on the form of an information infrastructure, such as when databases are compiled and linked or the seemingly endless pathways of the World Wide Web, it should not be treated as having a simply reflective role in the social world. Like other media forms, (news, television, film) information infrastructures do not just support cultural, political, social, and scientific processes. They play a constitutive role.

> It is politically and ethically crucial to recognize the vital role of infrastructures in the 'built moral environment.' Seemingly purely technical issues like how to name things and how to store data in fact constitute much of what we have come to know as natural. (Bowker & Star, 1999, p. 326)

Technologies and the classification systems that utilize them tend to make invisible the myriad of decisions that create them. We, the users, see only the interface: the front end of a particular technology. Hidden away inside are the attitudes, values, and politics that are written into code and "the arguments, decisions, uncertainties and processual nature of decision-making" (Bowker & Star, 1999, p. 187; see also Garfinkel, 2000). Put another way, the seemingly descriptive representations derived from information infrastructures in fact naturalize a whole set of practices, procedures, and ideological premises.

> Looked at historically, information seems basic to social life . . . In modern, literate cultures, artificial signs proliferate, and are frequently associated with social order itself. Signs tell us of distant events, places, persons and processes. Information is relational, connecting by reference persons and things . . . But whereas information might once have thrown light on reality, or even, through instructions or recipes, contributed to the

> transformation of reality, once technological devices become the predominant carriers of information, the distinctions blur. (Lyon, 2005, p. 225)

Information has moved from being factual, to technical, to a commodity and basis for social, technical, political, and cultural organization. Castells argues that it is not only the centrality of knowledge and information that characterizes the current technological revolution, "but the application of such knowledge and information to knowledge generation and information processing/communication devices, in a cumulative feedback loop between innovation and the uses of innovation" (Castells, 2000, p. 31). The informationalization of race acknowledges that race as a structuring device in society has not diminished in importance with the information age and has continuity with modernity. Race as information brings into the fore new forms of racialization that previous concepts, such as race as the body, race as nation, and race as culture do not. When cultural codes meet computer code, programming and distributed networks have become the new sites of negotiation over norms and the terms of social inclusion and exclusion. As Galloway and Thacker suggest about the logic of the network society, "codification, not reification is the new concern" (2007, p. 134). Where racialization has been about the reduction of cultural practices and social relations to objects, the informationalization of race looks to where bodies and practices become code. This could be in the form of state surveillance, bioprospecting, or self-expression online such as Facebook.

At all levels of society, from institutions to individual identities, information has become the material in which social and political meaning is constructed, new companies profit from, and states utilize to govern. This is a crucial development for the study of racial formations. This research seeks to examine how innovations and applications of communication technologies and the rise of information have produced new mechanisms of racialization in a post-civil rights context. Scholars have pointed out that race has not decreased in significance, but persisted and transformed into color-blind racism. As communication technologies play an increasingly centralized role in the everyday practices and organization of a range of social institutions and industries, there are a number of sites where we can see the informationalization of race at work, such as law enforcement, biomedical research, insurance, and marketing. While each would have their own set of technologies for information storage, classification, and surveillance, they have increasingly employed a similar array of technologies to their own institutional needs and goals. Where the microscope has been a central observational tool in the biomedical sciences, actuarial tables in insurance, and fingerprinting in law enforcement, data mining techniques and the technological infrastructure that it requires are commonly used across these different sectors. Their methods of observation, classification, and knowledge production have incorporated networked, digital, and informational processes. This research explores developments in biomedical research into DNA. In molecular biology, Mackenzie questions the usefulness and point of tracking computational processing of sequence data, which heavily utilize new media such as the Internet and databases for storage, analysis, and distribution: "Does not bioinformatics merely support the more decisive intellectual, social, political, cultural and economics events associated with contemporary biology and genetics" (2003, p. 316)? While scientists debate the accuracy or inaccuracy of scientific data, which is the outcome of computational and statistical routines, it is important to understand how those outcomes stitch together cultural assumptions, molecular particles, microprocessed bits and bytes, and historical context. In what follows, I explore how genomics is an effect of the informationalization of race.

Biology as an Information Science and the Human Genome

There has been a concurrent ascendance of the information society and genetic technologies (Capra, 2002; Castells, 2000; Meyers & Davis, 2003). Computer science and genetics have converged to the point where biologists, computer scientists, and engineers work side-by-side, borrowing from one another both methodologically and theoretically (Marturano & Chadwick, 2004). Modern science relies to a very large extent on computer simulations, computational models, and analyses of large data sets where genetic data points can number in the billions, such as in the case of the Human Genome Project (Gezelter, 1999). Genetic technologies are information technologies, since they are focused on the decoding and eventually reprogramming of DNA, the information code of living matter (Moody, 2004; Thacker, 2004). And, more importantly, without "massive computing power and the simulation capacity provided by advanced software, the HGP, would not have been completed — nor would scientists be able to identify specific functions and the locations of specific genes'' (Ibid.). In addition to this, some philosophical problems arise from the view that DNA and the Human Genome are pure informational concepts. In one sense, the convergence between the biological and computing might be thought to be associated with the massive use of computer technologies in biology. Computers are convenient tools for genome and protein sequencing. However, the bioinformatics paradigm has been central to the reorganization of disciplines such as molecular biology (Holdsworth, 1999). In the process, they have absorbed Shannon's technical notion of information.

Genome science has emerged as the next wave of human scientific research. There are a number of important ways that genomics differs from genetics. Genetics is the study of genes and the inherited differences and variation in DNA and its influence on phenotypes (Interview 1001; Interview 1022). The Department of Energy's Human Genome Project website defines genomics as "the study of genes and their function,"[3] as well as their interactions with environmental factors (Interview 1022). Genome Canada calls genomics "big picture science" whose central aim is to understand the "complexity of how genes interact with each other and their environment to make living organisms function."[4] Genomics is largely driven by the impetus to understand the differences in genomes, such as the variations in sequences at the same locations in a DNA strand, between different individuals and population groups, and the origins of complex diseases such as cancer. Where genetics looks at individual genes, genomics focuses on the entire collection of the 6 billion chemical bases, the A, C, G, and Ts, that are strung together in our 23 pairs of chromosomes. These more technical definitions based in biology miss some important technological, social, and cultural characteristics of genomics. One of the main challenges of genomics is the enormous amount of data involved in studying the 3 billion letters of code in a strand of human DNA, and the sheer number of relevant variables that has necessitated an inherent and enabling link to

---

[3] http://www.ornl.gov/sci/techresources/Human_Genome/glossary/glossary_g.shtml

[4] http://www.genomecanada.ca/en/info/DNA/genomics.aspx

computing. Developing from the convergence of molecular biology and computing science, genomics can also be defined as the computer assisted comprehensive study of all genes[5] (Interview 1001).

Like many new technologies, boosterism about its scientific potential leads into possible social benefits. Francis Collins, the former Director of the National Human Genome Research Institute (NHGRI; one of the National Institutes of Health) states about the promises of genomics for society:

> Genomics has been at the forefront of giving serious attention…to the impact of science and technology on society. Although the major benefits to be realized from genomics are in the area of health…genomics can also contribute to other aspects of society. Just as the [Human Genome Project] and related developments have spawned new areas of research in basic biology and in health, they have also created opportunities for research on social issues, even to the extent of understanding more fully how we define ourselves and each other. (Collins, 2003, p. 483)

Evelyn Fox Keller has shown how the importance and prominence of "gene thinking" grew during the 20th century in scientific and popular discourse (2000). Genome thinking may be growing at an even faster rate. "Personalizing" a Google News page with the keywords "genomics or genome," results in daily lists of stories in the 'mass' and scientific media, indicating we have entered the "genome era" (Bonham et al., 2005, p. 9). While genetics emerged from the insights of Mendel and the scientific practice of observation by sight, which are "analog" practices, genomics is emblematic of a cultural shift in scientific practice characterized by digital observation, data mining analysis, and computer automation. If the 20th century was the century of the gene, then the 21st is shaping up to be the century of the genome.

The International HapMap Project

A key site to examine the intersection of science, bioinformatics, communication technologies, and racialized identity is the next HGP, the International HapMap Project. The project is a multi-site, international venture between scientific teams in Canada, China, Japan, Nigeria, the UK, and the U.S. The project is in the process of mapping haplotypes of human genomes. Haplotypes are identifiable block-like patterns of the DNA nucleotides, A, C, G, and T. A genome consists of individually sequenced strings of nucleotides. The aim of the HGP was to create a comprehensive map sequence of an individual's DNA. In 2001, scientists shifted the focus to haplotypes in the genome. Haplotypes are "neighborhoods" of DNA that can be identified by common genetic variants in the genome sequence, revealing an underlying structure in a genome. Identifying the location of the variants, called single nucleotide polymorphisms (or SNPs), can facilitate and speed up subsequent research into the genetic origins of disease (International

---

[5] One of the HapMap participants, an internationally renowned geneticist, pointed out that there is a lack of agreement on the difference between genetics and genomics. While the interviewee felt that he was very clear about the differences, he recalled in meetings how close colleagues disagreed with his definition.

HapMap Consortium, 2003). Since the project began in 2001, Phases I and II of the project sampled groups from Europe, Africa, and Asia (China and Japan). Where the organizers of the HGP declared humanity's genetic similarity, HapMap searches for differences. Like the HGP before, communication technologies are at the heart of mapping, sequencing, data gathering, storage, analysis, and distribution.

Information sharing, data transfer, and analysis operate on a global scale from the sample collection sites in Africa, Asian, and North America to the sequencing and research centers. As the project's network of labs are scattered across the globe, the only way for genome scientists to construct a haplotype map of "major geographical groups," as they are named in the project, is to disassemble DNA into digital packets, send them through the Internet, and reassemble them in a SNP database. Each team of the HapMap project sequences a portion of the genome, such as Chromosome 2 and Chromosome 4p by the McGill University/Genome Quebec group in Canada, Chromosome 7p by the group at Washington University in St. Louis and the University of San Francisco. Others include both academic labs and biotechnology companies in China, Japan, and the United Kingdom. The data from each group is collected, curated, and stored for distribution in a database in Bethesda, Maryland, at the National Institutes of Health and released online for public access within 24 hours. The scientists' most important tools for the collection, databasing, and manipulation of genome sequences are communication technologies.

A leading geneticist from the HapMap Project stressed the under-appreciated role that communication technologies play in the development of genomic research (Interview 1001). Developments in computer power, databases, and the Internet have made the archiving, management, and distribution of the vast amount of data possible not only in local labs and internal networks, but also on a global scale. Genome research produces large amounts of information that bench scientists could never analyze by hand and "would be simply impossible to process without computers" (Interview 1009). This includes statistical calculations, which have required the development of new algorithms and software. Traditionally, biologists are not quantitative and are unaccustomed to working with such large data sets.

> And there has been over the last decade increasing, a trickle first and then increasing in flux of computationally sophisticated people into the field who have brought with them sophisticated computational methods for data mining and data analysis. But there's still a big divide that exists between most working biologists and any of those methods. And so it requires not just an intellectual shift, but also a real cultural shift because biologists are used to . . . the limiting step being their ability to collect data with their hands. (Interview 1001)

With the interaction of computer science and molecular biology, genomics has moved biomedical research to computational biology (Interview 1006).

> I think that, in fact, in many circles genomic research has been discussed as developing out of DNA technology. I think it is fair to say that the entire concept of genomics, which is really one of data rich studies in biology where you have archival quality data that is

> comprehensive and is shared freely, is as much about, if not more about, computers and the Internet as it is about DNA technology. (Interview 1001)

Accordingly, a number of respondents agreed that without the rich, flexible, and networked databases, as well as increasing speeds of computing, the impact of genomics would be trivial compared to what it is today. Developments in genomics are as much feats of technological management, archiving, and distribution as are advances in scientific knowledge (Interview 1001). Genomics simply would not be feasible by hand, making it impossible to analyze a single genome or make comparisons between individuals, groups, or even species (Interview 1003; Interview 1011). HapMap participants commonly discussed two technological developments in particular: databases/data mining and the Internet. New databases have been designed to store, analyze, and distribute the data and findings. Data mining techniques and "large, easily-accessible databases that would allow the extraction and comparison of data was absolutely essential for being able to put together any kind of sequence database" (Interview 1014). The Internet enables genome projects to move data between global locations and labs in the same building as well as provide open access from anyone interested in the data. These large-scale information infrastructures are the back end of scientific information. They are referenced but not normally discussed in scientific journals and media stories about scientific discovery. Further, they have become new technologies of difference in the informationalization of race and crucial to the new information science. Where racial science of the 19th century used grapeshot and mustard seed to measure supposed racial differences in skull volumes, genome researchers utilize digital genetic information to understand DNA variation between what many refer to as population or geographic groups.

The Shaping of the Internet and Databases in the HapMap Project

The HapMap consortium collected samples from "populations with ancestry from parts of Africa, Asia, and Europe" (International HapMap Consortium, 2003, p. 789). While project organizers deliberately decided to refer to the sample groups in terms of populations and not racial groups, the initial groups do match a traditional American taxonomy of race. When the National Human Genome Research Institute (NHGRI) of the National Institutes of Health (NIH) decided to build databases and a haplotype map, the scientists involved had to make a decision on which markers would be included. They decided that the groups would be labeled according to geographical rather than racial signifiers. An attendee told me that the discussions at the preliminary HapMap meetings held in the summer of 2001 were very open about the issue of racial identification (Interview 1008). When asked about the place of race in the schedule of items being discussed and whether or not it was an important issue or a marginal one, a bioethicist in attendance commented that it was "in the fabric of the meeting." Overall, the interviewees from the project are quick to point out that they do not use race but rather geographical ancestry to define the population groups. However, it would be remiss to overlook the significance of choosing the three major racial groups — African, Asian, and European — that have largely defined the social construction of race. When differences in a selected haplotype sequence (patterns of SNPs or single nucleotide polymorphisms) between these three groups are reported from digital DNA information, this can easily suggest the conclusion, especially in the context of a racialized society, that the differences are racial, genetic, and biological. For example, Hinds et al. (2005) report on whole-genome patterns of DNA variation between these three "population groups" and examine millions of SNPs. While this number seems rather large, the total number of individuals the samples collected from was 71, which is a very low sample size to be able to generalize to a larger population and assess any between-group variations. Further, when scientists collect and name the samples and enter them into databases, an example of turning race into information at the lab bench, the distinctions they make between the social and scientific meanings of race are not clear (Fullwiley, 2007a, 2007b). Scientists themselves often have trouble articulating a clear definition of race while at the same time rigorously racially naming genetic data. As Fullwiley reports on ethnographic work in a genetic lab, when she asked scientists "How do you define race?" Many of the informants "drew a blank" at the question and paused before answering, supposedly due to both confusion and caution around a controversial issue (2007b, p. 225). However, this also shows how the informational process of tagging, where DNA is organized into racial groupings, naturalizes the classification process and removes it from the analytical steps and social meaning.

To collect data for the database and analyze it, a new scientific process called "discovery science" was developed and is largely credited with the success of, first, the HGP and now the HapMap Project. Discovery science utilizes knowledge discovery in databases to map the haplotype blocks. Much like the basic assumptions of data mining described above, discovery science is "the idea that you take an object and you define all its elements and you create a database of information quite independent of the more conventional hypothesis-driven view" (Hood, 2001, online). In contrast to the meticulous method of making theoretically sound hypothesis before collecting and analyzing data, discovery science is more of a collect-first-and-ask-questions-later approach. The version adapted by Craig Venter and the private Celera HGP was called the "shotgun" method. According to Leroy Hood (one of the scientists who attended the early HGP meetings in the 1980s and the leader of the team form the California Institute of Technology who invented the DNA sequencer), the genome posed such technological problems not only with sequencing and mapping, but also in computation and analysis that a new paradigm — discovery science — was needed to tackle the enormous obstacles posed by creating and analyzing a comprehensive database.

The Internet enables genome projects to move data globally between international locations and local labs in the same building as well as provide open access to anyone interested in the information. Most importantly, the Internet allows researchers to upload data from collection sites and sequencing centers to the centralized database at the NIH. The HapMap Project follows an open source approach to data by making the findings available publicly on the Internet within 24 hours of the data being collected from the multiple sites. Distributing the data is akin to sending around "seven copies of the New York phone book" on a regular basis. "The impact would be trivial compared to what it is."

> You know that you don't need humans burning CDs, putting them in little boxes and mailing them to people, paying for the postage. I mean the cost of distributing data in the absence of this is, it's a cost of capital in human overhead and that, you know, humans are expensive, and probably not ideally suited for menial tasks such as preparing data for distribution, that's something computers do very well, and it's cheap. (Interview 1016)

Aside from being expensive and time consuming, transporting information via CDs limits the open source nature of genome databases, which have to be flexible and dynamic. The shared master version is constantly updated, instead of the changes from individual teams or individual labs storing their changes locally, waiting for the next round of data transfer. The Internet has increased the ability of scientists to undertake international collaborations. Also, cumbersome literature searches and the painstaking process of finding other researchers working on similar problems have been ameliorated through cyberspace. The changing technological procedures and culture of biology take place in the context of the wider social transformations of the information society. A key shift in the information age is the rise of color-blindness.

Discourses of Racial Realism and Color-blindness in Genomics

When organizers of the HapMap Project met at the National Institutes of Health in the summer of 2001 to discuss technical, scientific, and bureaucratic issues for this global undertaking, the naming of the four sample groups would be one of their most significant challenges and contributions to future genome research. For half a century, social and natural scientists across the world who argue that there is no biological basis for race often cite the post-WWII UNESCO "Statement on Race" (UNESCO, 1950, 1969). So when the groups from Africa, Asia, and Utah were chosen, this decision entered into the context of new debates about the relationship between biology, health, and race. Since the early 1990s, *Science* (Duster, 2005; Marshall, 1998; Sankar & Cho, 2002), *Nature Genetics* (Editorial, 2000), the *Journal of the American Medical Association* (Kaplan and Bennett, 2003; Osborne & Feit, 1992), the *New England Journal of Medicine* (Schwartz, 2001; Winker, 2006), *Genome Biology* (Risch et al., 2002; Petsko, 2004), and *Genomics* (Knoppers et al., 1996) have all featured editorials, letters, research studies, and commentary on the subject of race and science. These discussions have not only taken place in the medical and scientific journals, but also in other areas of health and biomedical research such as public health (Fullilove, 1998; Oppenheimer, 2001), pharmacogenomics (Nebert & Menon, 2001), and epidemiology (Editorial, 2000). A number of journals have instituted editorial policies on the use of race and ethnicity in research (British Medical Journal, 1996; International Committee of Medial Journal Editors, 2006; Journal of the American Medical Association, 2005; Nature Genetics, 2000). Further, the U.S. Department of Energy sponsored a special issue of *Nature Genetics* in November of 2004 that brought a number of leading scientists, social scientists, and experts in bioethics to discuss "'Race' and the Human Genome" (Patrinos, 2004).[6]

---

[6] These ethics and policy debates continue. *Genome Biology* recently published the 'Stanford 10,' which is a set of principles for the use of race in genetics by a multidisciplinary group of scholars from the natural and social sciences (Lee et al., 2008). *Science* issued a multidisciplinary group statement on the ethics and practices of genetic ancestry testing (Bolnick et al., 2007). The American Society of Human Genetics also put out a statement on direct-to-consumer genetic tests, although there was no mention of race or ethnicity (ASHG, 2007). Additionally, Lundy Braun and her colleagues questioned the use of racial categories in medicine (Braun et al., 2007), which was echoed by Craig Venter and his colleagues as they argued for using individual genomes instead of race in personalized medicine (Ng et al., 2008).

In spite of the many editorials, letters, and special issues, mainstream DNA research has been conducted using racial categories as a proxy for population groups. Some have strongly argued that biologically distinct groups do exist. This group of scientists, in forensic science, molecular biology, and population genetics, argue that racial markers can be identified in the human genome (Devlin & Risch, 1992; Evett, Buckleton, Raymond & Roberts, 1993; Jobling & Gill, 2004). While a number of scientists and doctors can be considered racial realists, arguing that human groups are organized genetically into a number of fundamental categories or clusters, other scientists and critical scholars argue that race is a social, political, and scientific construction. New scientific technologies are "prosthetically extending sight into nano-scales and can be linked to the impact of digital processing and other allied approaches to the body that allow it to be seen and understood in new ways, principally as code and information . . . skin, bone, and blood are no longer the primary referents of racial discourse" (Gilroy, 2000, pp. 44, 48; see also Nelkin & Tancredi, 1989, p. 15). Race is being constructed in terms of genetic information, rather than the traditional markers of skin, culture, or nation. The new genetic research is bypassing the skin in its search for the truth of who we are at the molecular level while digital imaging is re-imagining the body as code and information, rather than flesh and blood; a key development in the informationalization of race.

> The boundaries of "race" have now moved across the threshold of the skin. They are cellular and molecular, not dermal. If "race" is to endure, it will be in a new form, estranged from the scales respectively associated with political anatomy and epidermalization. (Gilroy, 2000, p. 47)

Gilroy refers to the process of constructing race at the genetic level as the molecularization of race (see also Fullwiley, 2007a). While discoveries are being made to cure previously untreatable diseases, the interaction between culture, technology, and science in genomic research has opened up old questions about the biological validity of race, the role of race in science, and science's role in the construction of race. Instead of finding evidence that we are all indeed the same beneath the skin, there are areas of genomic research that are attempting to discover new (and not so new) differences based on old assumptions (Duster, 2003, p. 146). Both Gilroy (2000) and Duster (2003) point to the relationship between the knowledge of genetic research and the technologies that make the body, below the skin, able to be seen. As mentioned above, one of the key differences between HapMap and the Human Genome Project is that the HGP concluded that we are all 99.9% the same at the genomic level while the HapMap Project looks for and compares differences between groups. This dilemma is not easily reconcilable. At times, genomic researchers directly confront the history of racial science and adopt color-blind lingo either out of a sense of ethics, justice, or good science, or in response to criticism of the reifying effects of their work. Others, however, may simply be inattentive to the terms they use. Scientists across the political spectrum grapple with ideology and data, taking positions between racial realism and color-blindness. Equipped with the Internet, databases, distributed networks, and data mining technologies, the scientific and biomedical journals have become their battleground.

Race talk in new genetics that takes place in scientific and medical journals is polished and the coded language that characterizes color blindness is much more difficult to pick up in highly educated elites than in everyday stories. One must be attentive to general frameworks, slippages, and contradictions in the formal narratives of the journal discourse and the less formal interviews. A major theme in the scientific literature is the discursive move to color-blind language in human genomics. Science has followed the social norms and the ideology of color-blindness in the post-civil rights era and there has been an effort by many researchers to replace racial categories with genetic terms that refer to ancestry. Racial classification systems that used terms such as Negroid, Mongoloid and Caucasian have become African, Asian, and European, which, on the surface, may appear more informational and descriptive than ideological.

Some researchers make concerted efforts to distinguish population groups from racial groups either by using geographically oriented concepts, such as "geographic ancestry," specific names for groups, or indicating where race is used to indicate environmental rather than biological factors (Bamshad 2005; Royal & Dunston, 2004). At times, however, there are even cases where scientists say that race is not biologically valid, yet subsequently use it as a categorical variable. Troy Duster remarked in an interview with MIT's *Technology Review* that this contradiction exists in the same article and, sometimes, in the mind of a single scientist (Rotman, 2005). A survey of the scientific literature shows a number of different semantic moves scientists utilize in trying to reconceptualize human populations using language that omits race. Terms such as "biogeography of human populations," (Tishkoff & Kidd, 2004) "continentally defined groups," (Burchard, 2003), "human genome variation," (Royal & Dunston, 2004), and "ancient geographic ancestry," (Tang et al., 2005) are but a sample employed to 'get past' race. Keita and his colleauges (2004) offer a whole host of alternatives: ethnoancestral, bioethnic, ethnobiohistorical, ancestral-ethnic, social-designation, biocultural, biopopulation, ethnosocial, ancestral, ancestor-historical, origin group, and ethnogeographical. The HapMap Project does not use racial terms for the four sample groups, but, instead, the color-blind language that "reflects both the ancestral geography of each population and the geographic location where the samples" were collected[7].

The move to color-blind categories is a strategy to avoid reproducing racial meaning into genetic research and to elude historical biases. Interview respondents from the HapMap Project suggested that findings from population studies become racialized through misinterpretation of the data. For them, the key interpretation of population studies for the general population comes from the media (Interview 1001). For example, a bioethicist gave the case of an article by Rosenberg et al. (2002) that studied genetic variation in samples taken from a wide range of populations throughout the world (Interview 1004). The researchers categorized the groups by genetic similarity and "were careful not to use the word race to describe the populations that they found were genetically similar to each other" (Ibid.). This research is especially significant as it is the first publication out of the Human Genome Diversity Project database and has been cited widely in the scientific community as evidence for the genetic basis for racial groups. When the results entered the public sphere, she noted "race was very prominently used in the lay

---

[7] "Guidelines for Referring to the HapMap Populations in Publication and Presentation" http://www.hapmap.org/citinghapmap.html

reporting of that particular research" (Ibid; see also Risch et al., 2002 & Wade, 2004). While the media certainly sensationalize stories, such as referring to BiDil as the first "race-based" drug (Editorial, 2005) or "black drug" (Malik, 2005), research has shown that journalists tend to "get it right" (Bubela & Caulfield, 2004; Condit, 2004).

Back in the scientific journals, Risch (2006) cites the Rosenberg study as evidence that genetic clusters do align with racial groups. Sometimes researchers mix up race and ethnicity in examining group differences in health issue. For example, Haiman et al. (2006) investigate differences between African American, Japanese American, Latino, Native Hawaiian, and white men and women in their Multiethnic Cohort Study of the differences in rates of lung cancer from cigarette smoking. There are similar "black/white studies all over the place" (Interview 1006). Because of the technological advances in the last 10 years, there has become an astonishing capacity for gene mapping and typing. The cost for high throughput has declined dramatically making big science projects, such as the HapMap Project, much more feasible economically. The existence of this technological infrastructure has enabled scientists to do certain types of comparative analyses simply because they can. Studies that compare geographically disparate groups and tend to follow an African/European/Asian model may not be theoretically sound. However, the fluid nature of the digital data, stripped from its messy political and material basis, enables it to flow across communication networks and through data mining software. Peeking inside the categories can reveal the informational process, where national identity, ethnicity, culture, or race, becomes code, Lee suggests, "researchers feel little pressure to be explicit about the meaning and significance of racial and ethnic identity in framing their research hypotheses" (2005, p. 2136). A bioethicist commented that we are seeing more of these types of studies being published because of the existence of three different data sets from three different people being fed into a computer (Interview 1017). When the surface of the hypothesis is scraped, however, the assumptions guiding the study are grounded in a particular view of the world. For example, Rosenberg et al. (2002; not the Rosenberg above) compare MTHFR C677T polymorphism frequencies between three groups: whites, Japanese, and Africans. The authors use a mix of racial, national, and continental identity markers. Interestingly, the "white" population is referred to as Israeli and Arab. Said (1978) has well documented the racialization of the Oriental Other in opposition to the white European. Also, there is no rationale for comparing the groups. However, the authors are from Israel, Japan, and Ghana. This suggests that these are comparisons of convenience rather than theoretically guided. While the HapMap database is a major source of genomic data for global researchers to conduct genome comparison studies, an important factor for the rise of such racial group comparisons in the last fifteen years, especially between African, Asian, and white groups, can, in part, be traced back to an obscure mandate from the U.S. Office of Management and Budget in 1977.

Statistical Policy Directive 15 set out the standard for racial and ethnic identity in the census: American Indian or Alaskan Native, Asian or Pacific Islander, black, and white; Hispanic origin and not of Hispanic origin. The formalization and standardization of population categories to better capture the racial and ethnic diversity of the United States can be viewed as an early chapter in the informationalization of race. This informational infrastructure forces citizens to imagine themselves as part of a particular racial group. Hidden in the basic racial codes is the history of political battles and discursive struggles over the meaning of race in America. The statistics that arise from the government databases are not merely descriptive, but can be used to prescribe new social programs and access to resources. These categories have been used by many state and federal agencies as minimum standards for measuring disparities in a number of areas, such as employment and education, as well as health. While racial categories are utilized as a mode of state surveillance, they are also deployed as an act of strategic essentialism to organize political voice for minority populations. In the 1990s, civil rights advocates turned their focus to the sphere of health using state statistics to illustrate health and health care disparities between racial groups. In 1993, President Clinton signed the NIH Revitalization Act that mandated the inclusion of women and racial minorities as subjects in publicly funded clinical research. While the Act did not state how to operationalize race, the OMB categories became the norm. In 2000, this trend was bolstered when Congress passed the Minority Health and Health Disparities Research and Education Act #106-525.

While the reasons the organizers chose the initial HapMap populations are quite complex, they involve both scientific and geo-political issues which fall under the regulatory regime of the NIH. The OMB categories may not have dictated the sample populations, but it can be viewed as a contributor to them. A direct result of these federal policies and the HapMap database is a trend in reporting health disparities between racial groups as well as genetic group comparison studies, also known as genome-wide association studies. An irony of the OMB categories that have been used since the 1993 and 2000 Acts is that some genetic researchers find their research constrained by them. A HapMap geneticist explains that, "one of the challenges really is the way that it forces you to put people into the categories" (Interview 1022). Others have suggested that the categories have allowed researchers to use race variables in an unreflective manner (Ossorio, 2005). Recently, an extraordinary news item in *Science* reported that "In Asians and Whites, Gene Expression Varies by Race" (Couzin, 2007), pointing out genetic distinction between races. Where earlier discussions were much more careful in using terminology such as "populations" (Couzin, 2002), this title shows how discourse has returned to racial realism.

Conclusion

Oscillating between discourses of racial realism and color-blindness, and articulated through databases and the Internet, the struggle in science over the meaning of racial identity is (re)surfacing at a time when technology and race are interacting in new ways. The HGP was supposed to provide the final evidence, a *fait accompli*, that race is not a biological fact, because, at the molecular level, we are all the same, and race is a social and political construct. Aided by cutting-edge technologies, Venter and Collins were to provide the scientific proof for what the UNESCO Statements on Race signatories put into policy 50 years before. What was eagerly anticipated to be the final blow to bio-race was, in fact, a moment of transition. Instead of signaling the destruction of the biological notions that Gilroy and others had been hoping for — the end of race — the HGP introduced a new digital form of biopolitics. The very next year, a number of constituencies met to plan the HapMap Project and the next big step in science genome research. Instead of building on the HGP findings about race, however, the focus turned to difference. To many, this sounded like far-too-familiar ground in the ongoing struggles to debunk or shore up bio-race. However, the technologies that Venter developed, and the algorithms and code that were written under Collins' direction, like all processes of technological innovation, are dependent on actors and social context to shape them. The debates in the science journals about the role of race have waned somewhat, and the HapMap database has been used in a number of studies that measure and mark differences between groups. Some fear that a genomics framed by racial categories will deepen already-held racial assumptions about difference and biology. This may be true. To understand these trends, it is crucial to highlight how current debates about racism, race, and science are not framed in the same manner as the ones that closed the last century. They are well served by and, perhaps, reproduced by the color-blind information age where communication technologies play an increasingly important role. Race has not simply been digitized.

The discursive terrain in this conflict of maneuver is far more complex, and now it is hyperlinked through a digital communication network where the production of power is informational. Contemporary racialization operates at the convergence of new communication technologies, digital networks, and color-blind racism. Communication infrastructures are comprised of processes of creating, storing, and distributing information. Genomic biotechnology plays a particularly important role in the informationalization of race. Flesh and DNA, culture, and politics become computer codes in the seemingly technical and neutral space of genomic databases. To avoid a return to the altar of biological determinism, debates about genomics, health, and race need to be expanded to incorporate discussion of technology that goes beyond simple mentions of the presence of supercomputers in genomic research and their usefulness in sequencing DNA faster and cheaper. This development is taking place at a dizzying pace; the *next* HapMap Project — the 1000 Genomes Project — is already underway, and the largest biotechnology and information companies have collaborated to form a personal genomics company, 23andMe. New media technologies, such as databases, are now the new sites for negotiation over social and scientific fact and the politics of racial representation.

◇◇◇

Appendix - *Code for Interviews*

Interview 1001	Population geneticist
Interview 1003	Site Project Manager
Interview 1004	Bioethicist
Interview 1006	Microbiologist
Interview 1008	Bioethicist
Interview 1009	Bioinformatician
Interview 1011	Bioethicist
Interview 1013	Medical doctor and geneticist
Interview 1014	Human geneticist
Interview 1016	Population geneticist
Interview 1017	Bioethicist
Interview 1022	Geneticist and Bioethicist

The HapMap Project includes a number of sub-committees that work in the various geographic locations and whose capacities range from collecting DNA samples to setting the project's ethical, legal, and social parameters. The list of personnel includes geneticists, lawyers, anthropologists, bioethicists, doctors, bioinformaticians, project managers, biologists, an NGO director, pharmacologists, and senior scientists for a leading biotechnology company. The interview protocol and procedures were conducted under IRB regulations of the University of Southern California. Identification of subjects is public knowledge and was obtained through the International HapMap Consortium publications. Subjects were recruited by letter, email, and telephone calls. A number of the individuals interviewed were Principal Investigators of the various research sites. A couple of Committee Chairs also participated. Only the participants' broad area of expertise is included to protect their identities. Telephone and on-site interviews were conducted and recorded by the author. Though the interviews were recorded, I also took notes to help guide questions in the interview and for reflection post-interview.

When the interview was completed, I would go over the notes and make annotations for issues and items that could be addressed in subsequent interviews and/or analysis. After having them transcribed, I checked the transcripts against the recordings for accuracy. I developed a coding schedule, which was constantly being refined in an iterative fashion. Then, I used a qualitative software program, Nvivo, to code the interviews. I coded three main sub-sections first, on technology and race. Further coding was necessary until the response had been categorized sufficiently, usually on no more than three levels. After identifying a broad number of sub-codes, I refined them, merging similar codes and eliminating some if there were fewer than three responses to the code. In creating a report outline, I looked at what was said, how it corresponded to the literature in the area (or not), and then developed report sections that would correspond to both the literature and the data.

References

Bamshad, M. (2005). Genetic Influences on Health: Does Race Matter? *JAMA, 294*(8), 937-946.

Banton, M. (1998). *Racial Theories*. 2nd ed. Cambridge, UK: Cambridge University Press.

Barkan, E. (1992). *Retreat of Scientific Racism: Changing Concepts of Race in Britain and the United States Between the World Wars*. Cambridge: Cambridge University Press.

Barker, M. (1982). *The new racism: Conservatives and the ideology of the tribe*. London: Junction Books.

Bolnick, D. A., Fullwiley, D., Duster, T., Cooper, R. S., Fujimura, J. H., Kahn, J., et al. (2007). The science and business of genetic ancestry testing. *Science, 318*(5849), 388-400.

Bonham, V. L., Warshauer-Baker, E., & Collins, F. S. (2005). Race and ethnicity in the genome era: The complexity of the constructs. *American Psychologist, 60*(1), 9-15.

Bonilla-Silva, E. (2001). *White supremacy & racism in post-civil rights era*. Boulder, CO: Lynne Rienner Publishers.

Bonilla-Silva, E. (2003). *Racism Without Racists: Color-Blind Racism and the Persistence of Racial Inequality in the United States*. Boulder, Colorado: Rowman & Littlefield Publishers, Inc.

Bowker, G., & Star, S.L. (1999). *Sorting things out: Classification and its consequences*. Cambridge, MA: MIT Press.

Braun, L., Fausto-Sterling, A., Fullwiley, D., Hammonds, E. M., Nelson, A., Quivers, W., et al. (2007). Racial categories in medical practice: How useful are they? . *PLoS Medicine, 4*(9), 1423-1428.

British Medical Journal. (1996). Style Matters: Ethnicity, race, and culture: guidelines for research, audit, and publication. *312*(7038), 1094-.

Brown, M., Carnoy, M., Currie, E., Duster, T., Oppenheimer, D., Shultz, M., & Wellman, D.. (2003). *Whitewashing race: The myth of a color-blind society*. Berkeley, CA: University of California Press.

Bubela, T. M., & Caulfield, T. A. (2004). Do the print media "hype" genetic research? A comparison of newspaper stories and peer-reviewed research papers. *Canadian Medical Association Journal, 170*(9), 1399-1407.

Burchard, E., González, E., Ziv, N., Coyle, S., Gomez, H., Tang, A., Karter, J., Mountain, E., Pérez-Stable, D., Sheppard, & Risch,N.. (2003). The importance of race and ethnic background in biomedical research and clinical practice. *New England Journal of Medicine*. 348(12), 1170-1175.

Castells, M. 2000. *Rise of the network society*. 2nd ed, *The Information Age: Economy, Society and Culture*. Oxford: Blackwell Publishers.

Castells, M. (2001). Informationalism and the network society. In *The hacker ethic and the spirit of the information age*, edited by P. Himanen. London: Vintage.

Césaire, A. (1972). *Discourse on Colonialism*. New York: Monthly Review Press.

Collins, F., Green, E., Guttmacher, A., & Guyer, M. (2003). A vision for the future of genomics research: A blueprint for the genomic era. *Nature*. 422(24), 835-847.

Collins, F. S. (2004). What we do and don't know about 'race', 'ethnicity', genetics and health at the dawn of the genome era. *Nature Genetics, 36*(11), S13-S15.

Collins, P. (2004). *Black sexual politics: African americans, gender, and the new racism*. New York: Routledge.

Condit, C. (2004). Science reporting to the public: Does the message get twisted? *Canadian Medical Association Journal, 170*(9), 1415-1416.

Couzin, J. (2002). GENOMICS: New Mapping Project Splits the Community. *Science, 296*(5572), 1391-1393.

Couzin, J. (2007). HUMAN GENETICS: In Asians and whites, gene expression varies by race. *Science, 315*(5809), 173-174.

Cubitt, S. (2000). The distinctiveness of digital criticism. Screen. 41(1), 86-92.

Devlin, B., & Risch, N.. (1992). Ethnic differentiation at VNTR loci, with specific reference to forensic applications. *American Journal of Human Genetics*. (51), 534-548.

Duster, T. (2003). *Backdoor to Eugenics*. New York: Routledge.

Duster, T. (2005). Race and reification in science. *Science* 307 (5712), 1050-105.

Editorial. (2000). Race, ethnicity and nationality. *Paediatric and Perinatal Epidemiology*, *14*(1), 13.

Editorial. (2005, June 19). The First Race-Based Medicine*. The New York Times*, p. 11.

Evett, I., Buckleton, I., Raymond, A., & Roberts, H. (1993). The evidential value of DNA profiles. *Journal of the Forensic Science Society*. 33(4), 243-244.

Eze, E. C. (Ed.). (1997). *Race and the Enlightenment: A Reader*. Cambridge: Blackwell.

Fanon, F. (1963). *The Wretched of the Earth*. New York: Grove Press.

Fanon, F. (1967). *Black Skin, White Masks* (C. L. Markmann, Trans.). New York: Grove Press

.

Fausto-Sterling, A. (2004). Refashioning race: DNA and the politics of health care. *Differences*. 15(3), 1-37.

Fayyad, U. (1991). On the Induction of decision trees for multiple concept learning (machine Learning). Doctoral dissertation, The University of Michigan.

Foucault, M. (1977). *Discipline and punish: The birth of the prison*. 2nd ed. New York: Vintage Books.

Fullilove, M. (1998). Comment: abandoning "race" as a variable in public health research--an idea whose time has come. *American Journal of Public Health*, *88*, 1297-1298.

Fullwiley, D. (2007a). The molecularization of race: Institutionalizing human difference in pharmacogenetics practice. *Science as Culture*, *16*(1), 1-30.

Fullwiley, D. (2007b). Race and genetics: Attempts to define the relationship. *Biosocieties*, *2*, 221-237.

Galambos, L., & Sturchio, J. (1998). Pharmaceutical firms and the transition to biotechnology: A study in strategic innovation. The Business History Review. 72(2), 250-278.

Galloway, A. R., & Thacker, A. (2007). *The Exploit: A Theory of Networks*. Minneapolis: University of Minnesota Press.

Gandy, O. (2006). Data mining, surveillance, and discrimination in the post-9/11 environment. In K. Haggerty & R.

Ericson, (Eds.), *The new politics of surveillance and visibility*. Toronto, ON: University of Toronto Press.

Garfinkel, S. (2000). *Database nation: The death of privacy in the 21st century*. Sebastopol, CA: O'Reilly.

Gezelter, D. (2005). Catalyzing Open Source Development in Science: The OpenScience Project. Accessed March 28. http://www.openscience.org/talks/bnl/img0.htm

Gilroy, P. (2000). *Against race: Imagining political culture beyond the color line*. Cambridge, Mass.: Belknap Press of Harvard University Press.

Gould, S. J. (1996). *The Mismeasure of Man*. New York: Norton.

Gray, H. (2005). *Cultural moves: African Americans and the politics of representation*. Berkeley, CA: University of California Press.

Haggerty, K., & Ericson, R. (Eds.). (2006). *The new politics of surveillance and visibility*. Toronto: University of Toronto Press.

Haiman, C., Stram, D., Wilkens, L., Pike, M., Kolonel, L., Henderson, B., & Le Marchand, L. (2006). Ethnic and racial differences in the smoking-related risk of lung cancer. *New England Journal of Medicine*. 354(4), 333-342.

Hammonds, E. (1997). New technologies of race. In *Processed Lives: Gender and Technology in Everyday Life*, J. Terry & M. Calvert (Eds.). New York: Routledge.

Hayles, N. K. (1999). *How We Becaome Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago: University of Chicago Press.

Hinds, D. A., Stuve, L. L., Nilsen, G. B., Halperin, E., Eskin, E., Ballinger, D. G., et al. (2005). Whole-Genome Patterns of Common DNA Variation in Three Human Populations. *Science, 307*(5712), 1072-1079.

Holdsworth, D. (1999). The ethics of the 21st century bioinformatics: Ethical implications of the vanishing distinction between biological information and other information. In *Genetic Information. Acquisition, Access and Control*, A. K. Thompson & R. Chadwick (Eds.). New York: Kluwer/Plenum.

Hood, L. (2001). Under biology's hood. *Technology Review*. September.

Hudson, K., Javitta, G., Burkee, W., Byer, P., & Committee, A. S. I. (2007). ASHG statement on direct-to-consumer genetic testing in the United States *American Journal of Human Genetics, 81*(3), 635-637.

Ignacio, E. N. (2005). *Building diaspora: Filipino community formation on the Internet*. Piscataway, NJ: Rutgers University Press.

International Committee of Medical Journal Editors. (2006). Uniform requirements for manuscripts submitted to biomedical journals: updated October 2001. Retrieved May 31, 2006, from http://www.icmje.org/index.html

International HapMap Consortium. (2003). The International HapMap Project. Nature. 426(18), 789-796.

Jobling, M., & Gill, P. (2004). Encoded evidence: DNA in forensic analysis. *Nature Reviews Genetics*. 5, 739-752.

Jordan, W. (1974). *The White Man's Burden: Historical Origins of Racism in the United States*. New York: Oxford University Press.

Journal of American Medical Association. (2005). Instructions for Authors. *JAMA, 293*(14), 1788-1796.

Kaplan, J. B., & Bennett, T. (2003). Use of race and ethnicity in biomedical publication. *Journal of American Medical Association, 289*(20), 2709-2716.

Keita, S., Kittles, R., Royal, C., Bonney, G., Furbert-Harris, P., Dunston, G., & Rotimi, C. (2004). Conceptualizing human variation. *Nature Genetics*. 36(11), S17-S20.

Kim, C. (1999). The racial triangulation of Asian Americans. Politics & Society. 27(1), 105-138.

Knoppers, B. M., Hirtle, M., & Lormeau, S. (1996). Ethical issues in international collaborative research on the human genome: The HGP and HGDP. *Genomics, 34*, 272-282.

Kolko, B., Nakamura, L., & Rodman, G., (Eds.). (2000). *Race in cyberspace*. New York: Routledge.

Krimsky, S. (1999). The profit of scientific discovery and its normative implications. Chicago-Kent Law Review. 75(15), 15-40.

Landow, G. (1992). *Hypertext: The convergence of contemporary cultural theory and technology*. Baltimore, MD: John Hopkins Press.

Lee, S. (2005). Racializing drug design: Implications for pharmacogenomics for health disparities. *American Journal of Public Health*. 95(12), 2133-2138.

Lee, S. S.-J., Mountain, J., Koenig, B., Altman, R., Brown, M., Camarillo, A., et al. (2008). The ethics of characterizing difference: Guiding principles on using racial categories in human genetics. *Genome Biology, 9*(7), 404.

Lipsitz, G. (1998). *The Possessive Investment in Whiteness: How White People Profit from Identity Politics*. Philadelphia: Temple University Press.

Loro, L. (1995). Everyone's talkin' in the 'multilogue'. *Advertising Age*. 66(35), 28.

Lyon, D. (2002). Everyday surveillance: Personal data and social classifications. Information, Communication & Society. 5(2), 242-257.

Lyon, D. (2003). *Surveillance as Social Sorting: Privacy, Risk and Digital Discrimination*. London: Routledge.

Lyon, D. (2005). The sociology of information. In *The Sage Handbook of Sociology*, C. Calhoun, C. Rojec & B. Turner (Eds.). London: Sage.

Mackenzie, A. (2003). Bringing sequences to life: How bioinformatics corporealizes sequence data. *New Genetics and Society*. 22(3), 315-332.

Malik, K. (2005). Is this the future we really want? Different drugs for different races*. New York Times*, Jne 18, Online.

Manovich, L. (1999). Database as symbolic form. Convergence: *The Journal of Research into New Media Technologies*. 5(2), 80-99.

Marshall, E. (1998). DNA studies challenge the meaning of race. *Science* 282 (5389), 654-655.

Marturano, A., & Chadwick, R. (2004). How the role of computing is driving new genetics' public policy. *Ethics and Information Technology*. *6*: 43-53.

Memmi, A. (1965). *The Colonizer and the Colonized*. New York: Orion Press.

Miles, R. (1989). *Racism*. London: Routledge.

Moody, G. (2004). *Digital code of life: How bioinformatics is revolutionizing, science, medicine, and business*. Hoboken, NJ: John Wiley & Sons, Inc.

Mosse, G. (1978). *Toward the Final Solution: A History of European Racism*. New York: H. Firtig.

Nakamura, L. (2002). *Cybertypes: Race, ethnicity, and identity on the Internet*. New York: Routledge.

Nakamura, L. (2008). *Digitizing Race: Visual Cultures of the Internet*. Minnesota: University of Minnesota Press.

Nature Genetics. (2000). Editorial: Census, race and science. *24*(2), 97-98.

Nebert, D. W., & Menon, A. G. (2001). Pharmacogenomics, ethnicity, and susceptibility genes. *The Pharmacogenomics Journal*, *1*(1), 19-22.

Nelkin, D., & Tancredi, L. (1989). *Dangerous diagnotics: The social power of biological information*. New York: Basic Books.

Ng, P., Zhao, Q., Levy, S., Strausberg, R., & Venter, J. (2008). Individual genomes instead of race for personalized medicine. *Clinical Pharmacology & Therapeutics, 84*(3), 306-309.

Omi, M., & Winant, H. (1994). *Racial Formation in the United States: From the 1960s to the 1990s* (2nd ed.). New York: Routledge.

Oppenheimer, G. (2001). Paradigm lost: race, ethnicity, and the search for a new population taxonomy. *american Journal of Public Health, 91*(7), 1049-1055.

Osborne, N. G., & Feit, M. D. (1992). The use of race in medical research. *Journal of American Medical Association, 267*(2), 275-279.

Ossorio, P. (2005). Race, Genetic Variation, and the Haplotype Mapping Project *Louisiana Law Review, 66*(Special Issue), 131-143.

Patrinos, A. (2004). 'Race' and the human genome. *Nature Genetics, 36*(11), S1-S2.

Petsko, G. A. (2004). Color Blind. *Genome Biology, 5*(12), 199.191-199.193.

Poster, M. (1991). *The mode of information: Poststructuralism and social context*. Chicago: University of Chicago Press.

Rasmussen, T. (2000). *Social theory and communication technology*. Aldershot: Ashgate Publishing.

Razack, S. (1998). *Looking white people in the eye: Gender, race, and culture in courtrooms and classrooms*. Toronto: University of Toronto Press.

Risch, N., Burchard, E., Ziv, E., & Tang, H. (2002). Categorizations of humans in bio-medical research: Genes, race, and disease. *Genome Biology, 3*(7), 2007. 2001-2007. 2012.

Risch, N. (2006). Dissecting racial and ethnic differences. *New England Journal of Medicine*. 354(4), 408-411.

Rosenberg, N., Murata, M., Ikeda, Y., Opare-Sem, O., Zivelin, A., Geffen, E., & Seligsohn, U. (2002). The frequent 5,10-methylenetetrahydrofolate reductase C677T polymorphism is associated with a common haplotype in whites, Japanese, and Africans. *American Journal of Human Genetics*. 70(3), 758-762.

Rosenberg, N., Pritchard, J., Weber,J., Cann, H., Kidd, K., Zhivotovsky, L., & Feldman, M. (2002). Genetic structure of human populations. *Science*. (298), 2381-2385.

Rotman, D. (2005). Race and medicine. *Technology Review* 107(4).

Royal, C., & Dunston, G. (2004). Changing the paradigm from 'race' to human genome variation. Nature Genetics. 36(11), S5-S7.

Said, E. (1978). *Orientalism*. New York: Vintage.

Sankar, P., & Cho, M. 2002. Toward a new vocabulary of human genetic variation. *Science* 298, 1337-1338.

Schiller, D. (1999). *Digital capitalism: Networking the global market system*. Cambridge, MA: MIT Press.

Schwartz, R. S. (2001). Racial profiling in medical research (editorial). *New England Journal of Medicine, 344*(18), 1392-1393.

Stepan, N. (1982). *The Idea of Race in Science: Great Britain, 1800-1960*. London: Macmillan.

Taguieff, P. A. (1990). The new cultural racism in France. *Telos, 83*(spring), 109-122.

Tishkoff, S., & Kidd, K. (2004). Implications of biogeography of human populations for 'race' and medicine. *Nature Genetics*. 36(11), S21-S27.

UNESCO. (1953). *The race concept: Results of an inquiry*. Paris: UNESCO.

UNESCO. (1969). *Four statements on the race question*. Paris: UNESCO.

Wade, N. (2007). Studies Find More Genes Underlying Diabetes. *New York Times*, April 26, online.

Wade, N. (2004). Articles highlight different views of genetic basis of race*. New York Times,* October 27, online.

Waldrop, M. Mitchell. 2001. Data Mining. *Technology Review* January/February:online.

Webster, F. (2002). *Theories of the information society*. London: Routledge.

Wellcome Trust Case Control Consortium. 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447 (7145):661-678.

Wellman, B. (2004). The three ages of Internet studies: ten, five and zero years ago. *New Media & Society*. 6(1), 123-129.

Wellman, D. (2003). Color blindness as color consciousness. Paper read at Color-Blind Racism Conference, October 3, at Stanford University, Palo Alto, California.

Winker, M. A. (2004). Measuring Race and Ethnicity: Why and How? *Journal of American Medical Association, 292*(13), 1612-1614.

Zarsky, T. (2003). "Mine your own business!": Making the case for the implications of the data mining of personal information in the forum of public opinion. *Yale Journal of Law & Technology*. 1-57.

……

Peter A. Chow-White
School of Communication
Simon Fraser University
8888 University Drive
Burnaby, BC V5A 1S6
Canada
Tel: 604.603.3794