Penultimate draft. Published in The Journal of Philosophy (2013), 110(6):331-351
If you do not have access to the final version feel free to ask me for a copy
adam.hochman83@gmail.com

AGAINST THE NEW RACIAL NATURALISM[*]

Adam Hochman

The idea that there are human races makes a claim about human biological diversity. The nature of this claim, however, is a matter of longstanding dispute. Scientists have presumed it to involve the taxonomic division of the human species into *subspecies*, understanding "subspecies" as the biological synonym of "race." Since the early fifties there has been growing scientific agreement—largely due to the UNESCO Statements on Race[1] and an influential paper by Richard Lewontin[2]—that there are no human subspecies, and hence that "race" is not a scientifically respectable category. Almost lost in the tide of support for Lewontin was a critique by Jeffry Mitton,[3] which demonstrated that Lewontin's statistical methods were insensitive to the correlation structure of the genetic data—structure that Mitton argued could be used to group people into traditional racial categories.

Mitton's criticism of Lewontin's work had gone relatively unnoticed until recently. Testament to this is an article by A. W. F. Edwards,[4] published in 2003, which makes the same argument that Mitton had made more than two decades earlier, but with no reference to Mitton.[5] The timing was not coincidental. In 2002 Rosenberg et al.,[6] using newly available data from the HGDP-CEPH Human Genome Diversity Cell Line Panel, identified six genetic clusters, five of

---

[*] I would like to give special thanks to Paul Griffiths for his helpful and encouraging comments on several drafts. I am also grateful to Pierrick Bourrat, David Braddon-Mitchell, Zvi Hochman, Frances Olive, Maureen O'Malley, Kristie Miller, Karl Rollings, Luke Russell, Karola Stotz and John Wilkins for their valued feedback. Thanks to an audience at the 2011 Sydney-ANU philosophy of biology workshop, and especially to Kim Sterelny for a particularly helpful question. This research was supported under Australian Research Council's Discovery Projects funding scheme DP0878650 and an Australian Postgraduate Award through the University of Sydney.

which roughly correspond to major geographic regions. This study confirmed Lewontin's finding that racial classification accounts for very little (around 5%) of the overall human genetic diversity, while appearing to offer some support to the view that when the correlation structure of genetic data is analyzed, clusters emerge which (at least roughly) correspond to a traditional racial taxonomy.

How to best interpret these findings is still being debated. The Mitton/Edwards critique of Lewontin has gained support from Neven Sesardic[7] in his recent attempt to revive race as a legitimate scientific category. He argues that philosophers, most of whom are social constructionists about race, attend insufficiently to the science on the topic. Yet, according to Sesardic, we can no more trust the words of scientists, whom he believes self-censor on this politically sensitive issue. "A poll about views of race," he quotes Henry Harpending as saying, "would be like a poll about Marxism in East Germany in 1980. Everyone would lie."[8] Philosophers tell, but do not know. Scientists know, but will not tell.

Against this mood of extreme skepticism, and against racial naturalism, I will argue that the science, when properly understood, actually supports social constructionism about race. My critique is based on two main arguments. The first is that the criteria race naturalists apply to humans are not consistent with those applied to define subspecies in nonhuman animals, and that no rationale has been given for this differential treatment. The second is that the continental clusters appealed to by race naturalists are biologically superficial groupings that should not be elevated to subspecies status. Race naturalists have been too hasty in their racial interpretation of genetic clusters. The science has moved on since the early seventies, but it has vindicated, rather than contradicted the view that there are no human subspecies.

I. RACIAL NATURALISM AND SOCIAL CONSTRUCTIONISM ABOUT RACE

Racial naturalism is the view that humans can be divided into subspecies, and that “race” is therefore a valid scientific category. In this paper I focus on *genetic* naturalism about race. There are other naturalistic accounts of race: Robin Andreasen’s[9] cladistic approach, Philip Kitcher’s[10] definition of races as reproductively isolated populations, and Massimo Pigliucci and Jonathan Kaplan’s[11] “ecotype” account. Genetic naturalism about race diverges from these accounts in three important respects, aside from its genetic focus. First, genetic naturalism about race is usually supposed to vindicate a traditional racial taxonomy, rather than to redefine racial categories based on new theories and new findings. Second, genetic naturalism about race has a *phenetic*, rather than a *cladistic* approach to taxonomy. In other words, it focuses on genetic (and supposedly genetically determined phenotypic) differences between populations, rather than on genetically inferred evolutionary relationships between populations. Third, this is the account of race that is mobilized to support racial claims of intellectual, cultural, and moral superiority (this is sometimes referred to as “ogre naturalism”).[12] Henceforth, “racial naturalism” will be used to refer specifically to genetic naturalism about race.

Racial naturalism depends not only on the validity of human subspecies classification, but also more generally on the taxonomic validity of the “subspecies” category itself. It is important to note that the subspecies concept has been the focus of considerable debate over the last century. In 1953 Edward Wilson and William Brown argued that “the subspecies concept is the most critical and disorderly area of modern systematic theory.”[13] This view is still commonly held. Robert Zink, for instance, found that “Mitochondrial DNA sequence data reveal that 97% of continentally distributed avian subspecies lack the population genetic structure indicative of a distinct evolutionary unit."[14] Phenetics, in particular, has attracted strong criticism.[15] Daniel Mulcahy describes a recent trend in eliminating subspecies altogether: "Discrete, diagnosable lineages are elevated to specific [species] status, while those that show clinal [gradual] variation and/or appear to represent ecological pattern classes are placed in synonymy with the parent species and the subspecific epithets are disregarded."[16] The debate surrounding racial naturalism does not take place against anything like a stable scientific backdrop.

For the purposes of this paper it will be useful to temporarily "black box" the question about the general validity of taxonomic divisions below the species level. We can proceed under the provisional assumption that "subspecies" is an orderly scientific category, and that for racial naturalism to be defensible "race" must offer a nonarbitrary and nonvague representation of human genetic diversity and population structure, and, as a result, lead us towards achieving far-ranging explanatory goals (for example, in biomedicine). Since biology's "modern synthesis" and the demise of species essentialism, there is no imperative to be essentialist about subspecies. However, as Guido Barbujani argues, "no classification is useful if the classification units are vague or controversial."[17] "Races" would have to be robust, biological kinds.

Social constructionism about race is the view that our racial categories are predominantly determined by social factors, even though some of their inclusion criteria will be biological. The title "social constructionism" has an unfortunate implication: that race, as a social construction, has no biological correlate. This is how race naturalists tend to interpret social constructionism.[18] Yet this is a straw-man characterization of the constructionist position. Lisa Gannett writes that for the constructionist, "race is socially constructed by enlisting biological differences and investing these with socio-cultural meanings."[19] Gannett's definition discounts an interpretation of social constructionism which would suggest that racialized groups have no biological correlates, while leaving open the possibility that “race” is correlated with only the most superficial (if sometimes medically and forensically useful) biological features.

The danger is to present the debate as a simple dichotomy, with social constructionists arguing that “race” is merely human projection (and prejudice) and race naturalists arguing that it is a biological fact, imposed on us by the world. A continuum is more reflective of the positions that are actually held, with “pure social construction” on one end, “exemplar natural kind” on the other, and everyone in the debate placed somewhere along the continuum. Yet even this is misleading, for two reasons. Firstly, the “no biological correlate” end of the continuum is not an extreme version of social constructionism, but a different view altogether. To hold this view one would have to argue that skin color, a prototypically biological feature, is not correlated with racialized groups. As social constructionists often point out, skin color and “race” do come apart, but nobody argues that people racialized as “white” have, on average, darker complexions than people racialized as “black,” or that skin color varies randomly. Secondly, a continuum from “pure social construction” to “natural kind” is misleading because it is unidimensional. It would suggest that there is one way of “carving the world at its joints” and that a racial taxonomy, at one extreme, fails to even scratch nature’s joints, and, at the other extreme, represents deep cuts in nature. This is metaphysically dubious. As Kitcher argues in his useful discussion on the future of “race,” “there is a nondenumerable infinity of possible accurate maps we could draw for our planet,” and there is “no feasible project of inquiry (singular) that aims at a complete account of our world, but rather many inquiries driven by specific questions we find it important to answer.”[20]

The metaphysical problem is best illustrated by its practical counterpart. Racial labels may be more or less useful proxies for biological variation in different scientific disciplines, and each discipline may support racial naturalism or social constructionism to a greater or lesser extent. It is crucial that these points be separated. For instance, forensic anthropologists are able to allocate skulls, with fairly high accuracy, to the racialized group with which the deceased would have been identified. Does this show that racial labels are useful proxies for biological variation in physical forensic anthropology? Yes. Does this support racial naturalism? No. Why? Because forensic anthropologists translate trait measurements to the "racial" taxonomies societies use, or would have used, to describe missing persons; they do not generally ask whether those taxonomies are scientifically valid. When they do it is in a research context, and their answer is firmly negative.[21] Forensic anthropologists are also able to allocate skulls to groups that are separated culturally, linguistically, politically, and historically, and at a finer grain than a racial taxonomy offers. "Race" latches on to human morphological variation, but in a crude way. Forensic anthropologists use racial classification, but this is because we, the public, classify our missing persons racially, not because it is a precise or privileged representation of human morphological diversity.

The debate between race naturalists and social constructionists is best framed not as a dichotomy, nor a continuum, but as a discussion worth having in various scientific, philosophical, and political contexts. This discussion may one day end in eliminativism about race. In the meantime, the social constructionist will expect correlations between racialized groups and various traits. Yet she will, in Gannett's words, expect such correlations to be "statistical not universal, local not global, contingent not necessary, and accidental not lawful, and expect their corresponding cuts in nature to be interest-relative not mind-independent, dynamic not static, indeterminate not determinate, many not few, overlapping not nonoverlapping, and superficial not deep."[22] Having framed the debate along these lines, let us consider the arguments for racial naturalism.

II. THE PRACTICAL AND THEORETICAL ARGUMENTS FOR RACIAL NATURALISM

In his 1972 article "The Apportionment of Human Diversity" Lewontin demonstrated that, on a locus-by-locus basis, around 85.4% of the overall human genetic diversity resides within any given continental population, so roughly within the traditional "racial" groups. Another 8.3% of our genetic diversity, he showed, is accounted for by genetic variation within racialized groups. The remaining 6.3% of the total human genetic diversity is specific to racialized groups. "Racial classification," Lewontin concludes, "is now seen to be of virtually no genetic or taxonomic significance."[23]

Lewontin's analysis of the apportionment of human genetic diversity has been confirmed by numerous studies.[24] However, his conclusion regarding racial classification has been the subject of some debate. The arguments against Lewontin's conclusion have two sources. One is practical, and derives from genetic-clustering studies; the other is theoretical, and challenges his statistical methods. I shall outline these arguments in turn.

The practical argument for racial naturalism claims that genetic-clustering studies empirically confirm the presence of human subspecies. According to Risch et al., "these population genetic studies have recapitulated the classical definition of races based on continental ancestry."[25] Indeed, genetic-clustering studies have shown that despite the small proportion of genetic variation separating continental populations, it is possible to assign some (geographically separated and not recently admixed) individuals to their (or their ancestors') continents of origin based on genetic data alone. Whether this indeed vindicates racial naturalism, or only supports a weaker notion of "biogeographical ancestry," is another matter (to which we shall return).

The most influential of the genetic-clustering studies is Rosenberg et al.'s 2002 investigation of the genetic structure of worldwide populations.[26] Using the statistical program STRUCTURE, Rosenberg and his team, having access to samples from 52 populations, attempted to infer worldwide population structure at five different grains of analysis. The 52 populations were divided into seven regions: Africa, Europe, the Middle East, Central/South Asia, East Asia, Oceania, and America. At the roughest grain of analysis, where the program was set to distinguish two groups, the clusters were anchored by Africa and America. When the program was set to distinguish five groups, genotypes from Europe, the Middle East, Central Asia, and South Asia were clustered together, and those from the other four regions were clustered separately. At the finest grain of analysis the Kalash of northwest Pakistan were added as a sixth distinct cluster, which the investigators attributed to their suggested European or Middle Eastern origins. Although these clusters do not correspond to any previously proposed racial taxonomy,[27] and many of the individuals sampled have substantially mixed cluster membership, this study is considered by race naturalists as strong evidence for the practical argument for racial naturalism.

The theoretical argument for racial naturalism was first proposed by Mitton and then later, independently, by Edwards. They argue that Lewontin analyzed the genetic data under the false assumption that a locus-by-locus analysis was sufficient to reveal all of the information necessary for racial classification. Edwards calls this "Lewontin's fallacy." Lewontin's reasoning was fallacious, according to Edwards, because a single-locus analysis is insensitive to correlations amongst loci—correlations which he believes can be used to group people into traditional racial categories, irrespective of low between-population differences on a locus-by-locus basis.

To illustrate how correlation structure can be useful for classificatory purposes I have constructed three figures, using mostly made-up data about maize and wheat yields (real data would be similar but messier). Figure 1 shows that when the weights of the maize and wheat yields are compared (in tons per hectare) they largely overlap. Figure 2 shows even greater overlap when comparing the rainfall that the maize and wheat yields received during the season in which they grew. Figure 3 shows that, despite these overlaps, when the measures are correlated the maize and wheat yields are clustered apart. The race naturalist believes that clustering studies are analogous: that while "races" overlap genetically, they can be clustered by analyzing the correlation data. It is the particular correlation structure of mainly species-wide alleles, claims the race naturalist, which causes "race."

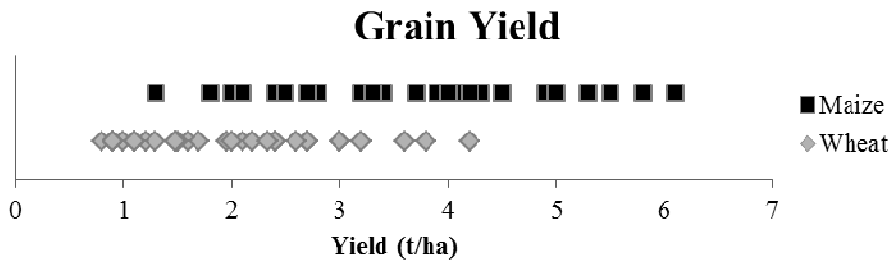

Figure 1

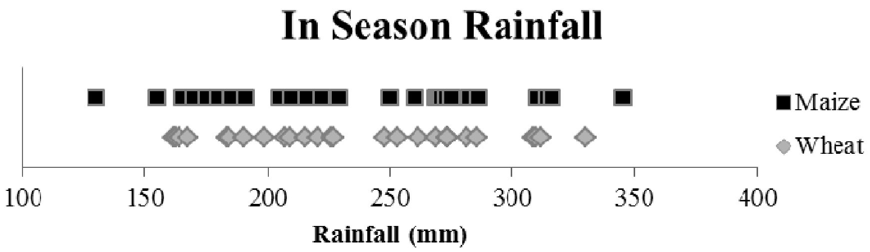

Figure 2

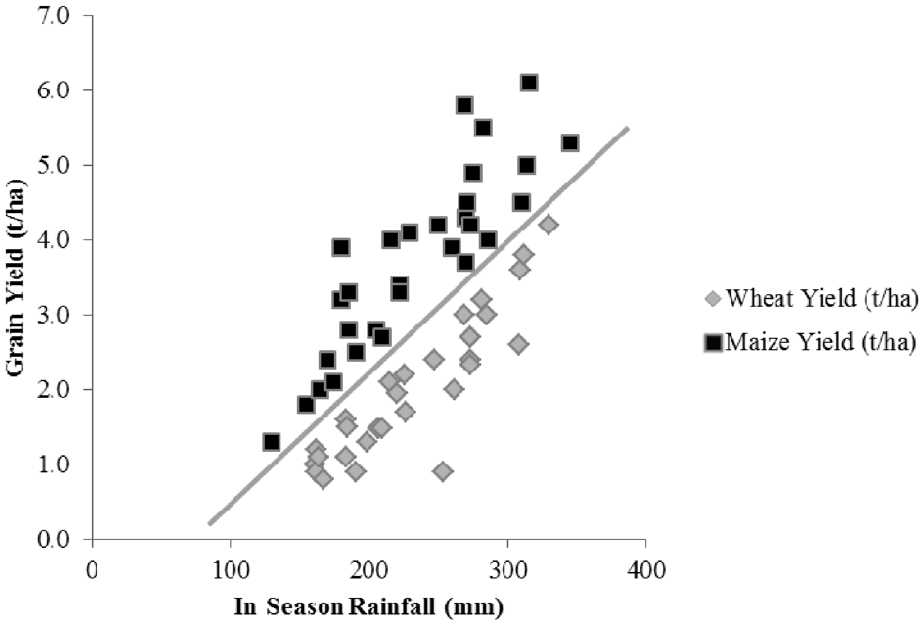

Figure 3

Before proceeding further, it will be useful to review the relevant statistics. The most common single-locus statistic used to measure genetic diversity or variation within a species is Wright's Fixation Index, or F_{ST}. It compares the amount of genetic variation within one or more subpopulations to that found in all populations combined (the total population). $F_{ST} = (T–S)/T$, where T represents the average difference between allele pairs selected randomly from the total population and S represents the average difference between allele pairs selected randomly from within a subpopulation. If mating in the total population is random, T and S will have equal values, and the F_{ST} estimate will be 0. In this case, the population has no genetic subpopulations. If there is some nonrandom mating (which there will be if individuals are geographically distant, for instance) the F_{ST} estimate will rise.

The standard criterion for subspecies division is an F_{ST} estimate of or over 0.25.[28] This criterion is, of course, subjective. Conservation biologists sometimes take a lower F_{ST} estimate to be significant, in order to protect endangered species. Nevertheless, the standard criterion appears to be agreed upon in the human case (race naturalists do not argue that the cutoff should be lowered, but that F_{ST} is the wrong statistic altogether). Human F_{ST} estimates are around 0.05 to 0.15, indicating that humans, under the standard genetic criterion, cannot be divided into subspecies.[29]

According to the new race naturalists F_{ST} is the wrong statistic for human subspecies classification. They appeal, instead, to the multilocus technique of genetic clustering. The most popular clustering program is STRUCTURE. This was the software used by Rosenberg and colleagues in their analysis of the genetic structure of human populations. The underlying methodology was first described in 2000 by Jonathan Pritchard, Matthew Stephens, and Peter Donnelly.[30] STRUCTURE allocates individuals to K populations, or clusters, where K is chosen in advance. Clusters are defined by their allele frequency correlation structures, and individuals are probabilistically allocated to one or more clusters. Joint cluster allocation indicates admixture (mixed population membership). Unlike with F_{ST} there is no standard criterion for delineating subspecies when clustering. STRUCTURE is understood to reveal “population structure.” The inference from “population” to “subspecies” is made by race naturalists.

To prevent confusion, I will use the terms "diversity" and "variation" when referring to the results of single-locus studies, and "population structure" when referring to those of clustering studies. While this terminological distinction is already in common usage, it is not always adhered to. This may make it appear as though the same thing is being measured in the two kinds of studies. As I have explained, this is not the case.[31]

III. IS 'LEWONTIN'S FALLACY' A FALLACY?

There is agreement with Lewontin, on both sides of the debate, that the proportion of genes that are unique to any given racialized group is very small. The arguments for racial naturalism target the conclusion Lewontin draws from this: that racial classification has no taxonomic significance. Contrary to Lewontin, Edwards argues that "The 'taxonomic significance' of genetic data in fact often arises from correlations amongst the different loci, for it is these that may contain the information which enables a stable classification to be uncovered."[32] To draw conclusions about racial classification based on data that have not been correlated is to commit "Lewontin's fallacy."

"There is nothing wrong with Lewontin's statistical analysis of variation," argues Edwards, "only with the belief that it is relevant to classification."[33] Notice that this claim about relevance begs the question. Edwards believes that single-locus statistics are not relevant to racial classification because they do not support racial classification. The question of relevance remains unanswered. If single-locus studies had found that populations were genetically distinct it seems as though Edwards would consider them relevant to classification. Edwards needs to fault the relevance of Lewontin's analysis of variation, where "relevance" pertains to the validity of the analysis, not to the contents of its results.

“Lewontin’s fallacy” is not a genuine fallacy. Mitton and Edwards are correct that Lewontin’s analysis of genetic variation was blind to correlation data that could potentially be used for classificatory purposes. Whether this correlation data would support human *subspecies* classification is another question. Edwards argues that Lewontin “used his analysis of variation to mount an unjustified assault on classification, which he deplored for social reasons.”[34] What Edwards seems to overlook is the fact that subspecies classification is a special case of classification in general. A classification scheme can reliably pick out kinds that are relatively superficial. Think, for instance, of amateur ornithology. Human subspecies, or “races,” are putative biological kinds. They cannot, by definition, be superficial.

To show that a fallacy has been committed there would need to be a forceful argument for using genetic clustering, and against using single-locus statistics, for subspecies classification. The fact that Lewontin’s analysis of human genetic diversity does not furnish us with a racial taxonomy is not reason to reject it. No argument has been given against the use of single-locus statistics beyond the suggestion that they are not powerful enough to provide evidence for racial naturalism. Moreover, no argument has been offered in support of the claim that genetic clusters distinguish subspecies, rather than a modicum of population structure, which of course can be present in the absence of subspecies.

I do not wish to be interpreted as arguing for the taxonomic validity of divisions below the species level, and even less for endorsing gene-centric biology. However, under the conventional population-genetics conception of subspecies, F_{ST} is more appropriate than clustering as a statistical tool for identifying subspecies divisions. The population-genetics approach to subspecies focuses on isolation and genetic differentiation. Genetic differences between groups develop under isolation because lost alleles are not reintroduced and new mutations are not spread. Under such conditions, the kinds of genetic discontinuities develop that lead biologists to apply the subspecies concept. These genetic discontinuities are represented by high F_{ST} values. However, when migration is high, species-wide genetic variation tends to be *clinal*, or gradual. Clinal variation indicates that gene flow, rather than isolation, has been the main evolutionary force shaping genetic diversity. This has certainly been the case for our species,[35] which explains our low F_{ST} values. The fact that F_{ST} measures genetic differentiation caused by reproductive isolation, while clustering finds cryptic genetic structure without lengthy periods of reproductive isolation, suggests that F_{ST} is the more appropriate statistic for delineating subspecies according to the population-genetic conception of subspecies that race naturalists champion.

Perhaps the most obvious objection to the idea of "Lewontin's fallacy," however, is that in nonhuman biology single-locus statistics are the standard tools used for subspecies classification. All Lewontin did was to shift the focus to humans. Surprisingly, Edwards does not foresee this objection to the idea of "Lewontin's fallacy." This may be because he has a weak notion of race, where a resemblance between racialized groups and clusters (when the clustering software is set to the right grain of analysis) is all that is needed to confirm a racial taxonomy. Edwards does not endorse racial naturalism as I have described it, because he does not argue that "races" constitute subspecies, only that they can be clustered. Edwards is no social constructionist, but his argument does not actually contradict social constructionism. Social constructionists expect some correlations between genes and racialized groups, just as they expect skin color and "race" to be roughly correlated. The real question is whether a traditional racial taxonomy picks out subspecies. Edwards has not answered this question. The fact that individuals from geographically distant and not recently admixed populations can be clustered according to their continental origins does not, on its own, constitute an argument for human subspecies division.

While Edwards is not, according to my definition, a fully fledged race naturalist, Sesardic has adopted his work to make a case for racial naturalism. Sesardic also avoids the above objection to the idea of "Lewontin's fallacy," but by claiming that multilocus statistical techniques (such as clustering) are used to define subspecies in nonhuman animals:

> In biology, the concept of "race" is often regarded as synonymous with "subspecies". Subspecies are populations of organisms that, despite belonging to the same species, differ among themselves with respect to frequencies of alternative alleles at a number of loci….Research has shown that, indeed, groups of people of significantly different geographical ancestries do differ from one another genetically: when compared on many genetic loci these groups have different frequencies of different alleles.[36]

Sesardic claims that he is simply applying the subspecies concept from nonhuman to human biology. However, this passage suggests that multilocus, rather than single-locus statistics are used to determine whether nonhuman species are divisible into subspecies. This is misleading, as the new multilocus clustering methods have not, to my knowledge, been used to redefine subspecies in nonhuman animals.

Sesardic implies that humans should not be treated as a special case—that what applies to nonhuman animals applies to humans, and *vice versa*.[37] Yet his characterization of nonhuman taxonomy is misleading, as clustering methods are not used to define subspecies in nonhuman biology. If we agree with his parity argument there is no "Lewontin's fallacy." There is either a "subspecies fallacy," and we need to disregard all subspecies divisions defined using single-locus statistics and begin again—this time with multilocus statistics—or there is no fallacy to speak of, in which case we should continue using single-locus statistics and abandon the idea of human racial classification. Race naturalists might prefer the former approach, but the fact that humans have no subspecies according to the standard genetic criterion seems a very bad reason to revolutionize taxonomy.

"Lewontin's fallacy" is about human subspecies classification. However, the inference from "genetic cluster" to "subspecies," fundamental to racial naturalism, stands unsupported. Correlation structure can indeed be useful for classification, but only when we have a clear interpretation of what is being classified. To illustrate this point let us return to the maize- and wheat-yield example. This example showed that the correlation data enabled crop classification. Race naturalists believe that subspecies classification is analogous. They argue that the genetic data need to be correlated for human subspecies classification, and that it is the correlation structure amongst mainly common genes which causes "race." The analogy does not bear scrutiny, however, because nobody thinks that what makes maize maize or wheat wheat is a joint probability based on yield weight and the amount of rainfall received! We can verify these crop clusters against the scientific definitions of maize and wheat. But what are we comparing genetic clusters to? We compare them to a folk racial taxonomy. As it happens, the "yield weight" to "in season rainfall" relationship is a fairly reliable differentiator between wheat and maize crops.[38] But genetic clusters, derived from gene correlation data, may not serve as proxies for "race." Genetic clusters may represent something entirely different. In the following section, this is what I argue.

IV. ARE THERE HUMAN SUBSPECIES?

Rosenberg and colleagues have dismissed racial interpretations of their work. “Our evidence for clustering,” they insist, “should not be taken as evidence of our support of any particular concept of ‘biological race’.” They continue, “The arguments about the existence or nonexistence of ‘biological races’ in the absence of a specific context are largely orthogonal to the question of scientific utility.”[39] Sesardic attributes such reticence in accepting racial naturalism to “reasons of the heart.”[40] However, what Rosenberg and colleagues offer is a testable scientific hypothesis: that the scientific utility of genetic-cluster studies is context dependent. Continent-based clusters do not provide the predictive power, it is implied, which we can expect from biological kinds. Genetic-clustering studies may inform research on human migratory history, but offer poor guides to the biological diversity relevant to biomedicine, for instance. If this is true—and at a continental grain of analysis it is certainly true[41]—it would be wrong to interpret genetic clusters racially, as picking out phenetic subspecies.

In their worldwide cluster analysis, Rosenberg and colleagues identify six genetic clusters. However, in their analysis of the genetic structure of continental populations, they find substantial within-continent structure. Structure that is hidden at a rough grain of analysis is revealed at a finer grain. One of the key problems for racial naturalism, as identified by Kitcher,[42] is that it suffers from a grain-of-resolution problem. The race naturalist cannot defer to clustering studies to number and name the races. The appropriate grain of analysis is unclear. Is there one American race, as suggested by a racial reading of the worldwide analysis, or five, the number of clusters identified in the within-continent analysis? The first suggestion confers an arbitrary privilege to a rough grain of analysis. The second suggestion could not be correct, as the individuals sampled were Colombian, Maya, Karitiana, Surui, and Pima—hardly a comprehensive sample of the numerous indigenous and meztizo peoples of the Americas. The Americas have hundreds of potentially genetically distinguishable populations. The second suggestion leads to a *reductio ad absurdum*, as the race naturalist would have to concede that we have hundreds, even thousands of "races." If we want to be consistent across species, similar extravagances would extend to the rest of the animal kingdom.

One prominent race naturalist, Armand Leroi, is willing to accept such a proliferation of races:

> there is nothing very fundamental about the major continental races; they're just the easiest way to divide things up. Study enough genes in enough people and one could sort the world's population into 10, 100, perhaps 1000 groups, each located somewhere on the map.[43]

Leroi, like Edwards, presents himself as a race naturalist. However, neither author appears to make a strong case for racial naturalism. Edwards fails to tell us what is fundamental about the major continental "races." Leroi tells us that they are not fundamental after all. If each possible cluster is a "race," and we could indeed distinguish thousands, the concept is virtually meaningless. Why continue to use it, especially given the devastating history of racial classification? Leroi contends that "no other noun seems to do the job,"[44] although he uses "groups" as a synonym in the quote above. Seeing as there is "nothing very fundamental about the major continental races," it would be more accurate to refer to them as "racialized groups," rather than "racial groups" or "races." We all have ancestries, but we do not—as I will continue to argue—belong to biological races.

The relatively small sample size of the Rosenberg et al. study points to another problem for racial naturalism. The idea that there are human races suggests that human biological (in this context genetic) diversity is best represented as clustered, or as forming discreet groups. As David Serre and Svante Pääbo[45] point out, a very bad way to test this hypothesis would be to sample from sites that are themselves clustered (geographically distant) and then ask if our genetic structure is best represented as clustered. This is just the criticism Serre and Pääbo leveled against Rosenberg and colleagues. To test the validity of clustered representations individuals would have to be sampled continuously from region to region. Such a sampling scheme would allow an alternative hypothesis—that the genetic structure of worldwide populations is clinal, or gradual, rather than clustered—to be evaluated. If our genetic structure is clinal, racial naturalism would be refuted. It is impossible to "carve nature at its joints" if it does not have any.

In response to Serre and Pääbo, Rosenberg et al.[46] provided evidence in support of a synthetic model of the genetic structure of human populations, integrating clinal and clustered representations. They showed, in favor of a clinal model, that "allele frequency differences generally increase gradually with geographic distance."[47] What role, then, do clusters play? Rosenberg et al. explain that it is "small discontinuous jumps in genetic distance—across oceans, the Himalayas, and the Sahara—that provide the basis for the ability of STRUCTURE to identify clusters that correspond to geographic regions."[48] This is fatal to racial naturalism. The genetic structure of human populations is almost entirely clinal. Clusters reflect slight discontinuities in the genetic structure of populations between geographical barriers to human dispersal. But why, then, do the clusters in Rosenberg and colleagues' study roughly resemble traditional "racial" groupings? The answer is simple: because we have roughly separated "racial" groups around these geographical barriers.

When using F_{ST} we can defer to a standard criterion in response to the subspecies question. The same cannot be said for the new genetic-clustering methods, as no such criterion has been formulated. We can, however, stipulate some conditions that would need to hold before seriously asking whether clusters represent meaningful biological units, such as subspecies. I will suggest four such conditions (they are not presumed to be comprehensive).

First, the range of allele frequency differences between clusters would have to be relatively uniform. There should not be more difference in genetic diversity and structure within one cluster than between that cluster and another. The second condition is that the number of clusters should not be arbitrary. As a third condition, the allele frequencies within a cluster should be relatively homogenous (not too clinal). The fourth condition is that there should be a large jump in genetic difference between clusters. If the first and second conditions are met, but not the third, clusters might still be meaningful biological units when there is a large jump in genetic distance between them (when the fourth condition is met). Alternatively, if all of the conditions but the fourth are met, the clusters may still be meaningful biological units. That is, a small jump in genetic distance might separate meaningful biological units when those units are internally homogenous, are nonarbitrary in their number, and do not vary greatly in the amount of genetic diversity and structure that they house.

None of the above conditions hold in the human case. Let us begin with the first condition. Because clustering is a product of discontinuities in population structure it is blind to smooth, clinal changes in the structure of populations. As a result, varying amounts of population structure (and genetic diversity) can reside within any one low-resolution cluster. This creates a dilemma for the new race naturalists, who interpret continental clusters racially. For the race naturalist, two within-continent “sub-clusters” should not be more different than their primary "racial" cluster and another "racial" cluster; this would be grounds for further subspecies division. Yet this is just what we find with Africa, for instance. Africa houses the most genetic diversity of any continent, and cluster analysis yields high within-continent structure.[49] As Yu et al.[50] have shown, there is a larger genetic difference among Africans than between Africans and Eurasians. The first condition fails in the human case.

The second condition stipulates that the number of clusters should be nonarbitrary. There needs to be a principled reason for deciding on an appropriate number of clusters. In the human case, for instance, the chosen number should not simply reflect the folk assumption that there are only a handful of races. As I have argued, racial naturalism is faced with a serious grain-of-resolution problem. It is unclear whether there would be five or five hundred races, according to a racial interpretation of clustering studies. The roughly continent-based clusters that race naturalists appeal to disappear when more genotypes are added and a finer grain of analysis is attempted. For example, Sarah Tishkoff and colleagues,[51] using worldwide samples from the HGDP-CEPH Human Genome Diversity Cell Line Panel and adding extra genotypes from other populations (mainly African and Indian), set STRUCTURE to find 14 clusters. More than half of them were found within Africa. The second condition fails in the human case.

Now to the third condition—that clusters should not be too clinal. It is important to note that clustered and clinal representations are not mutually exclusive.[52] Indeed, we should expect some discontinuities to arise from barriers to dispersal, such as oceans, mountains, and deserts. Yet such discontinuities are slight, and should not be overemphasized. As Handley et al. show,

> >75% of the total variance of pairwise F_{ST} can be captured by geographic distance alone. Adding information on genetic clusters to this model captures only an extra ~2% of the variance.[53]

Human genetic diversity and population structure is predominantly clinal. The third condition fails in the human case.

The fourth condition, recall, is that there should be a large jump of genetic difference between clusters. As I have already quoted Rosenberg and colleagues stating, the clusters they found were formed by "small discontinuous jumps in genetic distance."[54] What separates continental groups is not a large genetic leap, but a small jump, caused by geographical barriers to dispersal, long since crossed. The fourth condition fails in the human case.

There is nothing about clustering studies which suggests that clustering (discreet grouping) is itself appropriate as a general representation of the genetic structure of human populations. These studies, when properly understood, do not pose any threat to social constructionism about race. The clusters race naturalists appeal to do not represent subspecies. Traditional racial categories reflect human history and human prejudice more than they reflect human genetics.

V. CONCLUSION

If there is to be a productive debate between social constructionists and race naturalists, their respective positions will need to be clearly defined. I have stressed that social constructionism is not the view that racialized groups have no biological correlates. The point of social constructionism, it seems needless to say, is not that we all have the same skin color, the same shaped noses, and exactly the same genetic composition. It is that there are socio-cultural and historical reasons for our racial categories, and that our biological differences map poorly onto so-called "racial" groups.

It is less clear how to charitably characterize the claims of the new race naturalists. Is it their view that "race" has some genetic correlate, or do they defend a stronger claim? If racial naturalism is to challenge social constructionism, the latter needs to be true. However, it is not clear that this is the case. For Edwards, racial naturalism stands or falls on the issue of "stable classification."[55] But amateur ornithologists make stable classifications between birds which taxonomists do not separate into subspecies. Race naturalists need to tell us why "racial" classifications pick out subspecies, rather than more superficial groupings.

Even Sesardic, who makes the strongest claims amongst the race naturalists I have discussed, seems at times to endorse a rather weak view. For example, when referring to a clustering study by Tang et al.[56] he writes,

> Questions can be (and have been) raised about whether the same outcome would be obtained for other racial categories, or with a sample of people with more mixed ancestries, or on a more fine-grained scale, etc. My point is merely that in view of these new studies it becomes harder to accept the widespread but often unsubstantiated claim about the biological meaninglessness of race.[57]

The problems Sesardic points to here seriously undermine racial naturalism. He states that his aim is merely to show that race is not "biologically meaningless," but it is unclear what his target is—the idea that "race" has no biological correlate? No significant contributor to the debate defends this view. Sesardic attributes the claim that race is "biologically meaningless" to Robert Schwartz,[58] who indeed uses this phrase. The article Sesardic cites, however, is full of examples of how geographical ancestry is medically relevant. Schwartz argues that race is biologically meaningless only in the sense that it is too imprecise as a medical category, and he makes a convincing case against naïve racial profiling in medical research.[59]

We can see that race naturalists and social constructionists are talking past each other. When social constructionists say that race is "biologically meaningless," they mean that race is not a valid biological category. Race naturalists, however, interpret such phrases to mean that "race" has no biological correlate, a position which is clearly mistaken. As this is a mischaracterization of social constructionism, and one that can be easily avoided, the debate can move forward. The argument that "race" is not biologically meaningless—in that it has some biological correlate—misses its target, because social constructionists take this as a given. Race naturalists need to support a stronger position. Racial naturalism is not the view that racialized groups can be clustered genetically; it is the view that humans can be nonarbitrarily divided into subspecies. Race naturalists should either make the case for this view or abandon it.

Clustering studies, when properly understood, are entirely consistent with social constructionism. The continental clusters that race naturalists appeal to do not represent subspecies, in that they (a) vary greatly in the amount of genetic diversity and structure that they house, (b) are arbitrary in their number, (c) are not internally homogenous but clinal, and (d) are separated by only small jumps in genetic distance. When genetic diversity and population structure within a species are clinal, subspecies divisions are necessarily arbitrary. Human genetic diversity and population structure are almost entirely clinal, but there are some shallow cuts that allow clusters to be formed. These "scratches," which are located between geographic barriers that have slowed human migration, do not make human subspecies divisions any less arbitrary, because they are themselves arbitrarily located with regards to our genetic diversity and population structure. These clusters have more of a geographic significance than they have a biological significance. The fact that genetic clusters have been interpreted racially only supports social constructionism, in that it shows how readily we invest superficial biological difference with racial meaning.

Adam Hochman

University of Sydney

---

[1] The UNESCO Statements had this effect despite the fact that they rejected racist science, but not the concept of race, which was reconceptualized in population genetic terms. For both the 1950 and 1951 statements, see *The Race Concept: Results of an Inquiry* (Paris: UNESCO, 1952).

[2] Richard Lewontin, "The Apportionment of Human Diversity," in Theodosius Dobzhansky, Max K. Hecht, and William C. Steere, eds., *Evolutionary Biology, Volume 6* (New York: Appleton-Century-Crofts, 1972), pp. 381–98.

[3] Jeffry B. Mitton, "Genetic Differentiation of Races of Man as Judged by Single-Locus and Multilocus Analyses," *The American Naturalist*, CXI, 978 (March–April 1977): 203–12.

[4] A. W. F. Edwards, "Human Genetic Diversity: Lewontin's Fallacy," *Bioessays*, XXV, 8 (August 2003): 798–801.

[5] This was pointed out by Neven Sesardic in his "Race: A Social Destruction of a Biological Concept," *Biology and Philosophy*, XXV, 2 (March 2010): 143–62.

[6] Noah A. Rosenberg et al., "Genetic Structure of Human Populations," *Science*, CCXCVIII, 5602 (Dec. 20, 2002): 2381–85.

[7] Sesardic, *op. cit.*

[8] Henry Harpending, review of *Taboo: Why Black Athletes Dominate Sports and Why We're Afraid to Talk About It*, by Jon Entine, *Population and Environment*, XXII, 1 (September 2000): 83–89.

---

[9] Robin O. Andreasen, "The Cladistic Race Concept: A Defense," *Biology and Philosophy*, XIX, 3 (June 2004): 425–42.

[10] Philip Kitcher, "Race, Ethnicity, Biology, Culture," in Leonard Harris, ed., *Racism* (Amherst, NY: Prometheus, 1999), pp. 87–117. Kitcher now rejects this account of race in favor of a pragmatist approach. See his "Does 'Race' Have a Future?" *Philosophy and Public Affairs*, XXXV, 4 (Autumn 2007): 293–317.

[11] Massimo Pigliucci and Jonathan Kaplan, "On the Concept of Biological Race and Its Applicability to Humans," *Philosophy of Science*, LXX, 5 (December 2003): 1161–72.

[12] Coined by Ian Hacking in "Why Race Still Matters," *Dædalus*, CXXXIV, 1 (Winter 2005): 102–16.

[13] Edward O. Wilson and William L. Brown, Jr., "The Subspecies Concept and Its Taxonomic Application," *Systematic Zoology*, II, 3 (September 1953): 97–111, at p. 100.

[14] Robert M. Zink, "The Role of Subspecies in Obscuring Avian Biological Diversity and Misleading Conservation Policy," *Proceedings of the Royal Society of London B: Biological Sciences*, CCLXXI, 1539 (Mar. 22, 2004): 561–64, at p. 561.

[15] See, for example, Kevin de Queiroz and David A. Good, "Phenetic Clustering in Biology: A Critique," *The Quarterly Review of Biology*, LXXII, 1 (March 1997): 3–30.

[16] Daniel G. Mulcahy, "Phylogeography and Species Boundaries of the Western North American Nightsnake (*Hypsiglena Torquata*): Revisiting the Subspecies Concept," *Molecular Phylogenetics and Evolution*, XLVI, 3 (March 2008): 1095–115, at p. 1095.

[17] Guido Barbujani, "Human Races: Classifying People vs Understanding Diversity," *Current Genomics*, VI, 4 (June 2005): 215–26, at p. 215.

[18] Sesardic (*op. cit.*), for instance, criticizes "pure" social constructionism.

---

[19] Lisa Gannett, "Questions Asked and Unasked: How by Worrying Less About the 'Really Real' Philosophers of Science Might Better Contribute to Debates about Genetics and Race," *Synthese*, CLXXVII, 3 (December 2010): 363–85, at p. 375.

[20] Kitcher, "Does 'Race' Have a Future?" p. 299.

[21] See Stephen Ousley, Richard Jantz, and Donna Freid, "Understanding Race and Human Variation: Why Forensic Anthropologists Are Good at Identifying Race," *American Journal of Physical Anthropology*, CXXXIX, 1 (May 2009): 68–76.

[22] Gannett, *op. cit.*, pp. 382–83.

[23] Lewontin, *op. cit.*, p. 397.

[24] See Barbujani et al., "An Apportionment of Human DNA Diversity," *Proceedings of the National Academy of Sciences of the United States of America*, XCIV, 9 (April 1997): 4516–19; L. B. Jorde et al., "The Distribution of Human Genetic Diversity: A Comparison Of Mitochondrial, Autosomal, and Y-Chromosome Data," *The American Journal of Human Genetics*, LXVI, 3 (March 2000): 979–88; Rosenberg et al., *op. cit.*

[25] See Neil Risch et al., "Categorization of Humans in Biomedical Research: Genes, Race and Disease," *Genome Biology*, III, 7 (2002): 1–12, at p. 3.

[26] Rosenberg et al., *op. cit.*

[27] See Barbujani, "Human Races: Classifying People vs Understanding Diversity."

[28] See Sewall Wright, *Evolution and the Genetics of Populations, Volume 4: Variability within and among Natural Populations* (Chicago: University Press, 1984); Alan R. Templeton, "Human Races: A Genetic and Evolutionary Perspective," *American Anthropologist*, C, 3 (September 1998): 632–50; Rick A. Kittles and Kenneth M. Weiss, "Race, Ancestry, and Genes:

Implications for Defining Disease Risk,” *Annual Review of Genomics and Human Genetics*, IV, 1 (September 2003): 33–67.

[29] See Templeton, *op. cit.*, and Kittles and Weiss, *op. cit*.

[30] Jonathan K. Pritchard, Matthew Stephens, and Peter Donnelly, “Inference of Population Structure Using Multilocus Genotype Data,” *Genetics*, CLV, 2 (June 2000): 945–59.

[31] See also Jorde and Stephen P. Wooding, “Genetic Variation, Classification and ‘Race’,” *Nature Genetics*, XXXVI, 11S (November 2004): S28–S33.

[32] Edwards, *op. cit.*, p. 799.

[33] *Ibid.*, p. 800.

[34] *Ibid.*, p. 801.

[35] See Rosenberg et al., “Clines, Clusters, and the Effect of Study Design on the Inference of Human Population Structure,” *PLoS Genetics*, I, 6 (December 2005): 660–71; Lori J. Lawson Handley et al., “Going the Distance: Human Population Genetics in a Clinal World,” *TRENDS in Genetics*, XXIII, 9 (September 2007): 432–39.

[36] Sesardic, *op. cit.*, p. 148.

[37] See Gannett, *op. cit.*, and Kitcher, “Does ‘Race’ Have a Future?” for arguments against this view.

[38] This is because transpiration of water is necessary for photosynthesis and these crops have different photosynthetic pathways.

[39] Rosenberg et al., “Clines, Clusters, and the Effect of Study Design on the Inference of Human Population Structure,” p. 669.

[40] Sesardic, *op. cit.*, p. 156.

---

[41] For example, lactose intolerance affects Europeans and Africans; the missing CCR5 allele that confers HIV resistance protects Asians as well as Europeans; and the hemoglobin S allele that confers malaria resistance is present in sub-Saharan Africans, southern Europeans, Indians, and peoples from the Middle East.

[42] Kitcher, "Does 'Race' Have a Future?"

[43] Armand Marie Leroi, "A Family Tree in Every Gene," *Journal of Genetics*, LXXXIV, 1 (April 2005): 3–6.

[44] *Ibid.*, p. 5.

[45] David Serre and Svante Pääbo, "Evidence for Gradients of Human Genetic Diversity Within and Among Continents," *Genome Research*, XIV, 9 (September 2004): 1679–85.

[46] Rosenberg et al., "Clines, Clusters, and the Effect of Study Design on the Inference of Human Population Structure."

[47] *Ibid.*, p. 661.

[48] *Ibid.*, p. 668.

[49] See Sarah A. Tishkoff and Scott M. Williams, "Genetic Analysis of African Populations: Human Evolution and Complex Disease," *Nature Reviews Genetics*, III, 8 (August 2002): 611–21.

[50] Ning Yu et al., "Larger Genetic Differences Within Africans Than Between Africans and Eurasians," *Genetics*, CLXI, 1 (May 2002): 269–74.

[51] Tishkoff et al., "The Genetic Structure and History of Africans and African Americans," *Science*, CCCXXIV, 5930 (May 22, 2009): 1035–44.

[52] See Handley et al., *op. cit.*

[53] *Ibid.*, p. 435.

---

[54] Rosenberg et al., “Clines, Clusters, and the Effect of Study Design on the Inference of Human Population Structure,” p. 668.

[55] Edwards, *op. cit.*, p. 799.

[56] Hua Tang et al., “Genetic Structure, Self-Identified Race/Ethnicity, and Confounding in Case-Control Association Studies,” *The American Journal of Human Genetics*, LXXVI, 2 (February 2005): 268–75.

[57] Sesardic, *op. cit.*, p. 154.

[58] Robert S. Schwartz, “Racial Profiling in Medical Research,” *New England Journal of Medicine*, CCCXLIV, 18 (May 3, 2001): 1392–93.

[59] For a detailed response to Sesardic’s attempt to revive naturalism see Adam Hochman, “Racial Discrimination: How Not to Do It,” *Studies in History and Philosophy of Biological and Biomedical Sciences* (2013) http://dx.doi.org/10.1016/j.shpsc.2013.03.003.