ABSTRACT

Title of dissertation: INNATENESS IN THE SCIENCES: SEPARATING NATURE, NURTURE, AND NATIVISM

Mark Engelbert, Doctor of Philosophy, 2015

Dissertation directed by: Professor Georges Rey
Department of Philosophy

Scientists across the life sciences routinely appeal to notions of "innate" or "genetic" traits to explain developmental phenomena, and the idea of "innate" differences among people has figured prominently in some explanations of observed social inequality. This dissertation is an analysis of these concepts, which proceeds in two parts. Part I explores various philosophical issues related to the use of *innateness* as an explanatory concept, while Part II examines controversial claims that genetic differences among racial groups account for observed social inequality. I argue throughout that much disagreement about innateness arises from innocuous differences in explanatory goals and interests among different scientific research programs. Nevertheless, some proponents of genetic racial differences rely on understandings of "genetic" traits that conflict with the moral commitments of a just society.

Part I begins with arguments for a *contextual* and *pragmatic* approach to scientific explanation: in order for an explanation to be a good one, it must cite causes that are *relevant* to our interests in the explanatory context. I then apply this framework to biology and psychology, showing how different contexts call for different interpretations of

innateness. I conclude Part I by responding to arguments that aim to establish a single meaning for "innate"/"genetic" across all explanatory contexts.

Part II examines the use of "innate" and "genetic" concepts in developmental biology and population genetics, and applies the lessons of this examination to debates about alleged racial differences in genes for intelligence. I show that "hereditarians," who argue for innate racial differences, employ an explanatory framework that abstracts away from substantial complexity in developmental interactions between genes and environments. While this framework is adequate for certain purposes, it is poorly suited to designing interventions capable of eliminating racial IQ differences and attendant social inequality. I propose an alternative, *mechanistic* framework that promotes understanding of developmental complexity and design of effective interventions. I argue that a full commitment to racial equality demands that we adopt this latter framework, and to the extent that hereditarians resist doing so, their work exhibits some racist tendencies.

INNATENESS IN THE SCIENCES: SEPARATING NATURE, NURTURE, AND NATIVISM

By

Mark Engelbert

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:

Professor Georges Rey, Chair
Professor Peter Carruthers
Professor Susan Dwyer
Professor Mathias Frisch
Professor Jeffrey Lidz

ACKNOWLEDGEMENTS

I would like to thank all the members of my committee. I am grateful to Georges Rey for sharing his boundless enthusiasm for philosophy, and for the many enjoyable hours spent talking through the topics of this dissertation—and many others—in his office and his home. I have always appreciated—if not fully understood—Georges' unrelenting faith in me.

Peter Carruthers read countless drafts and in nearly every one of our meetings, provided several key insights that helped turn my assortments of jumbled thoughts into (what I hope have turned out to be) coherent arguments.

I am thankful to Sue Dwyer for many helpful conversations and, especially, for serving as a constant reminder that philosophy can and should strive to make a difference to how human lives go.

A well-taught seminar on causation with Mathias Frisch in my second year of graduate school stimulated my interest in the topic, which has remained with me throughout the writing of this dissertation.

Jeff Lidz graciously agreed to serve on the committee on short notice and taught me a great deal of what I know about language acquisition; both are much appreciated.

My parents, brothers, and sister have all provided me with many forms of support and encouragement over the years, for which I am deeply grateful. In addition, I have benefitted immensely from good conversation and good friendship with many of my fellow graduate students. Special thanks are due to Chris Vogel, Lisa Leininger, Brendan Ritchie, Xuan Wang, Andrew Knoll, Bess Picciuto, and Evan Westra.

My greatest debt is to Lina Khan, who from the day I began writing has provided innumerable kind words, hugs, home-cooked meals, and swift-kicks-in-the-pants, along with an uncanny ability to know which was needed at any particular time. She has my deepest thanks and my warmest love.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: CAUSATION AND EXPLANATION

1.1 Introduction

In this chapter I lay out the general theoretical approach I will take to the notions of *causation* and *explanation* in the context of scientific inquiry, which will set the stage for more detailed discussion of explanation in biological development in Chapter 2. Insofar as innateness claims are commonly offered and understood as claims about *what caused* a trait's presence, and insofar as innateness claims are meant to be in some way *explanatory,* it is crucial that we clearly specify our theoretical commitments when it comes to these two notions.

Following Davidson (1967), among others, I distinguish *causation* as a metaphysical relation from *causal explanation*, the practice of providing causal information as a way of making sense of phenomena of interest, and I argue that our explanatory practices are underdetermined by the metaphysical causal structure of the events we seek to explain.[1] I will explore two approaches to causal explanation, broadly construed: one is grounded in a *counterfactual* (and specifically, *manipulationist*) interpretation of causation, and the other is grounded in a mechanistic analysis of the *production* of phenomena of interest.[2]

[1] As will be explained in Section 1.5 below, this idea of "causal explanation" is a fairly weak notion—much weaker than the more robust notion of causation itself.

[2] Note that there are differences among proponents of the mechanistic approach in how they view the relationship between mechanistic explanation and notions of causality, and some would perhaps object to counting mechanistic explanation as a species of causal explanation. However, I see no disadvantage to

I will argue for the utility of *both* of these types of explanation, in that either of them might be the most appropriate explanatory strategy in a given circumstance. Thus, I defend a *context-dependent* (and specifically *contrastive* and *pragmatic*) approach to causal explanation. I will, however, argue that particular features of specific lines of inquiry can impose constraints on the kinds of explanatory strategies that will be appropriate in that domain.

The accounts of causation and of causal explanation defended here are familiar ones from the philosophy of science. Thus my aim here is not to develop new approaches to these much-discussed philosophical issues. Rather, my primary goal is to explicate the contrastive/pragmatic account of causal explanation in such a way that it is consistent with common-sense causal intuitions and everyday explanation. Doing so will serve two purposes. First, showing how my account comports with everyday explanatory practice will render the rather weak metaphysical notion of causation I offer less unsatisfactory than it otherwise would be. More importantly for the purposes of this dissertation, however, the discussion here will set the stage for later exploration of the role that different explanatory interests play in debates about innateness.

I begin in Section 1.2 by making explicit some very broad aspects of the general attitude I will be taking towards scientific investigation and explanation, which will inform the subsequent discussion in this and later chapters. Section 1.3 describes the two approaches to causal explanation (namely, the *counterfactual/manipulationist* approach

---

counting mechanistic explanation as a type of causal explanation in the weak sense intended here—the provision of information about an event's causal history.

and the *productive/mechanistic* approach) we'll be considering, while Section 1.4 provides a more thorough treatment of the manipulability approach. I close in Section 1.5 by integrating these various explanatory strategies into a general contrastive/pragmatic approach to explanation. The following chapter applies these various approaches to biological development and the question of "genetic causation."

1.2 The aims and limits of science: science as cartography

1.2.1 A modest realism

It will help to begin by making explicit the general attitude towards scientific inquiry adopted in this chapter, and indeed, throughout this dissertation. At the heart of this attitude is the idea that science is something done *by* and *for us*. That is, I take the purpose of scientific inquiry to be the fulfillment of various human-centered purposes, encompassing both idle curiosity about the universe around us and our practical interests. While I do not deny that there is a mind-independent world, or that humans can and do gain knowledge about that world through scientific inquiry, I take the particular shape of human scientific practice—the questions it formulates, its procedures for answering them, and the answers it provides—to be a reflection of the kinds of creatures we are. Creatures with an entirely different set of needs and interests, even if they shared the general impulse to understand and manipulate the world around them, would likely go about it in very different ways. This position is, therefore, a rather modest form of realism.

This conception of science has been helpfully illuminated by way of an analogy between scientific inquiry and cartography.[3] Maps can be accurate representations of reality in that they can depict the relative sizes, shapes, and locations of various topographical features of an area. Yet maps are necessarily selective in which features of an area to depict and how to depict them; Lewis Carroll has artfully parodied the idea of a "perfect map": a map that recapitulates every feature of the world in perfect detail is simply the world itself. Creating maps thus requires us to select from among the many different possible ways of representing the features of an area, a process that will reflect the interests of the mapmaker and those who will be reading the map. The map will therefore represent those features of the landscape that we find interesting, notable, or practically useful, and the variability of our interests entails that different maps will be appropriate for different purposes and on different occasions.

My claim is that science works the same way: just as there are multiple distinct yet accurate ways to chart a landscape with a map, so there are multiple distinct yet accurate ways to describe the world through science. In conducting scientific inquiry, we will continually make decisions not only about which phenomena to study, but also about when science has adequately explained the phenomena that interest us. I am deeply skeptical that there is any way of characterizing how such decisions are—or should be—made that does not make a fundamental appeal to human concerns and interests (where, again, these

[3] Philip Kitcher (2001, Chapter 5) has given the most thorough and enlightening presentation of this analogy, though see also Garfinkel (1981, pp. 146–147). Hempel (1960) attributes the original comparison to Max Weber, though it is not clear whether or where it appears in Weber's writing.

concerns and interests encompass both our desire to control and manipulate the world around us as well as sheer curiosity).

The perspective adopted here stands in opposition to views which hold that there is exactly one correct and complete way to describe the world (as it exists independently of human experience), and that science is in the business of generating that one description. Since this is not a dissertation about varieties of scientific realism, I will not spend time here explicitly defending my modest realism from its more ambitious competitors (though see Kitcher, 2001, for a detailed defense of a view sufficiently similar to my own). This note is meant to flag and make explicit an admittedly contentious theoretical assumption that underlies a great deal of what follows. Some of the arguments offered below, however, will help make clear why I find this view compelling.

1.2.2 Explanation as understanding

My conception of science as a kind of human-centered activity extends to scientific explanation as well: explanation is done by and for us. This idea fits naturally with the view that the hallmark of good explanations is that they facilitate *understanding* (De Regt & Dieks, 2005; de Regt, 2009; Gopnik, 2000). As a result, I would argue that what makes something explanatory is that it *feels* explanatory *to us*. Explanations, therefore, are a kind of psychological entity, and what facts explain which other facts is contingent and contextual—contingent upon the sorts of creatures we are and the kinds of curiosity we exhibit. There is a particular phenomenology associated with understanding a puzzling

phenomenon, and the production of this phenomenological state in us is the hallmark of explanation.

The view expressed here might seem to conflict with the idea that explanatory relations exist "out in the world," a view sometimes called the *ontic* notion of explanation (Craver, 2007; Salmon, 1984, 1989). However, when understood properly these notions can be made compatible. Craver distinguishes between explanations as representations (texts, pictures, models, and so forth)—the kind of thing that facilitates understanding and so satisfies my own criteria for explanatory success—and *ontic* or *objective* explanations:

> Other times, the term explanation refers to an objective portion of the causal structure of the world, to the set of factors that bring about or sustain a phenomenon (call them objective explanations). What explains the accident? The ice on the road, the whiskey, the argument, the tears, and the severed brake cables. . .Objective explanations are not texts; they are full-bodied things. They are facts, not representations. They are the kinds of things that are discovered and described. There is no question of objective explanations being "right" or "wrong," or "good" or "bad." They just are. (2007, p. 27)

As we can see, Craver's notion of *objective explanations* simply entails that there is an objective causal structure to the world, and that constructing adequate "explanatory texts" depends (at least sometimes) on discovering this structure. I do not deny any of this. I merely emphasize, in addition, that just as we must selectively represent certain features of a landscape in our maps, so must we selectively represent certain features of the causal structure of the world in our explanatory texts. The notion of explanation as understanding, which I endorse here, simply expresses a particular view of why representations of certain causal relations ("objective explanations," if you will) in certain ways (and not others)

constitute satisfactory explanatory texts: satisfactory explanatory texts are those that facilitate understanding.[4] Since for the most part we will be concerned with what Craver calls "explanatory texts," I will confine use of the term "explanation" to these representations, understood as descriptions of the world that facilitate understanding and serve to make puzzling phenomena intelligible.

Of course, much more needs to be said in terms of fleshing out the idea of explanation as understanding. Most notably, there will have to be some further constraints on good explanations beyond simply inducing the phenomenology of understanding in someone, somewhere. Clearly the phenomenology of understanding can be misleading: we might experience such a phenomenology even in response to a proposed explanation that, were we were paying closer attention, we would recognize as inadequate, and we should not be bound to accept the validity of a creationist's explanation for the presence of fossils simply because "God put them there to test our faith" *feels* like a good explanation to many people. But, as with the realism/anti-realism debates touched on above, a full consideration of the relevant issues would take us too far afield. Therefore, I trust that the notion of explanation as understanding will be sufficiently intuitive to allow us to proceed within that framework.

[4]Note that I do not mean to imply that Craver would deny that our explanatory interests must be taken into account when moving from objective explanations to explanatory texts—indeed, Craver endorses this thesis himself in emphasizing the contrastive nature of causal relations.

1.3 Two approaches to causal explanation

1.3.1 Distinguishing dependence and production

We are apt to apply the word *cause* in a wide variety of circumstances, and much philosophical work has been dedicated to systematizing our causal judgments, such that there are now a number of theories of what causation *is*.[5] Not surprisingly, however, for each theory that has been offered as providing necessary and sufficient conditions for causal relations to hold, counterexamples have emerged, such that each theory does a good job of accounting for our causal intuitions and our explanatory practices in some range of instances, but fails elsewhere. This is a problem if one is committed to the view that we need a complete and universal analysis of causation in order to say how we can engage in causal explanation at all. However, in keeping with the pragmatic and pluralistic approach to explanation pursued here, I prefer to think that there is nothing wrong with utilizing certain notions of causation in some circumstances, and other notions in other circumstances. I will develop and defend this thought—that different causal notions are explanatorily useful in different contexts—below, but for now let us continue exploring some of the basic features that seem to be common to our many explanatory practices.

As noted above, we are apt to apply the word *cause* quite widely, and there are a number of features of events—and the way we conceptualize the relations between events—that seem central to our sense of what it means for two events to be causally

[5]I should note that the ensuing discussion concerns *token* causation—that is, causation as a relation between event tokens ("Geraldine's heart attack caused her death")—rather than *type* causation ("Smoking causes cancer)."

related. For present purposes, I will focus on two of these features: on the one hand, the idea of *counterfactual dependence* seems fundamental to our understanding of causation (e.g., had the CEOs not signed the contract, the corporate merger would not have occurred; hence the signing caused the merger), and on the other hand, paradigm cases of causation typically involve *spatiotemporally continuous* processes linking the events, such that the cause *produces* the effect (e.g., the many causal interactions on a billiard table).[6]

Often enough, *dependence* and *production* occur together, such that, for events *c* and *e*, *e* counterfactually depends on *c*, and, in addition, there is a spatiotemporally continuous chain from *c* to *e* (consider again the billiard table, with *c* as the cue striking the cue ball and *e* as the eight ball falling into the corner pocket). However, it is easy to show that dependence and production come apart, and that treating either as a necessary and sufficient condition for a causal relation to hold between two events forces us to embrace bizarre results[7]. *Dependence* tends to fail in cases that involve "preemption": Billy and Suzy each throw a rock at a glass bottle. Suzy's rock gets there first and the bottle shatters, but it is not the case that the shattering of the bottle *depends* on Suzie throwing the rock—for if she had not thrown, or had she missed, Billy's rock would still have smashed the bottle. And yet we are strongly inclined to say that Suzie's throw, not Billy's,, caused the bottle to shatter (i.e., Suzy's throw *produces* the shattering event).[8]

---

[6] By *spatiotemporally continuous process* I mean something like Salmon's "mark transmission" or Dowe's "conservation of conserved quantities," which are discussed in Chapter 2.

[7] Much of the discussion here follows Hall (2004).

[8] The presentation of this example is meant to shed light on the general nature of the problem for supposing *dependence* is necessary and sufficient for causation, and so my description of it overlooks some nuances. For instance, the causal analysis of cases like this will of course depend on how events are individuated—i.e., one might argue that the shattering resulting from Suzy's throw is not the same event as the shattering resulting from Billy's. However, the example can be formulated (see Hall, 2004) in such a way that one and the same event is overdetermined by two distinct events, such that the outcome depends on neither.

Production, on the other hand, struggles with cases of "double prevention," or an event *c*'s prevention of some other event *b* which, had it occurred, would have prevented event *e.* Here imagine that Billy and Suzy are fighter pilots, and that Suzy successfully carries out a bombing mission, while Billy shoots down the enemy who, were it not for Billy's rescue, would have shot down Suzy. The bombing thus depends counterfactually on Billy's firing upon Enemy, and many would surely find it intuitive to say that Billy's firing was a cause of the bombing.[9] But suppose the dogfight between Billy and Enemy takes place hundreds of miles from Suzy and the bombing, such that there is no spatiotemporally continuous series that connects the dogfight and the bombing (indeed, Suzy might not even be aware of Billy's actions, or even of Enemy's existence). Billy's action (intuitively) helps cause the bombing, but doesn't produce it. Thus, just as it seems there can be causation without *dependence,* so too there can be causation without *production*.

Thus, in what follows I will explore two explanatory strategies, one (broadly) grounded in notions of *counterfactual dependence* (the manipulability theory of causation), and the other (broadly) grounded in attention to *productive* relations (the mechanistic approach to explanation in biology, to be presented in Chapter 2).

[9]There is a general question here about whether *absences* and *omissions* (e.g., Enemy's *failing* to shoot down Suzy) can be causes. Philosophers have debated the issue, but it doesn't matter for our purposes here, since the debate is really about whether absences and omissions have "genuine causal powers" and can be genuine constituents of the causal structure of the world; it is universally agreed that absences and omissions are *explanatorily relevant*. Our concern here is with the latter. (For discussion, see Barros, 2013; Beebee, 2004; Craver, 2007, Chapter 3).

1.4 The manipulability theory of causal explanation

The manipulability theory relies substantially on Judea Pearl's (2000) work, but has been given its most thorough philosophical defense by Jim Woodward (2003).[10] According to Woodward, causation is to be understood in terms of *manipulability*: at its core, the theory states that a causal relation holds between two variables[11], *x* and *y*, iff a change in one results in a change in the other. "Changing" a variable amounts to performing a (usually hypothetical) intervention that sets one of the variables to a particular value, and looking for a change in other variables (in effect asking: *what if things had been different?*)

I believe Woodward's view has a number of virtues that make it a very useful analysis of causation in many circumstances. Needless to say, the question of what is the *best overall* philosophical analysis of causation is a matter of some contention. While I cannot here provide a full defense of the manipulability theory against all of the objections that have been raised against it (e.g., Hiddleston, 2005; Strevens, 2007), I will briefly highlight a few of its advantages. First, note that Woodward's is not a *reductive* account of causation, and as such it makes no attempt to remove causal notions (e.g., a change in one variable *results in* a change in another) from the analysis of other causal notions. Some may find this non-reductivity unsatisfying; however, in eschewing issues of reduction, Woodward's approach has the quite useful feature of allowing us to see how scientific inquiry can generate robust systems of explanation even *without* a fully reductive analysis of causation. A second advantage is closely related to the first: as will be outlined below,

[10]All references to Woodward are to Woodward (2003)

[11]Strictly speaking, Woodward conceives of *properties* as the causal relata, but usually speaks of relations among "variables." I follow this usage.

Woodward's manipulability theory eliminates the need to appeal to *laws* for explanatory power (relying instead on the notion of *invariance*—to be explained below), thus providing an elegant account of how generalizations can be explanatory even if they fail to be laws (the latter notion being, of course, a thorn in many a philosopher's side).

As an intuitive example, consider the following: suppose that the variables *x* and *y* correspond to my having a disposable $50 and my going to a Jethro Tull concert, respectively. Each of the variables has a range of possible values. In the present case, each variable is binary: I either have $50 or I don't, and I either attend the concert or I don't. Imagine that the value of both variables is "0" (i.e., I don't have the money and don't attend the concert); the manipulability theory says that if there is a causal relationship between *x* and *y*, we should be able to change the value of one of the variables by changing the value of the other. So we could perform an intervention on *x* in order to set its value to "1" (i.e., someone gives me $50). If this manipulation leads to a change in the value of *y* (I go to the concert), then there is a causal relationship between *x* and *y*; if the value of *y* remains constant even after the intervention on *x*, then no such relationship exists.[12] (Note also that this intervention has revealed only that the causal arrow goes in one way—*x* is a cause of *y*—but we could easily consider the opposite intervention to see if *y* is a cause of *x*.) Of course, there is room for considerably greater complexity: relations may hold among more than two variables, variables may take any number of values, discrete or continuous,

[12]As will be explained more fully below, it must be assumed that all other causal factors are held fixed, i.e., that the generosity of the person who gives me $50 does not inspire me to do something magnanimous with the money, thus altering my desire to spend $50 on a concert ticket.

manipulations of one variable may lead to an increase *or* decrease in the value of a variable it causally influences, and so on.

There are a great many technicalities in Woodward's manipulability account, but here I review only two of them that will prove important later. Firstly, note that in Woodward-style causal systems, an intervention on a particular variable may be made (and indeed, *should* be made) while holding the values of all other *independent* variables constant. That is, we can manipulate *x* and *y* independently only if *x* is not a cause of *y* and *y* is not a cause of *x*. But importantly, when there are multiple variables linked in a causal chain, an intervention on any variable affects *all* the variables "causally downstream" of the manipulated one. To see this, consider the causal system in Figure 1.1:

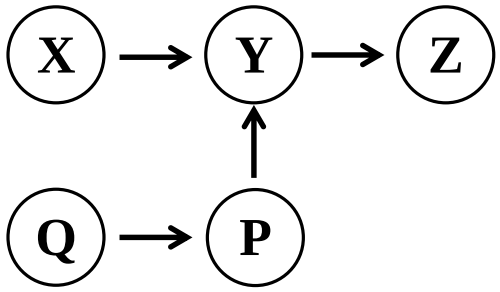

Figure 1.1: Woodward-style causal model

As we can see, there is only one intervention we can make (on Z) that would affect no other variables in the system, and only one other intervention (on Y) that would only affect *one* other variable. Any other single manipulation we might make (on X, Q, or P) will change the values of more than one variable. We cannot change the value of X or Q, for example, without changing the value of Y. Thus, it is impossible to consider the causal effect of X on Z *independent* of Y, for Y causally depends on X. In order to assess the relationship between X and Z, we must consider the effects of X on Y.

Secondly, causal relationships do not necessarily hold under all conditions—that is, whether or not a causal relationship holds between two variables may depend on the values of *other* variables. Woodward's term for this is *invariance*, and causal relations can be more or less invariant with respect to the range of conditions under which they hold. The notion of invariance proves particularly useful in meeting one of the main challenges to causal theories like manipulability. Because the manipulability theory maintains that causal relations between variables hold in virtue of what *would* happen to one variable if a manipulation *were to be* performed on another, it belongs to the class of *counterfactual* theories of causation (to be contrasted with, for example, causal process theories like those of Salmon, 1984 and Dowe, 2000). As a counterfactual account of causation, Woodward's manipulability theory faces a similar *prima facie* problem to that which plagues other such accounts: counterfactual accounts do not provide an obvious way of singling out one cause (i.e., what we might wish to call *the* cause) among the many things upon which a state of affairs might counterfactually depend (cf. Lewis, 1973). For instance, if I strike a match and it ignites, then my striking the match caused the ignition, because the ignition counterfactually depends on the striking (alternatively: an intervention on the striking variable results in a difference in the ignition variable). But by parity of reasoning, the presence of oxygen in the room caused the ignition as well, and, to say the least, it is not obvious how to tweak such theories so as to yield the "right" result of saying that the striking of the match occupies a special causal role that other causes do not. (I will return to this difficulty later on.)

This is where the notion of *invariance* comes in. Again, causal relationships can hold more or less invariantly, and we can see how this might relieve the tension stemming from our recognition that oxygen in the room was a cause of the match's ignition and our simultaneous sense that the match was a *special* kind of cause (which might be labeled in a variety of ways: "triggering cause," "critical cause," "preponderant cause," and so forth)—though the tension is perhaps not completely eliminated; see Section 1.6. The presence of the oxygen, we think, can be "held constant," and when we do so, the causal relationship between the match and the ignition is highly invariant (there are some values of some variables—wetness of the match, for example—that will disrupt that relationship, but these will be comparatively few). But in order to identify the oxygen as a "special" cause in this sense, we must hold the striking of the match constant. This amounts to saying that oxygen is a "special" cause of ignitions only when we can assume that matches are routinely struck, but the presence of oxygen varies. While there exist such contexts (e.g., employees at a match-production company that tests the hardness of match heads in a room that is supposed to be free of oxygen might well cite the presence of oxygen as the cause of an ignition), they are comparatively rare.[13] Thus, the invariance condition allows us to say that under "normal" conditions, the match causes the flame in a way that the oxygen does not.

This discussion of the counterfactual nature of the manipulability account, with its concomitant issues of invariance and "normal" conditions nudges us from the realm of causation to that of *causal explanation*, which I address in the next section.

[13] The example is from Garfinkel (1981)

1.5 Explanation: contrastive and pragmatic

1.5.1 Contrastivity

I will be assuming, in what follows, that a very large chunk of scientific explanation consists of *causal* explanation. I do not mean to claim that explanations can only be scientific if they are causal (much less that even non-scientific forms of explanation must be causal in order to be legitimate), but it would appear that most workaday scientific explanation is in some way aimed at describing the causal structure of the world. Importantly, philosophers of science routinely draw a distinction between *causation* and *causal explanation* (Beebee, 2004; Davidson, 1967; Woodward, 2003). *Causation* here refers to an ontological relationship that holds between things in the world (where "things" are, on various accounts, *properties*, *events*, *states of affairs*, or *variables*—I continue to use the latter term). As noted above, any counterfactual account of causation (like the manipulability theory discussed above), appears committed to the idea that there are a great many such relationships that hold between any particular variable and various antecedent variables. That is, there are many causes of the value of any particular variable *e*, because there are many other variables $c_1 \ldots c_n$ that, had they been different (i.e., had we intervened on them), would have resulted in a different value for *e*.[14] *Causal explanation*, then, refers to our practice of citing some, but not others, of these many causes when answering such

[14]In the next section I will take up the question of whether counterfactual theories are committed to counting *every* variable upon which an outcome counterfactually depends as a cause of that outcome—for now I assume that it does, as this assumption will allow us to gauge the effectiveness of the contrastive/pragmatic approach to explanation in helping us impose order on an extremely unruly conglomeration of causal relations.

explanatory questions as "Why does *e* have the value that it does?" or "What explains the value of *e*?" Unlike *causation*, which is, again, held to be an objective metaphysical relation that holds in the world, *causal explanation* appears to crucially involve pragmatic—rather than merely theoretical—components, in the sense that the choice of some causal factors over others reflects our own interests and the explanatory contexts we inhabit, rather than any objective feature of the causal structure of the world.

The notion of causal explanation, as distinct from causation, can be helpfully illuminated using cases involving omissions: suppose a lazy gardener fails to water his employer's plants, and the plants die. On a manipulationist approach to causation, the gardener's failure to water the plants counts as a cause of the plants' death iff there is some intervention on the variable corresponding to whether or not the gardener watered the plants (a variable that, let's suppose, takes one of two values: 0 = no, 1 = yes) that would have changed the outcome (i.e., resulted in the plants continuing to live). Since this condition is, by hypothesis, met (i.e., had the gardener watered the plants, they would have lived), the gardener's failure to water the plants is a cause of their dying. But, the familiar story continues, there is also an intervention on the variable that represents whether *I* watered the plants that would have changed the outcome, and the same is true for you, your barber, and Ronald Reagan's reanimated corpse (i.e., had any of us watered the plants, the plants would have lived). Hence, all of our failures to water those plants count as causes of the plants' death.

This "causal promiscuity" is generally thought to be unavoidable on any account that treats counterfactual dependence as central to causation. But although such theories

commit us to the existence of infinitely many bizarre causal relations, there is no reason our causal-explanatory practices need be contaminated by these rather disorderly elements. For our explanatory practices will reflect our practical interests, our judgments of relevance, and even our *normative* commitments (regarding, for example, who is *responsible* for keeping the plants alive).[15] Hence, in the vast majority of cases, citing the gardener's failure to water the plants will be regarded as an informative and appropriate bit of causal information, while pointing to Zombie Reagan's negligence will not.

On this account, then, causal explanation is a thoroughly pragmatic affair. There is, of course, some debate regarding just to what extent our explanatory practices are determined by such pragmatic considerations and how much they are constrained by the real causal structure of the world (again, this will be addressed in the next section). Nevertheless it is widely agreed that somewhere along the line our goals and interests will work their way into our practices of seeking and identifying causal relationships.

As a number of philosophers (e.g., Dretske, 1972; Garfinkel, 1981; Van Fraassen, 1980) have pointed out, one of the chief ways in which we often go about highlighting the causal factors that *are* relevant to us is by making explanations (and explanatory questions) *contrastive*. Indeed, some (Garfinkel, 1981; Northcott, 2008; Van Fraassen, 1980) have argued, persuasively in my view, that explanation is *fundamentally* contrastive, even if it doesn't always appear to be.[16] To say that explanations are fundamentally contrastive

[15] This last point will be especially important in Part II of the dissertation.

[16] Northcott (2008) actually argues that not only is *causal explanation* contrastive, but so is *causation itself;* i.e., causation is not a 2-place relation but a 4-place one, with all causal relations having the form [*c* rather than *c´* causes *e* rather than *e´*] (cf. Craver, 2007). This is an intriguing suggestion but not germane to our discussion here.

amounts to saying that, contrary to the way they are sometimes phrased, explanations do not offer an account of why some phenomenon occurred *simpliciter*, but rather why some outcome *rather than any of some relevant class of alternative outcomes* was realized.[17] (Note the term *relevant:* it is a basic feature of contrastive approaches to explanations that the selection of a contrast class will reflect our interests and our judgments of relevance.) On a contrastive notion of explanation, a key step in any explanatory project is determining what comprises the *relevant* class of alternative outcomes (often called the "contrast class" or "contrast space"), for the selection of a contrast class will constrain the sorts of explanations that can be offered to the explanatory question, in ways described below.

Garfinkel (1981) provides a helpful discussion of the relation between contrast classes and explanation, taking as a starting point a humorous (and apocryphal) anecdote involving the bank robber Willie Sutton. A priest visiting Sutton in prison is said to have asked Sutton why he robbed banks, to which Sutton supposedly replied, "Because that's where the money is." The disconnect between the information requested by the priest and that provided by Sutton illuminates the role contrastive focus plays in explanation, as well as the importance of stating contrast classes explicitly wherever possible when framing and

[17] There are in fact two ways of construing this "contrastivity condition." In what follows I do not distinguish between the two construals, since the difference appears to be merely a verbal issue regarding the use of the term "explanandum." But for completeness' sake: the contrastivity condition can be construed either as a constraint on how explananda must be stated, or as a constraint on how explanatory questions must be framed. That is, different ways of accounting for the contrastivity condition in a theory of explanation may or may not subsume the contrastive focus under specification of the explanandum. One might insist, then, that explanations can only be offered for explananda that are expressed as disjunctions of the observed and alternative outcomes (i.e., questions must be of the form, "why 'x rather than y, or z, or w,. . .'?"), or one might prefer to allow that the explananda of scientific questions themselves may be stated atomistically, but insist that *explanatory questions*, in order to be complete, must contain a contrastive clause that includes the relevant alternatives (i.e., "why 'x' rather than 'y' or 'z' or 'w'. . .?" Again, the difference appears to be merely verbal.

communicating explanatory questions. For the priest, the relevant alternative is *robbing nothing*, while for Sutton it is *robbing something other than banks*. Notice that Sutton's answer is not necessarily *wrong* as a response to the question as stated, because it is, presumably, true that if banks did not contain large quantities of money, Sutton would not rob them. Rather, Sutton's response simply reflects a different explanatory focus from the one the priest is assuming. (An alternative way to put this is that Sutton and the priest diverge on the issue of what is contained in the class of relevant contrasts for the state of affairs *Sutton robs banks*). Thus, Sutton's explanation would be a perfectly good one if he were asked the same question by a fellow thief with a penchant for robbing hot dog stands.

Closely related to the contrast classes that figure prominently in the present notion of explanation is another widely-discussed phenomenon, that of *background assumptions* (what van Fraassen, 1980, and Garfinkel, 1981, call "presuppositions"). Given the ambiguity of "why?" questions (here conceived as requests for explanations; van Fraassen, 1980) revealed by the Sutton story, what counts as an acceptable answer to such a question will be constrained by what we take to be "given" in the explanatory context, or as it's sometimes expressed, what we take to be the *background conditions* for our explanatory endeavors, or what we think we can *hold fixed* for our explanatory purposes. Such conditions can be (and sometimes are) expressed as "given that…" clauses. For example, the question Willie Sutton cheekily attributes to the priest can be stated: "*Given that* you are going to rob something, why do you rob banks?" And, had the priest been aware of the kind of wiseacre he was dealing with, he might have phrased his question thusly: "*Given that* one ought not rob anything at all, why do you rob banks?"

This highlights an important feature of background conditions (as expressed by "given that…" clauses), namely, that they determine what the contrast class of relevant alternative outcomes will be taken to be. The "*given that…*" clause of Sutton's reading of the priest's question (as formulated explicitly above) indicates that for Sutton, the relevant alternatives to his robbing banks all involve him robbing other things. By contrast, the implicit "*given that…*" clause of the priest's question indicates that the alternative outcomes within *his* explanatory framework involve Sutton robbing nothing. Whether they are made explicit or not, the background assumptions that we bring to an explanatory project determine what the contrast space of alternative outcomes is, which in turn determines exactly what explanatory question we need to ask and what will count as a good explanation for the phenomenon of interest.

I close this exposition of this contrastive theory of explanation with one caveat: while explicitly defining contrast classes for phenomena we wish to explain can go *some* of the way towards guiding our causal-explanatory practices through the thicket of counterfactual dependencies, they cannot go *all* the way. That is, even once we have identified what we take to be the relevant alternatives to our explanandum, we will *still* face pragmatic choices in deciding which of the many variables that, had they taken different values, would have yielded one of the alternatives in our contrast space. Van Fraassen (1980) offers the example of the question "Why is this conductor warped?" Possible contrast classes for the state of affairs assumed by the question (*this conductor is warped*) include conductors other than this one being warped (e.g., "Why is *this* conductor warped, rather than *that* one?') or this conductor retaining its shape.

However, even once it is made explicit what is the relevant range of alternatives, there is still considerable room for pragmatic considerations to inform what sorts of answers are appropriate and informative. For suppose it is stated that the contrast class for the question is the conductor retaining its shape. Depending on the context, the asker of the question might still be seeking any of a number of different types of information: perhaps she seeks information about who is *responsible* for the state of affairs (in which case "The technician did not employ the proper procedure" might count as a good explanation) or instead about the physical/mechanical conditions involved (here "There was moisture on the switches" might be the best explanation available).

The upshot of all this is that pragmatic considerations infuse our explanatory practices at multiple junctures. This, in turn, means that there will be much work to be done in teasing out just what contextual and pragmatic considerations are in the background of any particular line of empirical investigation (and whether these same contextual features are shared among different theorists in a field). And in particular, as regards the subject of this dissertation, we will want to probe the features of the explanatory contexts in which claims about *innateness* or about *genetic causation* are made. I will begin such an analysis shortly in Chapter 2, where I examine some of the causal facts of biological development (and in particular the causal role of genes), but first I will comment on the question—alluded to earlier—of whether we ought to endorse the view that there are *no* metaphysical distinctions to be made among the many causes of a particular event. That is can we really not draw metaphysical distinctions among the infinitely many remote and bizarre

conditions and omissions upon which an event counterfactually depends? Is it really just pragmatics all the way down?

1.6 Are causal relations really *that* unconstrained?

I have said that the practice of causal explanation—of citing certain states of affairs (and not others) as causes or explanations of other states of affairs—is an ineluctably pragmatic enterprise. That is, if we want to know which of the many things that count (metaphysically speaking) as causes of a particular event or state of affairs is properly regarded as *the* cause, the answer is not provided by the metaphysical relationships between the different events. Rather, my account holds, selecting one cause over others is a choice that depends on our interests, goals, normative commitments, and so on. Even if this is right, however, we might still ask the question: is there *any* role for metaphysics—rather than just pragmatics—in determining which causes are "better" candidates for the role of *cause* than other candidates?

We can distinguish a range of positions as to the answer to this question: two extremes and a variety of intermediate stances. At one extreme there is the claim that for every event, there is exactly one cause that occupies a metaphysically privileged position—it is, "from the point of view of the universe," *the* cause. This position I take to be discredited by the foregoing discussion of Woodward's manipulability approach to causation (a counterfactual account): if causation is a metaphysical relationship involving counterfactual dependence, we seem to be stuck with some degree of causal promiscuity, thus opening the door for pragmatic considerations. (Consider: given the options of

explaining the demise of the flowers in terms of the gardener's negligence or the unseasonably warm weather, would we really want to commit ourselves to one of these being the *real* cause?) At the other extreme, however, we have the view that every zany, bizarre counterfactual upon which a given event depends counts *equally* as a cause of that event. Thus, from the point of view of the universe, the destruction of the cathedral in Dresden during the Allied bombing was every bit as much caused by the *absence* of flying kangaroos snatching the Allied bombs from midair as by the Allied commanders' order to drop bombs on Dresden. According to this view, our decision to cite the command—rather than the absence of the kangaroos—as the cause of the destruction is merely a reflection of our interests and our subjective judgments of relevance.

Occupying the space between these two extremes are a number of possibilities. One might suggest, for example, that the factors upon which a given event counterfactually depends partition into two classes: those that are metaphysically relevant and those that are not. This suggestion would likely hold, for example, that events like the order to conduct the bombing and the pilots' pulling the triggers fall into the category of relevant causes, while the absence of flying kangaroos does not. This view could then concede that *within* the class of metaphysically relevant causes, our pragmatic concerns determine which we choose to focus our attention on and/or select as *the* cause (when, for example, we are deciding who is a potential target of an indictment for war crimes). But, this view would hold, it would be a mistake for us to direct our focus to the kangaroos, for a belief that the presence or absence of flying kangaroos is causally relevant to the destruction of the cathedral is simply *false*.

Alternatively, one might suggest that the metaphysics of causation allow for assigning *degrees* of relevance, such that some counterfactual dependencies are *more* relevant than others in the causal history of an event, and so constitute objectively (but only *ceteris paribus*—see below) *better* candidates for causal explanation. This view could stand on its own, holding that from the point of view of the universe, *all* counterfactual dependencies are causes (including the kangaroos), but that some are only infinitesimally relevant. Or it could be combined with the bifurcation view, such that some dependencies are simply outside the range of metaphysically relevant causes, but *within* the relevant range there are gradations of relevance, with some causes being prime candidates for causal explanation and some just barely meeting the threshold for relevance. Either way, such a view would again countenance a role for pragmatic considerations: causal metaphysics provide certain constraints and/or guidelines for causal explanation, but we are at liberty to choose from among the list of plausible candidates depending on the pragmatic considerations at play (perhaps with the proviso that the larger the deviation from the top of the list, the stronger must be the pragmatic reasons that count in favor of explaining a particular event in terms of a particular cause).

How might we go about spelling out the criteria by which some causes are more metaphysically relevant than others? One possibility is presented by David Lewis's (1973) account of causation, which relies on a possible-worlds approach to counterfactuals, complete with a metric for determining the closeness of various possible worlds to the actual world (thus allowing for the selection of *the* cause—the cause that is absent in the closest possible world in which the effect fails to occur). However, the prospects for

constructing a coherent metaphysical-relevance view in terms of the proximity of different possible worlds seem to me rather dim (for largely the reasons laid out in Woodward, 2003).[18]

As we can see, the theoretical options for describing the relationship between metaphysical and pragmatic constraints on causal explanation are many, and the choice among these options will depend on some fairly arcane philosophical considerations. Fortunately, however, I think that that for our purposes here, settling the question of whether or not there are *any* metaphysical constraints on our causal explanations is inessential. In the end, I'm happy to remain agnostic on the question of whether there is a coherent way of spelling out a "metaphysical relevance" view, for the essential point for my purposes is that with regard to the sorts of causal questions that concern us (e.g., the destruction of the cathedral and, in subsequent chapters, the ontogeny of biological and psychological traits), it seems clear that there is *some* range of possible explanations in any particular case among which we might choose on pragmatic (contextual) grounds, as I will now explain.

Consider first the case of the destruction of the cathedral in Dresden. Notice, first of all, that ordinary discourse permits a certain degree of imprecision here. When presented with the question "What caused the destruction of the cathedral?" or "Why was the cathedral destroyed?" it seems appropriate to respond "the Allied bombing." The

[18] Briefly, Lewis's account proposes that, in order to identify the closest possible world in which a hypothesized cause is absent, we should imagine a "small miracle" that simply erases the hypothesized cause from existence. As Woodward points out, there are multiple ways of envisioning such miracles and hence multiple close possible worlds to consider, none of which is obviously closer to the actual world than the others.

appropriateness of this response, however, merely reflects the fact that this explanation is sufficient for many of the explanatory contexts we ordinarily inhabit. But for serious investigators of the event—historians, architects, ballistics experts, just-war theorists, and so forth—further questions will remain as to why the cathedral was destroyed at the precise time and in the precise manner that it was, and whether anyone is legally or morally culpable for the destruction. And here one might cite the pilots' pulling of the trigger, the Allied commanders' issuance of orders, the temperature resistance of various building materials, the technician who pulled an overnight shift after detecting a problem with the bomb-release mechanisms of several planes, or various of Hitler's strategic decisions that left the city without sufficient air defenses to ward off the attack. All of these, it seems to me, are perfectly good explanations for the destruction of the cathedral given the right explanatory context (i.e., they are all good answers to the questions "what caused the destruction of the cathedral in Dresden?" or "why was the cathedral in Dresden destroyed?").

Perhaps this set of explanations is in some way metaphysically "privileged" over explanations involving flying kangaroos, and perhaps not. But importantly, it seems to me it would be difficult to maintain that any of the above explanations are metaphysically privileged over the others in the set, such that we could say, for example, that the historian's focus on who gave what orders is somehow less metaphysically accurate than the architect's concern with design features of the cathedral. I contend that once we move beyond the realm of everyday, informal causal ascriptions (e.g., "the cathedral was destroyed because of the bombing") and into the domain of providing serious scientific (or even historical)

explanations, there will be *some* class of dependencies—things we might have manipulated in order to change the outcome—that share roughly comparable, and significant, degrees of metaphysical respectability. And as we will see in the next chapter, this dynamic holds as much for analyzing causal structure in biological development as for demolished buildings.

CHAPTER 2: EXPLANATION AND BIOLOGICAL DEVELOPMENT

2.1 Introduction

This chapter explores the implications of the account of causation and explanation developed in Chapter 1 for the study of biological development. I will argue that, based on the account of causation and explanation I have offered, it is theoretically baseless—if not outright incoherent—to describe the emergence of a particular trait—psychological or otherwise—in a particular individual as "caused more by genetic than environmental factors" or "more genetically than environmentally determined." As with the bombing of Dresden, our everyday causal talk tends to uncritically accept explanations for developmental outcomes that appeal to genes as "critical" or "preponderant" causes. But, again as with the Dresden bombing case, I wish to argue that such naïve appeals to genes as causes mask considerable causal complexity inherent in biological development, and rely on an oversimplified notion of causal explanation. A close examination of the details reveals the same sort of complicated causal story—and attendant context-dependence of the appropriateness of various explanations—as we saw in the case of the Dresden bombing.

This is a claim about the causal metaphysics of organismal development. However, in spite of the causal parity between genetic and extra-genetic factors in development, I argue that in many cases citing genetic factors (to the exclusion of environmental ones) can in fact be *explanatory*—indeed, in many contexts an appeal to genes may well constitute

the *best* explanation for a particular trait. Nevertheless, given the kind of complexity we are likely to observe in biological systems, calling a trait "genetic" will often obscure rather than illuminate the causal structure of developmental processes. As an alternative, I suggest that a mechanistic approach (which need not identify uniquely influential causes or apportion causal responsibility) will supply superior explanations for many phenomena of biological development, though again this "superiority" is defeasible in that it will not necessarily hold in *all* explanatory contexts.

It's important to note at the outset that my remarks about genes and development in this chapter are restricted to the development of *biological*—as opposed to *psychological*—traits; Chapter 3 takes up the latter. (Some may balk at the idea that there is a clear distinction between the two types of traits, but I will present arguments in the ensuing chapters that not only can such a distinction be drawn, but that drawing it is essential to making sense of the explanatory practices of those studying biological and psychological development.)

2.2 Explanation and complexity in biology

Let's begin by reviewing the kinds of causal scenarios covered in the last chapter that demonstrate the independence of *production* and *dependence* as types of causation, as well as the need to include both in our explanatory repertoire. There is *production* without *dependence* when Billy and Suzy both launch their rocks at a bottle but Suzy's gets there first (overdetermination). Likewise, there is *dependence* without *production* when Billy shoots down the enemy who would have prevented Suzy's bombing (double prevention).

Importantly, the kinds of problems for both *dependence* and *production* approaches to causation revealed by the exploits of Billy and Suzy do not arise only in the rarefied air of philosophical discourse, populated by known counterexample-philes. Rather, these same issues arise in many genuine cases of scientific investigation and explanation, particularly in areas highly germane to the project of this dissertation, viz., gene expression in biological systems. Two examples will illustrate.

Overdetermination occurs in cases of *genetic redundancy* and *robustness*. One of the most powerful methods in the toolkit of molecular geneticists is the "knockout (KO) experiment," which involves inactivating (one or both copies of) a gene and observing the phenotypic outcomes. Thus, KO experiments allow us to determine the functions performed by particular genes by showing us what happens when they are prevented from performing those functions.[1] Varying degrees of developmental disruption (including death) are observed in KO experiments, but in some cases there appear to be no phenotypic differences at all.[2] This could simply mean that the knocked-out gene simply has no appreciable effect, except that in many such cases, specific functions (often important ones) for the relevant proteins have already been established on independent grounds. It turns out that biological systems can often compensate for an inactivated gene in a variety of ways. Sometimes there are simply redundant copies of the same gene that become active when

[1]In speaking of the "functions" of a gene, I do not mean to endorse the idea of evolutionarily *proper* functions of the sort defended by Millikan (1989); I have in mind, rather, causal-role functions in Cummins' (1975) sense.

[2]There is some disagreement concerning just how robust developmental pathways can be in the face of a gene knockout. And clearly, if we individuate our events very finely, it will not be the case that the developmental outcomes arising from activation of the "backup" genetic pathway are "the same" as if the primary pathway had performed its function. I'm assuming that there's an explanatorily interesting degree of abstraction at which the outcome does not depend on the particular (primary) genetic factor initially identified as its cause.

one is knocked out, but sometimes the gene is embedded in a complex network of mechanisms that allow for the same output to be produced even if the primary pathway is disrupted (Greenspan, 2001). These sorts of causal networks, which are ubiquitous in biology (Edelman & Gally, 2001), exhibit a form of *overdetermination*: in most cases a specific gene initiates a causal sequence that eventuates in a particular phenotypic outcome, but if this gene is inactivated, a *different* causal sequence will be initiated that produces the same result. So even if we have identified a causal chain from gene to phenotype, it may not be the case that the phenotypic outcome counterfactually depends on the activity of that gene.

A real-life case of *double prevention* can be found in the metabolization of lactose by *E. coli*. Francois Jacob and Jacques Monod (1961; see also Morange, 1994, Chapter 13) first modeled the mechanism by which the metabolization of lactose in *E. coli* takes place. According to the Jacob and Monod model, the metabolization of lactose occurs via the activities of a trio of enzymes (which are coded for by structural genes[3] known as *lacZ*, *lacY*, and *lacA*) that are produced by *E. coli* only when lactose is present. Under ordinary circumstances (i.e., when lactose is absent), a protein present in the *E. coli* nucleus binds to the regions of DNA containing *lacZ*, *lacY*, and *lacA*, thereby repressing or inhibiting the activity of these three genes. This protein is therefore called a *repressor protein*. In lactose-rich environments, however, allolactose (an isomer of lactose also present in such environments) binds to the repressor protein and prevents it from performing its usual

[3]Genes can be separated into two types: *structural genes* and *regulatory genes*. Regulatory genes code for elements (proteins or RNA) that themselves regulate the expression of other genes. All other genes (e.g., those that code for proteins involved in metabolization processes, like the *lac* genes) are structural genes.

inhibitory role. When the usual inhibitory activity of the repressor protein is disrupted by the allolactose molecules, the structural genes (*lacZ*, *lacY*, and *lacA*) that produce the lactose-metabolizing enzymes become active, synthesizing the appropriate enzymes and enabling the metabolization of lactose.

Now, suppose we are interested in explaining the presence of the lactose-metabolizing enzymes. Their presence owes to the fact that the repressor protein is no longer serving as a preventative agent, because the repressor protein is itself *prevented* from acting in a preventative role. Hence: double prevention.[4] The upshot is that while explaining the presence of the metabolizing enzymes seems to require citing the presence of allolactose, *there is no spatiotemporally-continuous chain linking the allolactose molecules to the synthesis of the amino acids that make up the enzymes.*

There are two lessons I think we should draw from these two examples. The first is that, as Mitchell (2009) has emphasized, biological systems are *complex* in a way that complicates, if not precludes, traditional causal thinking (Mitchell stresses this particularly in relation to examples of genetic robustness). Much of the importance of this point will emerge in later chapters.

The second lesson is that in cellular biology, just as in philosophical thought experiments, sometimes thinking of causal relations strictly in terms of *dependence* or strictly in terms of *production* fails to capture the relevant aspects of the causal structure of the phenomena we investigate. Again, this result would be deeply problematic if we

[4] If it helps, here are the equivalencies between the present example and the case of Billy and Suzy the fighter pilots: presence of metabolizing enzymes = Suzy's bombing; repressor protein = Enemy; presence of allolactose = Billy's firing.)

were inclined to think that all causal explanation must fit a singular mold. But given a pluralistic and context-dependent conceptualization of causal explanation, this is precisely what we should expect. And fortunately, we have at our disposal an explanatory strategy (the mechanistic approach) that allows us to incorporate *both* of these types of causal relation into an integrated description of how and why the observed phenomenon arises.

2.3 The mechanistic approach to explanation

The mechanistic approach to explanation bears some relation to the *causal process* view of causation (Dowe, 1992; Salmon, 1984), in that it focuses on the processes by which phenomena are *produced*. Despite this similarity, the mechanistic approach contains distinct advantages that enhance its suitability for the purposes of analyzing biological phenomena

The first advantage that mechanistic analysis brings over the causal process view is that mechanistic explanations need not satisfy Salmon's strict constraint of spatiotemporal continuity. Thus while it is doubtful that Salmon's or Dowe's views could account for the explanatory relevance of allolactose to the presence of lactose-synthesizing enzymes in *E. coli* (since there is a break in spatiotemporal continuity between the former and the latter, just as there is a break between Billy's firing and Suzy's bombing), a mechanistic analysis of the phenomenon can simply incorporate the complete sequence of events beginning with the binding of allolactose molecules to the repressor protein and ending with the presence of lactose-synthesizing enzymes within the cell.

Secondly, while there are some variations of the mechanistic approach that are couched in causal terms (e.g., Glennan, 1996), the mechanistic strategy can be fruitfully pursued even if it is not construed as a theory of causation, as has been demonstrated by Machamer, Darden, and Craver (2000; henceforth MDC). Rather than seeking to identify particular elements as causes of a particular phenomenon, the mechanistic strategy in the MDC tradition explains phenomena by *describing the activities of the components of a mechanism from start to finish* (where "finish" means, roughly, the emergence of the phenomenon of interest). This flexibility with causal analysis makes the mechanistic approach particularly well-suited to explaining double-prevention phenomena like the metabolization of lactose—more so than similar "productive" views that are couched in explicitly causal terms. An approach (like that of Salmon or Dowe) that ties explanation to causal analysis in terms of spatiotemporally continuous processes forces us to rule out the activity of the allolactose. Since the mechanistic approach is not concerned with separating causes from non-causes, however, it faces no difficulty in including the activity of allolactose in its explanation of the phenomenon. The mechanistic approach gives us the freedom and flexibility to include in our model those entities and activities that seem important for our understanding of how the phenomenon comes about.

An immediate worry one might have about the mechanistic approach to explanation is that insofar as the mechanistic approach merely *describes* the events that lead up to the phenomenon of interest, this approach cannot be *explanatory* at all. That is, one might insist that *describing* events is one thing, and *explaining* them quite another. However, this worry is misplaced for two reasons. Firstly, a crucial component of providing a mechanistic

explanation is to *specify the explanandum*—the phenomenon of interest—at the outset. This ensures that the *explanans* (the entities and activities that produce the phenomenon) differs from the *explanandum*, and so a mechanistic explanation is not a mere redescription of the phenomenon to be explained.

Secondly, I would argue that any philosophical theory of explanation must be grounded in actual scientific practice (cf. Craver, 2007). In particular, absent evidence to the contrary, we should assume that scientists working in a particular domain do in fact provide explanations for the phenomena they investigate, and that the things that they treat as explanations actually do constitute explanations. And as MDC point out, attention to the explanatory practices of cellular and molecular biology reveals that explanation in these domains very often, if not always, proceeds via the modeling of mechanisms. Thus, we have good reason to believe—and little reason to doubt—that the mechanistic strategy provides us with good scientific explanations, at least within the context of the explanatory projects with which many biologists routinely concern themselves.

2.4 Causation, explanation, and context in biological development

In this section I expand the discussion of pursuing explanation in the face of complex biological systems, and I further develop the argument that traditional causal analysis is unsuitable for many explanatory contexts in the study of development, and that mechanistic approaches offer a promising strategy for meeting the explanatory needs of developmental biology.

2.4.1 Context matters

Let's begin with some general observations about the role of context in explaining biological traits. I demonstrated in Chapter 1 how the bombing of Dresden Cathedral admits of multiple causal explanations, depending on what explanatory question is being asked. Now consider the project of explaining why humans have two eyes at the front of the head. As with the Dresden case, ordinary language allows for some imprecision: for casual observers it is perhaps sufficient to say that humans have two eyes "because of the genes they carry." And this informal explanation is fine as far as it goes, but note that it leaves out a great deal about *which* genes are causally relevant to the building of eyes as well as *how* they do it and what we might do to manipulate the outcome. So we might reasonably expect a genuine scientific consideration of the question to demand more. And here again, the appropriateness of different explanations will depend on precisely what question is being asked and for what purpose.

The reason for this interest-relativity is that, as we saw in the previous chapter, seemingly-innocuous "why" questions can exhibit considerable ambiguity insofar as they are consistent with multiple sets of presuppositions and multiple contrast classes. Cowie (1998) provides a useful illustration: a geneticist might seek to explain why humans typically have two eyes—rather than eight—by appealing to the genetic differences between humans and jumping spiders, and to explain why these eyes are on the front of the head—rather than on the sides or back—by appealing to the genetic differences between humans and birds. An epidemiologist, in contrast, might seek to explain why humans typically have two normal-shaped eyes on the front of the head—rather than the larger,

elongated eyes towards the side of the face that are characteristic of fetal alcohol syndrome—by appealing to the (usual) absence of alcohol from the prenatal environment. Again, neither of these explanations seems metaphysically privileged—neither researcher is necessarily "missing" something about the causal structure of development—it is simply that the researchers are engaged in different projects and asking different questions (more specifically, their questions presuppose different background conditions and hence include different contrast classes).

2.4.2 Causal analysis and gene-environment interaction[5]

As a further example, I present below a series of discoveries concerning the development of sexual behavior in male rats. As we will see, the interplay between genetic and extra-genetic factors in this instance is quite complex, in such a way as to frustrate attempts to apportion causal responsibility to genes and environment. However, by presentation of this example I do not mean to suggest that *all* cases of biological development are similarly resistant to simple causal analysis—many are plausibly not. Rather, I present the example as a sort of illustrative worst-case scenario: to the extent that biological development turns out to be like this in general (and I will comment below on how prevalent this sort of developmental story is likely to be), the general utility of apportioning causal responsibility to genes and environments in the explanation of biological traits will be minimal.

[5] The term "gene-environment interaction" here is intended in the everyday sense of *causal* interaction, not the sense of *statistical* interaction between genetic and non-genetic variance that arises in quantitative behavioral genetic analysis, which is discussed in Chapter 7.

Developmental psychobiologist Celia Moore has provided insight into the complex interplay of genotypes and environments in laboratory rats (C. L. Moore, 1984, 1992). Mother rats routinely engage in licking of the genital areas of their pups, and the amount of licking provided has been shown to vary with the levels of various hormones secreted by the pups. Thus, insofar as the amount of maternal licking is an environmental stimulus, different pups within the same litter can be exposed to substantially different environments. The important thing to notice is that the nature of the environmental conditions depends on variable genetic factors: the neonatal environment a pup experiences depends on the activity of its genes (i.e., the amount of licking a pup receives varies with the degree to which the genes that control its hormone levels are expressed). (Even these genetic factors are not, of course, free from extra-genetic influence; pups gestated by mothers in a crowded, stressful environment exhibit diminished expression of the relevant hormone-controlling genes, and thus elicit less maternal licking by adoptive mothers than do controls.)

These environmental differences—differences in levels of maternal licking—are associated with differences in a wide range of behaviors in pups, many persisting into adulthood. One notable finding is that gene expression in the spinal cord of male pups is substantially influenced by levels of maternal licking, and the ability to perform various sexual behaviors in adulthood depends on proper development of spinal cord nuclei. Thus, male pups who receive inadequate amounts of licking show a variety of abnormal sex behaviors (C. L. Moore, 1984).

How should we classify the development of sexual behavior in male rats? Is it genetically-caused or environmentally-caused? As a first pass at answering this question, we can approach the problem by specifying the causal relationships among the different aspects of development, and analyzing them according to the manipulability approach described Chapter 1. Here we have a causal system comprising five variables: the prenatal environment (P); genes that regulate production of hormones that elicit maternal licking (H); maternal licking (M); genes controlling development of spinal cord nuclei (N); and finally sexual behavior in maturity (S). For simplicity's sake, let us assume that each variable has only two values (stressful or not for the prenatal environment, on or off for the genes, present or absent for licking and for normal sexual behavior). Of course the relationships among these variables are, in reality, more nuanced, but the basic causal structure shown Figure 2.1 below still holds. Our graph looks like this (for clarity, the genetic variables appear in boxes, the environmental variables are encircled, and the behavioral outcome is in regular type):

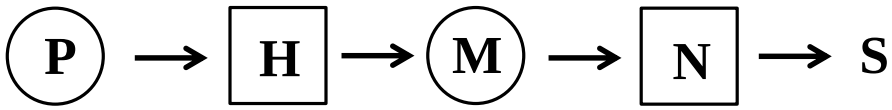

Figure 2.1: Causal model of rat pup development

In pursuing a Woodward-style analysis of the system, we can begin by inquiring as to the outcomes of performing interventions upon the different variables. Suppose we intervene upon N (disrupting gene expression in the developing spinal cord of a normal

male pup). As a result of this intervention, the pup displays abnormal sexual behavior. Hence, N counts as a cause of S, so perhaps we should count sexual behavior as a "genetic" trait, since we have identified a direct causal relationship between a genetic factor and a trait of the organism.

The problem with this suggestion is that N, while representing a genetic factor, is directly caused by an *environmental* variable (maternal licking); this relationship can be ascertained by imagining an intervention on M, which would change the value of N (licking causes gene expression in the spinal cord). Furthermore, this sort of pattern (whereby environmental stimuli trigger gene expression that yields physiochemical changes) is rampant in biology. To cite but a few examples, the effects of exercise on physiological systems are mediated by the expression of genes that effect adaptational changes (e.g., running long distances triggers expression of genes that build additional capillaries to carry blood to muscles in the legs), as are the effects of stressful situations on the body's response to future stresses (see, e.g., Sabban & Kvetňanský, 2001). And, in a recent finding, Tung et al. (2012) identified a series of genes that wreak long-term changes to the immune systems of macaques (e.g., causing chronic generalized inflammation) in direct response to being introduced to a new social group (and hence occupying the lowest rung of that group's hierarchy).

It seems safe to say that none of the traits identified above (capillarization optimized for endurance running, responses to stressors that emerge late in life and are contingent on past stressors, and chronic inflammation) should receive unqualified classifications as "genetically caused" if they arise in the manner specified. And yet, in each case, we could

make a difference to the emergence of the trait by performing interventions on the genome after the environmental stimulus has occurred but before the trait has emerged. Thus, it cannot be the case that *any* time genes are a proximal cause of the emergence of a trait, the trait is (uniquely) genetically caused. Rather, as these examples demonstrate, when the gene expression that causes the emergence of a trait is itself caused by environmental stimuli, labeling the trait "genetic" obscures, rather than illuminates, what seem to be the relevant features of the trait's causal etiology.

Now let us consider H and M. Can either of these be given a privileged explanatory role? Both intuitively and formally, the answer seems to be no. Intuitively, the presence of the environmental factor, maternal licking (M), *depends on* the presence of particular genetic factors for hormone secretion (H) (this is in contrast to other environmental stimuli, such as the presence of oxygen, water, or predators, which the organism has no role in creating), and the influence of genes on S is *mediated by* an environmental factor (M). And if we consider making interventions in the system, we see that with respect to the value of S, intervening on H and intervening on M yields precisely the same result.[6] Thus, it appears that applying a manipulationist causal framework to this system simply cannot identify *any* of the variables—genetic or environmental—as a privileged cause of the behavioral outcome S. The upshot is that in cases of this sort there is no way to escape causal parity between genotype and environment when the detailed developmental story is considered.

[6] I am ignoring certain subtleties here regarding the individuation of events (or variables or states of affairs—whatever are considered to be the causal relata). The precise behaviors exhibited by rats subject to interventions on H and M may well differ in some respects, but they are alike in that they deviate from the normal outcome.

Here we can see the utility of the mechanistic approach to explanation for application to this sort of case. A mechanistic analysis of the present case would include not only a linear sequence of causally-relevant variables, but also an account of *how* these components interact: if effect, this amounts to "filling in" the arrows of the above diagram (which merely indicate causal relevance) with descriptions of the *activities* enacted by the relevant entities. Again, mechanistic analysis need not be causal analysis in the sense of enumerating the causal antecedents of an event and/or singling out one cause as a metaphysically privileged "critical" cause. Characterizing the mechanism by which rat pups come to exhibit particular sexual behaviors makes no requirement that we distinguish between causes and non-causes, let alone that we single out a particular cause as *the* cause. Rather, the mechanistic approach licenses us to describe in detail how the components of the developmental system—broadly construed—interact with one another, just as I have done above. Such a mechanistic description of the system, it seems to me, constitutes a far better explanation of the phenomenon than we are likely to get by attempting to apportion causal responsibility to the different components of the system.

Again, however, the superiority of the mechanistic story does not hold in an absolute sense, since the appropriateness of any explanation is governed by features of the explanatory context. But nevertheless there are several reasons why it seems legitimate to consider the mechanistic story a superior explanation *most of the time*. First, recall that I have characterized explanation as largely a matter of *facilitating understanding*. It seems clear that one who knows not just which genetic variables and which environmental variables are relevant to the development of the rats' sexual behaviors, but also *how these*

components interact, understands the rats' development more fully, and is also better poised to make sense of the pattern of experimental results (e.g., the long-range dependence of sexual behavior on prenatal environment). Second, knowing the mechanistic details of this developmental process affords a wider range of prospects for manipulation and control.

So while a mechanistic explanation of the rats' development is perhaps not *inherently* superior to labeling the traits "genetic" or "environmental," it seems safe to say that the mechanistic approach provides distinct advantages in what are quite likely to be the most common explanatory contexts. Thus, it remains the case that whether something counts as "genetically caused" is, *strictly speaking*, a matter of the explanatory context—in this case, how much we are "zooming in" on the mechanistic details or abstracting away from such details (i.e., "holding them fixed") to consider more course-grained facts about associations between, e.g., particular alleles and particular eye colors. But nevertheless, we can still say that for a great many (perhaps most) explanatory purposes in the context of developmental biology, we are better off not trying to apportion causal responsibility. Rather, it seems our biggest explanatory payoff will come from detailed mechanistic accounts of the entities and activities that produce the phenomenon. In effect, the ontic nature of developmental processes—as complex, mutually-influencing systems—*fixes the pragmatics of the explanatory context* in such a way as to privilege mechanistic explanations.

2.5 The fate of "genetic" traits

For all I have said about the futility of attempting to partition genetic and environmental causes of development, and of labeling the outcomes of complex, interactive developmental processes as "genetic traits," I acknowledge it would be a black mark on an account of explanation in biological development that it precluded such common-sense truisms as *eye color is a genetic trait in humans*. We should therefore see what can be said within the present framework by way of accommodating these sorts of statements.

2.5.1 Some traits are genetic (in most contexts)

To begin, it's worth pointing out that at *some* level of analysis, every instance of biological development—including eye color—will exhibit a kind of causal parity between genetic and extra-genetic factors. This is more than the well-worn point that every trait results from both genes and the environment, which often amounts to nothing more than the claim that the presence of some-or-other genetic or environmental conditions is necessary for trait development. Rather, I am here pointing out that by "zooming in" far enough on the details of biological development—by attending to the ways that nucleotide strings and protein molecules interact with cellular machinery, the mechanisms regulating gene expression, and so forth—we will reach a point where genetic factors (i.e., DNA strands) and extra-genetic ones are simply interacting and mutually-influencing components of a complex causal system. So at *some* level of analysis even eye color will have a causal story resembling that of sexual behavior in rats, and there may well be contexts in which this level of analysis is the most explanatorily appropriate one.

Is this sufficient to rule out classifying eye color as a genetic trait? It depends how the claim is interpreted. If the claim *eye color is a genetic trait* is interpreted as a claim that those causal antecedents of eye color that reside within the genome occupy some sort of *metaphysically privileged* place in the causal etiology of the trait, then I think the claim is simply false, for all the reasons laid out in the previous chapter and in this one. But this does not mean that the claim, interpreted more weakly, is explanatorily illegitimate; indeed, as should be clear from the preceding discussion, I think that the search for such "metaphysically privileged" causes *anywhere* in science is a fool's errand and presents too high a standard for explanation.

The lessons learned in the Dresden case will apply here as well. Denying that the metaphysics of causality specify a unique cause for any event (since there will be multiple variables upon which any event counterfactually depends—or alternatively, multiple interventions that would have modified the outcome, à la Woodward), does not mean anything goes. To say that within a certain range the choice of which causal factor to cite as the explanation of a given phenomenon is a pragmatic one is not the same as saying the choice is arbitrary and/or completely relative. In other words, choices of explanatory contexts are subject to some degree of *rational criticism*. And this cuts both ways: while someone who declares that sexual behavior in rats is genetic is clearly *missing something,* we can similarly point out that someone who insists that blue eyes cannot be said to be genetically caused—because the HERC2 gene must interact with non-genetic elements in order to influence eye pigment—is being obtuse (even if the claims of both these hypothetical scientists are metaphysically legitimate).

The claim about metaphysical parity among an event's causes defended throughout the last two chapters is a pretty weak one: there will be *more than one* possible causal explanation for any given event, among which we must choose on pragmatic grounds. But while the constraints are ultimately pragmatic—i.e., dependent on the explanatory context—there is likely to be large overlap in the pragmatic features governing the appropriateness of different explanations on different occasions, and this will be truer of some causal relations than of others. In other words, and to apply the point to the present case, it's quite plausible that many of us will go through our entire lives without ever setting foot in an explanatory context in which facts about a creature's genotype do not constitute a good explanation for its eye color (or indeed, in which it makes more sense to describe an ignition as being caused by the presence of oxygen than by the striking of a match). But this does not, again, *metaphysically* privilege these explanations.

So is eye color "caused by the genes"? Yes, if only in the minimal sense that genes are among the causal antecedents of an individual's eye color being what it is. But by the same token, strictly speaking eye color is caused by environments as well (in particular, by extra-genotypic factors within the developing organism, with which DNA segments causally interact), and there may well be contexts (e.g., among scientists studying the molecular and cellular mechanisms by which the iris comes to have a particular melanin concentration) in which these factors constitute the most salient causal antecedent to an individual's eye color. But for most everyday as well as scientific purposes, it will be explanatorily sufficient to say that an individual's eye color is a result of their genes, even

if it is still inappropriate to say that their eye color was caused *more* by their genes than by any environmental (extra-genotypic) factors.

2.5.2 How do we tell which traits are genetic?

We have looked at different sorts of cases in biological development, and I have concluded that *sometimes* it is legitimate to give an explanation for a trait couched in genetic terms (eye color), and *sometimes* it is illegitimate (sexual behavior in male rats). How are we to tell one case from the other?

My view on this is that since the legitimacy of an explanation is determined by the explanatory context, which is in turn determined by our explanatory interests, there is little that can be said *in general* by way of an answer to the question of how to tell if a trait can be "legitimately" described as a genetic one—our explanatory interests, again, vary substantially from case to case and it's not clear that this variation is particularly systematic. But, *very* roughly speaking, we can say that when a trait is like eye color, in that there is a (relatively) direct causal path from genotype to phenotype, and where the link between genotype and phenotype is (to use Woodward's term) *invariant* across a wide range of environmental conditions that are of interest to us, then the trait is, all else equal, a good candidate for being described as "genetic." Traits that are more like the rat pups, where these conditions do not hold, we are better off foreswearing the "genetic" label.

Now, just how these two types of traits, and the multitude of cases that would fall intermediate between them, are distributed in nature is a wide open empirical question and a matter of considerable scientific debate. Proponents of so-called "Developmental

Systems Theory" (DST) in biology (Gottlieb, 1995; Gray, 1992; Meaney, 2001b; Oyama, Griffiths, & Gray, 2003; Oyama, 2000) are fond of citing cases like the rat pups and declaring that the notion of a "genetic trait" is hopeless; others point out that relatively few such cases have been discovered and described in detail, and argue that a few examples of interestingly complex developmental processes do not undermine the common scientific practice of studying development under the assumption that genes are, for all practical purposes, relatively self-contained and independent causal factors (Sesardic, 2005).

Despite the rhetorical posturing on this issue, it seems clear that we are nowhere near being in a position to say very much about the epistemic question at hand—our understanding of biological development lacks the degree of breadth and detail that would be required to start sorting developmental processes into "rat pup" and "eye color" bins and then counting the results. But a few remarks can be made. One is that developmental systems researchers like Celia Moore (of rat pup fame) and Gilbert Gottleib (1991) have certainly enhanced our understanding of particular developmental processes and modes of coaction between genes and environments through detailed experimental work. Hence, DST's embrace of causal parity in epigenesis seems, at worst, one of several successful research paradigms in biology.

Another point is that interactive systems like those observed in the rat pups may turn out to be much more prevalent than we would have thought, because as it turns out such systems are often quite good at hide-and-seek. Many cases of mutual causal influence among genes and environments reflect highly stable relationships, with developmental regularity within a species often being as much a matter of organisms creating similar

environments for themselves as of similar genes building similar traits (a phenomenon sometimes known as "niche construction"). Thus, there may well be any number of "relevant" environmental variables (i.e., variables that, if manipulated, would make a difference to developmental outcomes) that exhibit almost zero *actual* variation in natural environments, but this only because the genotypes of the organisms have a hand in fixing the values of these environmental variables. For instance, it is unclear how much actual variation exists in natural populations of rats vis-à-vis maternal licking, but even if there were none (because of the absence of variation among pups in the genes that stimulate licking), our understanding of the development of the male rats' sexual behavior would be significantly impoverished if no one had ever thought to investigate the relevance of maternal behavior, and the trait were simply labeled "genetic." The lesson to be learned here is that the absence of "natural" variation in a population—and *even a strong correlation of particular genes with universally-manifested traits* (as discovered through experimental manipulations)—does not necessarily mean that the trait is best classified as "genetic" (e.g., on the analogy with eye color). Detailed experimental work is needed to uncover the relevant mechanisms.[7]

Finally, as several psychologists and philosophers have pointed out (Griffiths & Machery, 2008; D. S. Moore, 2008), many of the ways in which genes and environments influence one another are subtle and often non-obvious (though they might make sense in

[7] Clearly this point can be pressed too far if one insists—no matter the available evidence—that labeling a trait "genetic" is illegitimate (or at least premature) because some heretofore unconsidered environmental variable may be playing a role. However, once the developmental mechanisms have been explored experimentally and reasonable attempts to detect gene-environment interactions have been made, the case for labeling a trait genetic will be much stronger—this point will be especially important in Chapter 7.

hindsight). For example, the posture of a mother rat while nursing her young (prone or arch-backed) turns out to have dramatic effects on the expression of genes regulating various aspects of the pups' neural development, which in turn influence a number of adult behaviors (Meaney, 2001a). Thus, it is easy to see how we might mistakenly think that once we have observed a consistent relationship between a genotype and phenotype, invariant across manipulations of a number of the more intuitively plausible environmental variables, we could conclude that the emergence of the trait given the genotype is invariant across all but the most ancillary background conditions. As the preceding examples show, however, we may well have significantly underestimated the pervasiveness of the sorts of complex gene-environment interplay that frustrate attempts to identify genes that serve as explanatorily privileged causes.

To summarize the points raised in this section: we can, in principle, distinguish between traits that are "genetic" and those that aren't (though the difference, being rooted in pragmatic considerations, may seem unduly *un*principled to some). However, we may not be in a position to say which is which for a great many traits of interest (and moreover, at this point it's difficult even to guess at the relative frequencies of tractable cases like eye color versus more unruly ones like the rat pups). These considerations, I think, recommend a general policy of caution and an openness to revision in the application of the label "genetic" to biological traits, but nonetheless we can, sometimes at least, legitimately identify genetic traits.

CHAPTER 3: INNATENESS IN COGNITIVE SCIENCE

3.1 Introduction: aims and approach

The notion of innateness is thought to play a crucial role in a number of well-developed theories in cognitive science, but there is no consensus on the matter of what innateness *is*. In other words, cognitive scientists—especially developmental psychologists and linguists—routinely claim that certain features[1] of the mind are innate, but it is not immediately apparent what these claims amount to. There are a number of proposed analyses of innateness on offer, but again, no consensus on which of these theories—if any—best captures the use of the term in nativist research programs in cognitive science. Thus, cognitive science (and the philosophy thereof) are in need of a theory of just what we commit ourselves to when we describe some psychological trait as innate.

This chapter aims to provide such a theory. My approach takes for granted that the notion of innateness performs some explanatory work in contemporary psychological theory (as its practitioners clearly believe), and so I will seek to discover what innateness claims amount to by looking at how the concept is *employed* in the explanatory projects of mainstream cognitive science. I conclude that the best analysis of innateness as an explanatory concept in cognitive science can be found in a family of views that count innate

[1] Since a precise account of what sorts of mental entities are potential satisfiers of the predicate "x is innate," is beyond the scope of this chapter, I will employ—interchangeably—a variety of catch-all terms: *cognitive/psychological features, traits, structures*, and so forth. Here I rely on the reader's intuitive sense of what kinds of things we might call innate; I believe this will suffice for the purposes of the present discussion.

traits as those that arise in the *absence* of certain kinds of information processing on the part of the organism. The most prominent recent exemplar of this kind of view is Samuels' (2002) *Primitivism* account (i.e., innate traits are those that are *psychologically primitive*). The present work defends the general approach exemplified by Samuels' account, while also refining the primitivist theory so as to resolve problems with Samuels' formulation.

The account offered in the present chapter should be of interest both to specialists in cognitive science and to anyone with an interest in the scientific study of human nature. The account is important for three specific reasons. Firstly, my account of innateness provides a coherent foundation for the scientific study of human nature. One of the most successful research programs in all of cognitive science has been the identification of a wide variety of early- and universally-developing human psychological traits (Carey & Spelke, 1996; Feigenson, Dehaene, & Spelke, 2004; Landau & Gleitman, 2009; Scholl & Leslie, 1999; Spelke & Kinzler, 2007; Wynn, 1998). This research program (referred to herein by the broad term "nativist cognitive science," or, where appropriate, the narrower terms "nativist psychology" and "nativist linguistics") is conceived of by its practitioners as the search for human nature, in that it seeks to uncover both that which is distinctively human about the human mind, and that which is universally shared among members of our species. This research program constitutes our best hope for a scientific account of human nature, and understanding human nature therefore requires understanding how *innateness* is understood in nativist research.

Secondly, even though the account of innateness offered here is confined to *psychological* innateness, this account forms an integral part of a broader theoretical

account of innateness across different scientific fields. This broader account, developed throughout the first four chapters of this dissertation, demonstrates that although the various phenomena referred to as "innate" across different scientific fields and different explanatory contexts are *disunified*—the term picks out different phenomena in different contexts[2]—each of these phenomena have a legitimate role to play in certain explanatory projects. The complete account in which the present work is situated thus makes sense of a wide swath of scientific discourse that would otherwise appear inconsistent or unintelligible.

Finally, my account of innateness provides a response to those (e.g., Oyama, 2000) who argue that because (as they claim) the notion of *genetic determination* is incoherent (or at least ill-defined), the nativist research program in cognitive science is fundamentally misguided. I share much of this skepticism about the idea of genetic determination, but importantly, in this chapter I show how the notion of innateness utilized in nativist psychology and linguistics can be spelled out without reference to genes or genetic determination. Thus, I set nativist cognitive science (i.e., the scientific study of human nature) on firm theoretical footing, independent of whether the idea of genetic determination can be given a coherent and useful scientific interpretation.

Here is a road map of the chapter: Section 3.2 lays out some preliminary matters, including some assumptions I will be making—but not defending at length—regarding the nature of explanation in cognitive science, as well as some constraints on an adequate notion of innateness for cognitive science. Section 3.3 sets out the Primitivist account and

[2] Another way of expressing this point is to say that the term "innate" does not pick out a single natural kind.

refines the version offered by Samuels. Section 3.4 takes a close look at the role played by innateness in some actual recent nativist theorizing, specifically the construction of a Poverty of the Stimulus Argument for innate syntactic competence, and also analyzes the debate that has ensued between nativist and anti-nativist camps about that argument. It is in the course of this analysis that I defend the claim that the Primitivist view best captures the notion of innateness utilized in cognitive science.

3.2 Preliminaries

3.2.1 Assumptions

Insofar as my argument is concerned with providing an analysis of innateness in cognitive science, it will help to begin by identifying some basic theoretical commitments I adopt regarding this branch of science. There is insufficient space here to defend these assumptions (though of course I think they can be defended); I lay these out primarily because the Primitivist view of innateness I defend is, I think, significantly more transparent and more compelling in light of these assumptions.

Assumption 1: Psychology does not reduce to biology (or to physics).

I adopt a non-reductive view of science, and of psychology in particular. For essentially the same reasons originally set out by Fodor (1975), and elaborated by Rey (1997), I take it that psychology is a "special science," in the sense that there are interesting scientific generalizations that can only be had at the psychological level. In particular, there

are regularities in the behavior of intelligent creatures that could not be satisfactorily explained without reference to intentional states. In short, certain phenomena for which we desire scientific explanations must be given psychological explanations.

Assumption 2: The "psychological level" is characterized by intentional states.

I also adopt an account of cognition that takes the *intentional* or *representational* nature of mental states to be what is distinctive of the psychological level (Harman, 1990; Rey, 1997). According to this representational account, phenomena at the psychological level are constituted by *representational states* that are causally efficacious in the behavior of the organism.[3] Mainstream cognitive science, I take it, is in the business of generating scientific explanations by developing theories about the principles that govern the functioning of these states and their role in generating behavior. To give a "psychological explanation" for something is to explain it by appeal to the representations that constitute or cause it. Similarly, a "psychological process" is a process that involves the activity of intentional states.

[3] This should not be taken as an endorsement of the view that *non-physical* mental substances or properties are causally efficacious. Rather, the causal efficacy of mental states is perfectly consistent with those states being entirely physical, even if they do not *reduce* to physiological states. Mental states can be physically realized by, for example, computational or connectionist systems (Churchland, 1996; Fodor, 1975; Marr, 1982; Rey, 1997; Smolensky, 1988). Indeed, much of the description of mental operations throughout this paper is couched in computational terms; however, I am entirely agnostic about how best to characterize the physical realization of mental states (i.e., in computational or connectionist terms)—any computational language in the ensuing discussion can be translated into equivalent connectionist terms, if the reader is so inclined.

Assumption 3: Linguistics is a branch of psychology.

This assumption should be relatively uncontroversial, as it has been a guiding assumption for mainstream linguistic theory at least since Chomsky's *Aspects* (1965). Nevertheless, as there have been some recent challenges to this idea (Devitt, 2008), it is worth stating explicitly. Since my case study for the role of innateness in linguistic theory is drawn from the literature on language acquisition, it is important to highlight the fact that modern linguistic theory (including the study of language acquisition) takes its primary task to be the discovery of the representational facts that underlie human linguistic competence. Thus, our search for the "psychological notion of innateness" or the "notion of innateness used in cognitive science," is also the search for a notion of innateness for linguistic theory.[4]

3.2.2 The explananda of innateness claims

As a further constraint on theories of innateness in cognitive science, I argue that innateness claims are aimed at providing *individual*-level explanations. That is, the primary explanatory focus of innateness claims is providing explanations for the emergence of traits in individuals. This conception of the explanatory focus of innateness claims conflicts with—and if my arguments here are successful, invalidates—the otherwise plausible-

[4] Interestingly, Chomsky himself has at certain points offered an explication of his views that renders my three assumptions inconsistent. That is, he has claimed that the science of linguistics is *not* committed to the existence of mental representations of linguistic entities, which would mean that linguistics is not a branch of psychology, if psychology is what I have claimed it to be (the study of representations). However, I am employing here what I take to be the "standard" interpretation of Chomskyan linguistics, on which the explanations offered for linguistic phenomena are committed to the existence of mental representations (see the exchange between Chomsky and Rey in Antony & Hornstein, 2003).

sounding idea that what makes something innate is that it is universal among normally-developing members of a species. My claim about the individual-level explanatory focus of innateness claims is grounded in the presumption that whether a trait is innate for a creature is a matter of the facts of that creature's development, independent of facts about the distribution of traits in the population. I motivate this presumption with the following thought experiment.

Suppose that a human embryo is placed in an incubator that simulates typical human gestational conditions and is launched into space. The incubator reaches a sufficiently earth-like planet just as the fully-formed infant is ready to emerge. This infant therefore exists as a population of one. Could this infant have innate psychological traits? It would seem that the standard interpretation of nativist cognitive science is committed to an affirmative answer. Several of the more well-known suggestions for innate traits (e.g., those involving language and theory of mind) are tricky insofar as they concern phenomena that are largely social, so let us consider the proposal that humans have innate beliefs about many properties of physical objects (Baillargeon, 2002).[5] Could this be true of our infant singleton? Suppose it is a fact about the infant that it does not expect solid objects to be able to pass through one another. It seems that this infant has innate beliefs about physical objects in exactly the same way nativists about intuitive physics usually mean the claim. Crucially, we do not need to advert to any facts about the distribution of intuitive physics

[5] One might well take a view about *belief* (e.g., one that takes natural language to be a necessary condition for having beliefs) that entails that the representations infants entertain about physical objects do not qualify. But the example works the same whether we call the relevant representations "beliefs" or not.

in the infant's population (for there is no such population); if the infant's beliefs are innate, they are innate because of *intrinsic* facts about its development.

There are construals of the notion of a *population* that do not require spatial contiguity, so one might object to the claim that the infant constitutes a population of one, and insist instead that it is a member of some population of humans back on Earth, (presumably the one that included its parents). Enter Swampbaby. Swampbaby exists in a universe in which humans never evolved, but when lightning strikes the swamp an embryo in an incubator spontaneously forms, an embryo that is identical to a typical human one. When Swampbaby emerges from the incubator, it has the usual beliefs about physical objects. Insofar as one is able to form intuitions about this bizarre scenario, I think it still seems right to say that Swampbaby has its beliefs about objects innately.

Thus it appears that the standard claims about innate characteristics offered in nativist cognitive science should be construed as claims about those traits arise in the development of individual creatures.

3.2.3 The explanatory two-step

Clearly, however, there is a tension between my claim that the proper explananda of innateness claims are individual-level phenomena (i.e., instantiations of psychological traits in particular individuals) and the idea that innateness claims play a role in a genuine scientific research program (nativist psychology and linguistics). For we might suppose that whatever "science" is, it fundamentally concerns the identification and explanation of interesting *generalizations*. And indeed, proponents of nativism in psychology often state

that the purpose of their investigations is to explain observed regularities and generalized phenomena, such as the fact that the vast majority of human children in a linguistic community end up speaking the same language as their parents and each other, and that these languages seem to share a common basic structure. If nativist psychology is, as a scientific discipline, concerned to provide generalized explanations of pervasive regularities, then how can it be true, as I have claimed, that the central claims of nativist psychology—that this or that psychological trait is innate—are aimed at such local, individualized phenomena as the emergence of particular traits in particular individuals? The concern here is that insofar as innateness claims concern extremely localized phenomena, they will, even if they are true, be of no more scientific interest than a claim about when and why a particular leaf fell from a particular tree. Indeed, we might wonder why, if the purpose is to explain individual-level phenomena, nativist psychologists and linguists bother to conduct experiments with large samples of individuals!

We can see how deep this tension goes if we reconsider the motivation for advancing the idea that innateness claims are directed at individual-level phenomena in the first place, namely the thought experiment provided above. If the lesson drawn from the thought experiment—that the very framework in which nativist cognitive scientists work is committed to the possibility of innateness claims being true at the individual level—is right, then the problem of identifying the proper explananda of innateness claims does not arise merely as a quirk of my own pet theory of innateness. Rather, the puzzle arises simply from theoretical commitments made by the theorists themselves.

Fortunately, however, I think this tension can be resolved. The key is to recognize two important points. One is that although claims about individual-level phenomena and claims about population-level phenomena are logically independent, there are nevertheless *evidentiary* connections between them in many cases (more on this point in a moment). The second point is that we can distinguish between the phenomena that particular scientific claims are intended to (and indeed can) explain, and the phenomena that motivate scientists to take up their investigations in the first place.

Consider the case of human language (the lessons will apply equally to other areas of nativist cognitive science): surely one reason why linguists have sought to study language, and why they utilize the particular tools they do in order to study it, is because it is a striking fact that there is so much regularity in human linguistic behavior. Moreover, it is evident that if linguistics failed to provide any insight as to *why* there is so much regularity, it would fail to satisfy the intellectual curiosity (or the practical interests, where these are involved) of those who study language—indeed, if linguistics provided no insight on this question, it is hard to imagine that it would have lasted very long as a research program. But, crucially, this is all consistent with the thesis that claims about *innateness*, per se, do not explain these regularities. For it can be true that while the overall explanatory itch that linguistics seeks to scratch concerns population-level phenomena, the explanatory work done by one of its central concepts (that of *innateness*) is confined to the individual level.

In other words, the fact that there is such striking regularity in human linguistic behavior can serve to tip off investigators that there is something worth studying in the

vicinity, but this does not mean that innateness claims must serve a *direct* explanatory role in explaining those regularities (though as I argue below, they may play an *indirect* explanatory role). We can state, therefore, that innateness claims are *made true* by the fact that individual creatures have the psychological properties that constitute innate traits, but they are *interesting* because, for reasons that make sense in light of our knowledge of biology, many members of a given species can be expected to—and manifestly do—share precisely these psychological properties.

This, then, is the explanatory two-step of nativist psychology: we are often motivated to investigate questions about the etiology of psychological traits by observations of striking regularities among individuals; that is, our explanatory questions begin at the population level. We then descend to the individual level to make claims about the etiologies of traits that inhere in individuals. Upon devising plausible explanatory stories for how traits have arisen in the individuals studied, we combine these stories with our knowledge in other areas (e.g., evolutionary biology) to make sense of the population-level phenomena we began with: our knowledge of biological species and of how evolution works suggests that we can expect the individual-level explanations we have devised (those that posit innate structures) to be highly conserved within the species and to arise from similar processes of biological development. In this way, we can generate satisfying explanations for phenomena of interest, even though we remain committed to innateness being an individual-level phenomenon. When the explanatory *modus operandi* of nativist cognitive science is seen in this light, we can recognize that nativists' claims that they are interested in explaining generalizations (or as it's often put, "human nature") are quite

correct, insofar as their discoveries of innate traits (in individuals) are generated in the overall service of explaining observed regularities.

3.2.4 A qualification

Before we proceed, I should note an important qualification to my claim that we are searching for *the* notion of innateness used in cognitive science. As stated, this claim is too strong, insofar as it implies that no cognitive scientist would ever have reason to employ a notion of innateness other than the one I defend (the Primitivist analysis). For as I argued in Chapter 1, scientific explanations are both *contrastive* and *pragmatic*: what counts as a good explanation in a given circumstance will depend on various features of the context, including the explanatory question being asked and the interests of those pursuing the explanation (which will combine to determine which alternative outcomes are relevant) (Garfinkel, 1981; Van Fraassen, 1980; for specific applications of this idea to *innateness* and related notions, see Birch, 2009; Gannett, 1999; Wendler, 1996).

Explanations in cognitive science that appeal to innateness will exhibit this same context sensitivity; different construals of innateness may well be appropriate in different explanatory circumstances even within a single scientific domain. The upshot is that any claim about what notion of innateness is appropriate in a given domain will have to be qualified to countenance the possibility that particular contextual factors will demand a different explanatory focus. Thus I do not wish to claim that I will be identifying the *only* notion of innateness that could *ever* be of use to cognitive scientists. There may well be particular explanatory contexts within the normal scope of cognitive scientific research that

call for the use of a different notion of innateness. There are too many pragmatic variables that go into determining the appropriateness of a particular explanation in a particular context for us to rule out, a priori, the possibility that a psychological researcher might have occasion to employ a notion of innateness that rests on, say, genetic determination.

What then of my thesis that purports to identify "the notion of innateness employed in cognitive science"? I think that properly qualified, the thesis is still true. I think that *most* mainstream cognitive scientific theorizing takes place in explanatory contexts where "innate" can be taken to mean "psychologically primitive." Moreover, as I demonstrate below, the Primitivist account comports extremely well with Poverty of the Stimulus arguments. Given that these arguments are one of the most central features of modern nativist psychological and linguistic theories, this suggests that Primitivism captures the notion of innateness utilized by cognitive science in its "core" or standard explanatory contexts.

3.3 The Primitivist view of innateness

The Primitivist family of views, again, seeks to capture the notion of innateness only as it appears in cognitive science. There are a variety of ways of spelling out the basic idea behind Primitivist views, but fundamentally these views hold that psychologically innate traits are those that emerge in development *without the need to process information from the environment*. A traditional way of formulating this idea is to suggest that what is innate is *not learned* (learning is, presumably, a way of processing information from the environment). Thus, if a child *learns* that water is wet (as a result of experiences with water,

or as a result of being taught this by her parents), her belief that water is wet is not innate. On the other hand, if it is true that no such learning process is needed for the child to understand that solid objects cannot pass through one another (if, for example, the child has this belief even prior to any experience with solid objects), then according to this approach the belief is innate.

3.3.1 Samuels and "psychological processes"

Picking up on this intuitive idea, Samuels (2002) has offered the notion of a *psychological process* to capture the sort of etiology innate traits are supposed to lack. That is, according to Samuels innate cognitive traits are psychological structures whose emergence cannot be explained by appeal to any psychological process (where again, what is distinctive of a psychological process is that it involves representational states). In other words, innate psychological traits are those whose acquisition has no psychological explanation—they are *psychologically primitive* (i.e., they exist at the psychological level and can play a role in psychological explanations, but in order to explain their presence we must appeal to phenomena outside the scope of psychology—presumably biological phenomena).

Returning to the example above, a child's belief that water is wet is not innate if it arises as a result of an inductive generalization (a psychological process) over a series of individual experiences with water (i.e., token representations with the content "*this* water is wet" are combined with inferential rules to yield a standing belief-like representation

with the content "water is wet").[6] The child's beliefs about the properties of solid objects, however, are innate because no such psychological explanation can be given for them—they arise, let's suppose, simply as a result of biological maturation in the brain.

As we'll see, I think Samuels' specific formulation of innateness as primitiveness has several shortcomings, but it's worth first defending this notion against an objection raised by Mameli and Bateson (2006). The objection is that the idea that innate traits are those that have no psychological explanation is unacceptable because it defines innateness *sociologically*. The worry here is that if we define innateness in terms of what "psychology explains (or can explain)," then what counts as innate will reflect not only the biological and psychological properties of the organism, but also highly contingent facts about the sociological organization of science—in particular which researchers, institutions, experimental methods, research questions, and so forth are counted as part of "psychology."

However, the assumptions outlined above in Section 3.2.1—that psychology does not reduce to some lower-level area of study (e.g., physiology or physics) and that it studies a well-defined (again, non-reducible) kind of phenomenon (mental representation)—should make it clear why the objection is unfounded. For as I—and, to all appearances, Samuels—are using the term, "psychology" does not refer, as Mameli and Bateson assume, to the sociological institution typified by university departments with the label "psychology," but to a scientific discipline defined by its interest in a particular kind of

[6] Of course it is expected that child would engage in this cognitive process *implicitly*, rather than explicitly representing the steps of the inferential procedure.

phenomenon, that of mental representations. This phenomenon would exist even if there had never arisen a specific sociologically-defined scientific field dedicated to studying it, and so its status as part of the "furniture of the world," amenable to scientific study, is independent of any particular sociological facts of how we go about studying it. Thus, the claim that innate traits are those that are "not explainable by psychology" should be understood simply as a claim that a certain kind of natural phenomenon does not feature into the causal-etiological explanations of those traits, not as a claim (as Mameli and Bateson construe it) that certain kinds of scientists and not others are willing or able to explain the traits in question. This objection, therefore, rests on a confusion about the central claim of Primitivism.

While Samuels' formulation of innateness as primitiveness survives this objection, it is subject to other difficulties. Specifically, the notion of a "psychological process" is too broad a characterization of the sorts of processes we wish to exclude from the acquisition of innate traits. To see this, consider how the psychological/non-psychological distinction might be applied to individual perceptual events. A core question in scientific psychology is *at what point in the perceptual processing stream does something deserving of the term "representation" arise*? That is, the process of perception is, if nothing else, a process by which non-representational sensory activation (e.g., luminance gradients on the retina) is used to generate representational states (e.g., a representation of a flower in one's visual field). It is (well) beyond the scope of this chapter to take up the issue of where in the processing stream representational content first appears, but the key point is that at *some* point in the perceptual processing stream, there must arise a representation where before

there was none. Such a representation (say, of an *edge*) would be a psychological entity, but it would have no psychological explanation; there would be no need to appeal to any psychological (i.e., representational) phenomena to explain how the representation was acquired—the explanation is couched entirely in physical (or at least physiological) terms.

Thus, the "no psychological explanation" approach to defining innateness—which excludes representations of empirical matters—would therefore seem to count the representation of an edge as innate. For even though some of the information used to generate the representation is experiential (or empirical, or exogenously provided), the edge-y representation still lacks an etiology in psychological processes (i.e., processes involving representation). Thus, Samuels' notion of Primitivism, which again centers on the very broad notion of *psychological process*, yields an unacceptable result: every single instance of perception that yields representational content yields an innate psychological state!

3.3.2 An alternative version of Primitivism: the "starting state" hypothesis

This shortcoming in Samuels' version of Primitivism—which relies on the notion of a *psychological process*—pushes us in the direction of what I take to be the most promising version of the primitivist strategy, namely the idea that innate features of the mind constitute a *starting state* for cognitive systems, i.e., the state of those systems prior to their engagement with information from the environment (Gross & Rey, 2012). These states need not be *temporally* initial—the organism may need considerable time to mature before certain psychological systems come online—and the starting states for different

systems may emerge at different times (so some innate traits might be present at birth, but others might not emerge until significantly later). Rather the idea is to identify an initial state of a cognitive system that is *explanatorily relevant* for the project of explaining the acquisition of some psychological structure (where "psychological structures" include, for example, token concepts, token doxastic representations, and large-scale competences like the ability to recognize whether a given linguistic string is consistent with a hypothesized grammar). The (innate) starting state of a system is a state of that system that successful explanatory projects in the study of cognitive development take as primitive for the purpose of explaining how the mature psychological states of the organism emerge from the organism's interactions with its environment.

To take a simple example, suppose that, as has been suggested, human infants innately possess a representations of all the phonemes that might appear in the totality of humanly-possible languages, as well as a system for "pruning" this repertoire in response to early linguistic experience (e.g., by collapsing phonemic distinctions that do not carry meaning in the ambient language). (This theory is meant to account for the fact that all infants are sensitive to the difference between, for example, the *l* and *r* sounds in English or the hard and soft *v* sounds in Spanish, but (generally speaking) adult Japanese speakers cannot perceptually distinguish English *l* and *r*, and adult English speakers do not perceive the difference between hard and soft Spanish *v*.) Under the "starting state" idea of innateness, what it means to say that the repertoire of phoneme representations and the pruning system are innate is that successful explanations of human competence when it comes to phonemes (e.g., a particular learner's pattern of sensitivities to phonemic

distinctions at various points in development) operate by *assuming* the existence of the complete representational repertoire and the pruning system, and demonstrating how these elements *plus* early linguistic exposure yields the mature (i.e., "pruned") repertoire. The representations and the pruning system are taken as *psychologically primitive* in generating a satisfactory explanation of an interesting psychological phenomenon, in that they are the most fundamental *psychological* components of that explanation—the remaining components of the explanation are either further, *non-primitive* psychological structures (e.g., representations of the phonological and/or semantic properties of incoming linguistic strings) or *non-psychological* elements (such as, perhaps, facts about physical maturation of the brain and/or perceptual apparatuses).

As detailed above, I take the "starting state" approach to the Primitivist account of innateness to be the most promising version, given the problems we observed with Samuels' "no psychological process" formulation. Importantly, however, I am more interested in establishing the general promise of the Primitivist strategy than with narrowly defending any particular version of it. It may well turn out that unforeseen problems will doom the "starting state" account I've offered, and that the problems noted with Samuels' formulation can be solved, such that the best way of cashing out the promise of Primitivism is to move back towards *psychological processes* as the key theoretical notion. With this in mind, I turn now to a discussion of the role of innateness in some of the theories, arguments, and debates in contemporary cognitive science. The subsequent analysis will generally be couched in terms of the "starting state" version of Primitivism, but will not necessarily eschew discussion of *psychological processes* altogether.

3.4 Case Study: Innateness and the poverty of the stimulus

What follows is an analysis of nativist theory in action. I scrutinize both an argument made by theorists in the nativist tradition in linguistics—in which the notion of innateness plays a prominent role—and the debate with anti-nativists that has ensued.[7] The motivation for proceeding in this fashion is two-fold. First, looking at the ways in which innateness figures into how nativist theorists construct an argument—what they take to be evidence for what, what they believe follows from the claim that something is innate, and their (perhaps implicit) assumptions about how language acquisition occurs—will give us insight into what is at the core of the notion of innateness to which practicing researchers commit themselves. Second, on the assumption that when theorists in different camps argue about whether something is innate, they are arguing about something *substantive*—that is, they do in fact have a common notion of innateness in mind and the debates are not merely verbal[8]— looking at what actually underlies a disagreement about innateness should give further evidence about what comprises the core of the notion of innateness employed—and fought about—in cognitive science.

[7] The terms *nativist* and *anti-nativist* are somewhat infelicitous, in that all sides in these debates must—and generally do--countenance a substantial degree of innate cognitive structure in order to explain human learning (these debates, as noted by Cowie (1998), are, in many cases at least, probably best understood as conflicts regarding the domain-general versus domain-specific nature of learning processes). Nevertheless, I'll stick to the *nativist/anti-nativist* nomenclature for reasons of familiarity.

[8] It is of course a live possibility that some or all nativist debates *will* turn out to be merely verbal; nevertheless, it seems to me most methodologically expedient to assume that these debates are in fact substantive until such an assumption proves genuinely untenable. (Cf. Samuels', 2002 "Significance Constraint".)

3.4.1 Poverty of the stimulus arguments

Our case study centers on an exemplar of the well-known family of nativist arguments known as *poverty of the stimulus arguments* (henceforth POSAs), and the loci of disagreement surrounding the argument. A few introductory remarks about POSAs are in order. As they are typically conceived, POSAs seek to establish some sort of innate structure to the mind by emphasizing that the knowledge acquired in some developmental process cannot be reached from information available in the environment via general learning procedures. While POSAs can be (and are) applied in a broad range of psychological-developmental contexts, their most well-known application is in the domain of language acquisition, from which our case study is drawn.

Now, there is some measure of disagreement surrounding just how POSAs are supposed to work. Points of disagreement include (*inter alia*): (i) what conclusions POSAs aim to establish; (ii) what conclusions POSAs *can* establish; and (iii) which premises often included in POSAs comprise the "core" of the argument, and which are supplementary or are best taken as support for other propositions (which may be theoretically related to, but are ultimately distinct from, the conclusions putatively established by the POSA). Thus, although reference is often made to "*the* argument from the poverty of the stimulus," it is not clear whether there is any such unique argument. Nevertheless, following Laurence & Margolis (2001, p. 220), I suggest that POSAs can be usefully viewed as series of "related, mutually supporting claims" that tend to follow a common inferential structure.

That there are actually two related but distinct forms of POSA: one (largely) a posteriori and the other (largely) a priori (Cowie, 1998; Samuels, 2002). Our case study

falls into the a posteriori category, but it's worth first outlining the general contours of the a priori version as well. The more a priori arguments tend to emphasize that only *negative evidence* about what structures are *not* part of the language can be of help to learners in particular situations. Specifically, if a learner settles on a grammar that is consistent with the sentences she has thus far encountered, but that is ultimately incongruous with the correct grammar[9], only negative evidence can steer her back towards the correct principles. For example, suppose a learner of English forms the hypothesis that expressing a grammatical subject is optional, as it is in, say, Spanish (e.g., *Ella fue al teatro* and *Fue al teatro* are both acceptable ways of saying 'She went to the theater'). This hypothesis is consistent with any sentence of English the learner will have heard. This learner's grammatical competence would count **Went to the theater* as a grammatical sentence of English, even if she had never heard a construction with a null subject or produced one herself. Thus, without negative evidence (i.e., evidence that sentences with a null subject are ungrammatical in English), the child will have no way to recover from the mistaken hypothesis.

Negative evidence could take a variety of forms, including spontaneous explicit instruction about ungrammatical forms or corrections of children's ungrammatical utterances. But, according to this version of the POSA, such negative evidence is generally not available, and so if children arrive at the correct grammar (as they invariably do), then innate constraints, again, must have had a hand in preventing the child from settling on a mistaken hypothesis about the language. Whether negative evidence is typically available

[9] Here "correct grammar" simply means the grammar of other speakers in the learner's linguistic community.

to language learners is, of course, an empirical matter (this is why this variety of POSA is *largely* a priori). There are those who are more optimistic than Chomsky and his followers about the availability of negative evidence (e.g., Cowie, 1998), but proponents of this sort of argument emphasize that given the *consistency* with which children arrive at the correct grammar, the appropriate negative evidence must *always* be available to *each* learner (and for each learning problem that requires it); this seems implausible given the diversity of childrearing methods, including differences in adults' propensity to correct children's ungrammatical utterances (Laurence & Margolis, 2001).

More generally, a priori POSAs seize on the Goodmanian insight that our experiences are consistent with a veritable infinitude of "grue-some" concepts that are never even considered. And just as no child learning color words considers the hypothesis that the term "green" might mean "green if observed before year 2500 and blue if observed thereafter," there is no evidence that any child considers such hypotheses as that the subject of a sentence may be fronted only if the number of words in the sentence is even, or less than nine, or any of an infinite number of other bizarre alternatives. The fact that we all end up using concepts like "green," rather than "grue," and that the variation in human languages is but a tiny subset of the infinite number of logically possible languages, suggests significant innate constraints on these acquisition processes.

In contrast, in presenting an *a posteriori* POSA, linguists studying language acquisition will compare what language learners appear to know about some specific grammatical feature of the ambient language with what could plausibly (or even possibly) be extracted from the evidence available to them (i.e., the linguistic strings to which they

are exposed). If learners appear to know more than could be extracted from the available data, then it is concluded that some form of innate constraint has guided the learner in arriving at the correct grammar. (The next subsection explores in detail a specific example of such an argument.)

It's important to note that although, again, POSAs concerning *language acquisition* might be considered the arguments' paradigmatic exemplars, these arguments—particularly the a posteriori variety—are widely applicable and widely utilized in the study of cognitive development. In many cases, the argument is pursued not by claiming that it would be *impossible* for organisms to develop an observed competence with only general-learning mechanisms, but rather by comparing the sophistication of the competence against the timeline of its emergence. It is difficult to deny that certain competences that are present literally at birth—such as face recognition in human neonates (Bushneil, Sai, & Mullin, 1989) or predator recognition among newly-hatched turkeys (Goth, 2001) were not acquired by learning. But even when the presence of a competence cannot be established quite this early—such as understanding of false beliefs, which appears at least by 15 months (Onishi & Baillargeon, 2005), or understanding of addition and subtraction at 5 months (Wynn, 1992)—a plausible a posteriori POSA can still be constructed if the competence is sufficiently sophisticated that it is unlikely it could have arisen from general-learning processes (for a series of such arguments, see Spelke, 1994). Thus, a posteriori POSAs feature prominently in nativist theorizing throughout cognitive science. What I have to say about *language-centered* POSAs applies equally to these further uses of the arguments.

3.4.2 Case study: NP structure and anaphoric use of "one"

A paradigmatic a posteriori POSA is offered by Lidz, Waxman, and Freedman (2003). Lidz et al.'s analysis concerns the structure of noun phrases (NP) and the rules that govern the referent of the pronoun *one*. Consider sentence (1):

(1) I'll play with this red ball and you can play with that one.

For adults, the referent of *one* is *red ball* rather than just *ball*. But this is only possible if *red ball* is treated as a distinct constituent of the NP *the red ball*, as can be seen in the comparison between two possible structures of NP in Figure 3.1:

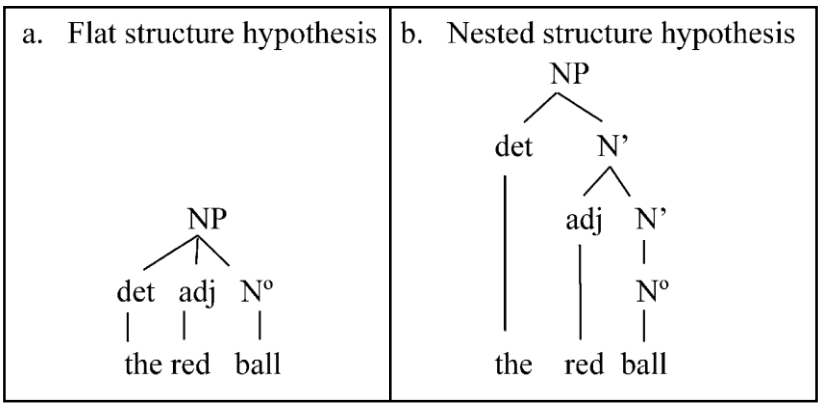

Figure 3.1: Hypotheses about the structure of NP

Let us call the hypothesis that NP has structure (a) (flat structure) **H1**, and the hypothesis that NP has structure (b) (nested structure) **H2**. Only **H2** treats *red ball* as a constituent, and both structures are, as Lidz et al. note, consistent with many sentences of English (e.g., *The red ball is on the floor*). Thus there is a question as to how and when children manage to rule out **H1** and settle on **H2** as the correct structure of this sort of NP

in English, and how and when they come to know that *one* is anaphoric to *red ball* rather than *ball* (it is assumed that all children do so, as there is no evidence that idiolects differ in this respect).

Lidz et al. investigated whether 18-month-old infants have command of the hierarchical structure of NP and the anaphoric properties of *one*. Infants in the study were shown a picture of an object (e.g., a bottle, shoe, or bear) at the same time they were presented with a sentence that referred to the object with a determiner phrase that included both an adjective and noun (e.g., *Look! A yellow bottle.*). The infants were then shown a picture containing both the original object and another object of the same type but a different color (e.g., a yellow bottle along with a blue bottle). Infants in the control group heard a neutral phrase (*Look! Now what do you see?*), while infants in the experimental group heard a phrase with an anaphoric use of *one* (*Look! Do you see another one?*). Infants in the control group looked longer at the novel stimulus, while infants in the experimental group looked longer at the familiar stimulus. Previous studies having established (i) that infants tend to prefer looking at novel stimuli, all else being equal, and (ii) that infants tend to prefer looking at stimuli that match the linguistic input (if such is available), the researchers concluded that these 18-month-old infants understood that *one* is anaphoric to, e.g., *yellow bottle* rather than *bottle*, which in turn suggests that they have settled on **H2** regarding the structure of NP.

Lidz et al. also performed a corpus analysis (using databases of child-directed speech) to determine the availability of data that would help learners establish the correct

structure of NP and the correct referent of *one;* such data could come, for example, in the form of sentence (2) being uttered in a situation in which Max has a blue ball:

(2) Chris has a red ball but Max doesn't have one.

According to the corpus analysis, sentences like (2) occur at about the same frequency as do ungrammatical utterances containing anaphoric uses of *one*, making such sentences, from the learner's perspective, indistinguishable from mere noise in the linguistic data. Given that infants at this young age appear to have ruled out certain logically possible hypotheses about the syntactic structure of English, and given that the data available to them appears not to contain sufficient evidence for an unbiased learner to extract the relevant structure, Lidz et al. conclude that learners must not have considered **H1** in the first place; that is, there are *innate* constraints on the hypothesis space of language learners.[10,11]

[10] While Lidz et al. conclude that **H1** is not even considered, this strong conclusion does not seem warranted by their data. It might be the case that **H1** *is* part of the hypothesis space of the learner, but is assigned a much lower prior probability than **H2**, such that structure (b) constitutes a "default" hypothesis, and direct evidence would be needed to steer the learner towards **H1** (e.g., sentences of the form *Chris has a red ball and Max has one too*, in a situation in which Max has a blue ball). Nevertheless, these prior probabilities would constitute innate constraints on the hypothesis space just as much as the total exclusion of structure (a), so Lidz et al.'s overall thesis would be unaffected by the weakening of this conclusion. Hence in what follows I will refer to their conclusion that "**H1** is absent from the hypothesis space," but this can be read as "**H1** is absent from the hypothesis space or assigned low probability."

[11] It's worth noting that the specific grammatical principle chosen by Lidz et al. as the basis for their POSA is perhaps not the ideal case for establishing the innateness of a grammatical principle, since it is not clear that the constraints on the referent of anaphoric *one* are entirely grammatical, rather than pragmatic (i.e., it may be possible for adult speakers of English to envision scenarios in which *one* could be understood to refer to "ball" rather than "red ball"). Other examples (such as children's understanding of the constraints on the use of contractions like "wanna"—e.g., Crain & Thornton, 1988) might serve as more compelling elements of the nativists' arsenal. But the present example has the advantage of being the subject of an especially clear and tractable debate between nativists and non-nativists.

What can we say about the explanatory role of innateness at work in this argument? To begin with, let us apply some of the lessons we learned earlier about contrastivity and pragmatics in explanation. Our earlier remarks on contrast classes can assist us in identifying what, exactly, is the *explanandum* in this instance. It might be thought that, as it's sometimes put, the phenomenon to be explained in language acquisition is the child's coming to know the correct grammatical principles.[12] But given the contrastive nature of explanatory questions, we know there must be something more to the story. What, in this case, is the range of relevant alternatives to the explanandum (i.e., *the child comes to know the correct grammatical principles*)?

To answer this question, notice that as Lidz et al. frame the learning problem, learners must decide between two different hypotheses about the structure of NP and the anaphoric referent of *one*. The data that would allow learners to make the correct choice between these two hypotheses are, ostensibly, not available. If the matter were left at that, it would be mysterious why learners consistently end up knowing the correct grammatical principles rather than settling on one of the two different hypotheses according to chance. Thus, it appears that what Lidz et al. take to be the relevant explanandum—that which the innateness claim is supposed to explain—is the fact that learners end up with the grammar they do *rather than any of the other logically possible alternative grammars consistent with the linguistic data available to them*. When framed in this way, it is clear how Lidz et

[12] In saying that infants—or adults—"know" principles of grammar, I of course do not mean to suggest that they consciously represent these principles. I use the term "knowing grammatical principles/rules" as shorthand for "having a language faculty that generates and interprets linguistic strings in accordance with these principles."

al.'s answer—that there is some antecedent structure[13] to the learners' hypothesis space that makes the adoption of **H1** impossible or unlikely— constitutes an explanation of the revised (contrastive) explanandum: if these alternative grammars are not part of the hypothesis space (or are assigned very low prior probability), then it is no mystery that learners do not adopt these grammars with anything near the frequency that would be predicted if, e.g., **H1** and **H2** were assigned equal prior probabilities in the hypothesis space.[14]

The addition of this contrast space to the explanandum may seem trivial, or so obvious as to not be worth stating, but note that there are, in fact, contexts in which we might well consider different contrast classes to be the relevant ones. Consider two other contexts in which we might wish to explain a child's adopting a particular grammar, but with a different comparison in mind instead. If, for example, we are comparing the same normally-developing child to a child who—like the tragic case of Genie—has been given virtually no linguistic input at all, we might well ask the question, "Why did this child develop the grammar she did, *rather than developing virtually no grammatical competence*

[13] Note that there is an inference on Lidz et al.'s part from "antecedent structure" to "innate structure" (i.e., it is possible that children's language acquisition device acquires the structure required to adopt the correct grammar in the face of ambiguous evidence via some process that is inconsistent with that structure's being innate). Whether this inference is valid turns on what exactly we take psychological innateness to be (e.g., the inference is valid if "innate" amounts to "not learned" and there is good reason to believe that the "antecedent structure" could not have been learned—for example, if 18 months is insufficient time for the required learning to occur). But of course the proper construal of psychological innateness is precisely what we are hoping to discover. Since my approach relies on analysis of the ways nativists *use* the notion of innateness, I will, for the time being, take Lidz et al.'s assumption—that the antecedent structure must be innate structure—at face value.

[14] It is of course possible that some of the infants in Lidz et al.'s study had in fact settled on **H1** for the time being (the prediction for such infants would presumably be that at the test phase, when they are given a phrase with anaphoric use of *one*, they would attend to the novel stimulus, given the baseline preference for novelty and the fact that both stimuli would match the linguistic input under **H1**). Lidz et al. do not report whether any infants displayed this pattern, but in any case there would be no principled way to distinguish any such instances from experimental noise—babies are, after all, a very noisy bunch.

at all?" Or, if we are considering the one known case of a particular genetic mutation that affects linguistic capacities specifically, we might wish to compare our linguistically-normal child, possessed of a non-mutated copy of the FOXP2 gene, with a child who has the specific language impairment (SLI) that is associated with having a mutated copy of FOXP2 (see Lai et al., 2001, and Enard et al. 2002 for details about FOXP2 mutation; Watkins et al., 2002, for a description of the associated phenotype). Here we might wish to ask: why did this child develop the grammar he did *rather than one that includes irregular use of inflectional morphemes?*

Again, according to our contrastive account of explanation, the contrast class we select for an instance of a child adopting a particular grammar will inform and constrain the sorts of explanations that are appropriate. In the case of comparing a normally-developing child to one deprived of linguistic input, citing the mere presence of *any* linguistic data, or (depending on exactly what our interests are) the presence of emotional nurturance might serve as a proper explanation. In the FOXP2 case, we would presumably cite the presence of a particular genetic factor as the explanation of the child's grammatical competence.[15]

Returning to the original explanatory context, however—that of Lidz et al.'s hypothesis, in which the contrast space includes children adopting faulty grammars consistent with the primary linguistic data—we can say the following: *insofar as* most linguistic theorizing takes place in a shared explanatory context, one in which the relevant

[15] Skepticism about "genetic determination" as a *general* explanatory theory—a skepticism I espouse elsewhere—is of course consistent with acknowledging that citing genetic factors can constitute a good explanation in some contexts.

alternatives to learners arriving at the correct grammatical principles involve learners arriving at *other grammatical principles consistent with the primary linguistic data*, then innateness claims will have the same explanatory role as they have in Lidz at al.'s POSA: they make the otherwise mysterious fact that learners adopt the correct grammatical principles un-mysterious.

3.4.3 The debate

Regier and Gahl (2004) and Foraker and her colleagues (Foraker, Regier, Khetarpal, Perfors, & Tenenbaum, 2009) have both offered critiques of Lidz et al.'s analysis of young children's competence with anaphoric uses of *one*. While these critics do not dispute that even young children possess such competence, they do argue that it is not necessary to posit the particular innate constraints postulated by Lidz et al. in order to explain this competence. Rather, the critics argue, this competence could be acquired by a Bayesian learner who assigned equal prior probabilities to the flat-structure hypothesis (**H1**) and the nested-structure hypothesis (**H2**).

Both Regier and Gahl and Foraker et al. object to Lidz et al.'s acquisition story, arguing that Lidz et al. have overlooked the possibility that the *absence* of certain grammatical constructions from the linguistic input available to learners could provide the necessary evidence to rule out **H1** and settle on **H2**. Both sets of researchers designed a model of a hypothetical language learner whose task was to determine the proper use of anaphoric *one* and the proper structure of NP. Both models gave equal prior weight to the nested-structure hypothesis and the flat-structure hypothesis (i.e., the hypothesis that in the

sentence "I'll play with this red ball and you can play with that one," *one* may refer to a blue ball), were equipped with a Bayesian procedure for evaluating hypotheses, and were fed a standard diet of child-directed speech (in Foraker et al.'s case, the input to the model was a subset of the corpus analyzed by Lidz et al.). Both models learned the correct syntactic principles with relatively limited input, owing, the authors say, to the *absence* from the input of certain grammatical constructions or utterance-environment pairings that would be predicted by the flat-structure hypothesis (but not the nested-structure hypothesis). These researchers conclude that Lidz et al. were mistaken in their assertion that this aspect of children's syntactic competence could not have been extracted from the available linguistic evidence, and, consequently, in their application of the (a posteriori) POSA. According to these critics, a hypothesis space that assigns low or zero probability to the flat-structure hypothesis is not necessary to explain learners' convergence on the correct grammatical principles.

Let us take stock of the debate between these two camps in light of our earlier discussions of the nature of scientific explanation and of the role of the concept of innateness in this realm of psychological research. Importantly, my aim here is not to declare a winner, but rather to identify precisely *what* is under dispute.

To begin, let us note a number of points that are *not* under dispute. First is the claim that the process of language acquisition—or indeed, the acquisition of the syntactic structure of NP—is heavily constrained. Both Regier & Gahl and Foraker et al. openly acknowledge that their models employ a variety of built-in assumptions (e.g., Bayesian reasoning principles) without which learning would be impossible (or at least substantially

more difficult), as well as various kinds of specifically-linguistic knowledge (e.g., the ability to recognize syntactic types, such as nouns, modifiers, and so on). Likewise, there is no dispute over whether *something* is innate; Lidz et al.'s critics do not say this explicitly, but it is widely acknowledged that *some* innate structure is necessary for any learning to take place at all (it is plausible, for example, to read the critics as supposing that Bayesian reasoning principles are innate). Rather, the debate between Lidz et al. and their critics appears to center on two key questions: (1) whether the procedure for acquiring the syntactic knowledge in question is *language specific* (Lidz et al.) or *domain general* (Regier & Gahl, Foraker et al.); and (2) the respective prior probabilities assigned to different elements in language learners' hypothesis space.

We will focus on the latter point of contention, as this is the one that most directly and immediately concerns questions of innateness. Lidz et al.'s critics, again, do not say what they *do* think is innate, but let us suppose, for simplicity's sake, that they would grant the existence of an innate hypothesis space of possible grammars. We may then rephrase point (2) to make explicit the role of innateness in the dispute: according to Lidz et al., human children innately possess a hypothesis space that either lacks or assigns low probability to the hypothesis that NP has a flat structure; according to their critics, the innate structure of children's hypothesis space is such that the flat-structure hypothesis is both present and assigned the same prior probability as the nested-structure hypothesis. In other words, these theorists disagree about the *innate structure* of the hypothesis space of human language learners.

We can quickly put to rest any worries that these theorists are merely talking past one another because they are inhabiting different explanatory contexts (e.g., having in mind different contrast classes for the observed outcomes), as was the case between Willie Sutton and his priest. Rather, all appearances are that these opposing theorists are operating within a shared explanatory context and are seeking to explain the same phenomenon: why do learners, from a very early point in development, embrace the correct syntactic principles *rather than any of the other principles consistent with the data?* Again, according to Lidz et al. the explanation for this fact is that the "other principles" are innately excluded from the hypothesis space. According to their critics, however, the proper explanation is that human infants are good Bayesian reasoners, and there is no need to posit the innate constraints described by Lidz et al.. Hence we have a straightforward disagreement in which one side believes that a particular trait is innate (a hypothesis space with structure *S*), with the opposing side denying that this same trait is innate. The upshot is that both Lidz et al. and their critics appear to have in mind a particular notion of innateness, and are merely in disagreement about what the empirical evidence suggests about which things meet the criteria for innateness.

Given that we have a straightforward disagreement about what is innate (and not about what is *innateness*), such that the two opposing sides mean the same thing when they talk about innateness, we can ask what, precisely, these two sides mean. I think that the core disagreement between these two camps can be captured by the "Psychological Primitiveness" account of innateness. This view, again, holds that whether a trait is innate in the psychological sense depends on whether *processing of information from the*

environment is necessary for the emergence of that trait, or if instead that trait may be counted as part of the "starting state" of a system in the creature's psychology.

My claim here is that the interlocutors in the debate over acquisition of the structure of NP are, as the Primitivist analysis of innateness would have it, arguing about what the starting state of the human language faculty (or at least, those aspects of it that underlie understanding of anaphoric *one*) looks like. Again, the explanandum here is the emergence of syntactic competence in human infants—a state of affairs in which the language faculty operates in accordance with the nested-structure hypothesis about NP. The debate between the nativists and anti-nativists concerns whether this phenomenon is best explained by positing an initial (or *primitive*) state in which only the nested-structure hypothesis is considered and the flat-structure hypothesis is absent (or assigned much lower probability)—as the nativists would have it—or by positing an initial state in which the nested- and flat-structure hypotheses are both present and assigned equal probabilities—as presumed by the anti-nativists. Nativists think the preference for the nested-structure hypothesis is innate because they think an adequate explanation of the children's observed competence requires taking this preference as a primitive, irreducible component of the language acquisition process. Anti-nativists consider the preference *non*-innate because they think the best explanation for the children's competence *derives* the "mature" state of the 18-month-olds' language faculty from the combination of experiential input and a *yet-more-basic* state of the system (a state in which the nested- and flat-structure hypotheses are assigned equal probabilities).

So it appears that the Primitivist account of innateness, and in particular the idea that innate traits are (part of) an explanatorily-relevant starting state, captures what is at issue in the debate between nativists and anti-nativists about grammatical competence. It's worth noting that this result holds even if we consider the debate through the lens of Samuels' original "no psychological process" account. That is, the debate can be seen as a disagreement about the process by which the language faculty has arrived at its "mature" structure in the 18-month-olds, and in particular whether an appeal to *psychological* processes is necessary. A full analysis of the notion of a "psychological process" in beyond the scope of this chapter, but I take it that if anything is a psychological process, checking hypotheses against incoming empirical evidence according to Bayesian principles is one. Hence, according to the anti-nativists, the explanation for the structure of the language faculty in its mature state appeals crucially to a psychological process (i.e., Bayesian reasoning). According to the nativists, on the other hand, the explanation for the structure of the language faculty (at least with respect to the structure of NP) makes no appeal to any such process. We do not, according to the nativist, need to appeal to any cognitive process to explain why learners' hypothesis space comes to assign low or zero probability to **H1**, for this was the state of the hypothesis space before the language acquisition process even began.

Thus, regardless of the particular version of the Primitivist account we appeal to, the Primitivist strategy succeeds in capturing the contours of this debate about innateness, and shows us why, for example, the participants in the debate would appeal to the sorts of evidence they do in building a case for their respective positions. In light of the foregoing,

I contend that the central notion at issue in the debate between nativist and anti-nativist theorists in language acquisition is that of the *psychological primitiveness* of particular features of the structure of the language faculty.

3.5 Conclusion

Assuming my analysis is correct, we have now identified a notion of innateness that accounts for the explanatory practices of nativist cognitive science, and characterizes what is at issue in debates about whether particular psychological traits are innate. Notably, the explanations given by nativist cognitive scientists (e.g., Lidz et al.), do not make any overt appeal to *biological* properties of the children, such as their genetic makeup. So, as noted in the introduction, the analysis offered here lends no particular support to the idea that innateness should be defined in biological terms (e.g., as "genetic determination"), as is often suggested. Many explanatory projects in nativist theorizing seem to get by just fine without referencing genetic facts at all.

Of course, this does not preclude the possibility that nativist scientists do in fact think of innateness in genetic terms (indeed, as we will see in the next chapter, many say explicitly that they do), nor does it preclude the possibility that psychological primitiveness is just a particular species of genetic specification, such that the Primitivist analysis can be subsumed under a more general biological account of innateness. But in light of the fact that successful explanations can be devised in nativist theories without appealing to genetic facts, we would need an *argument* for why we should collapse psychological innateness under a genetic theory. The next chapter is dedicated to examining such arguments.

CHAPTER 4: THE DISUNITY OF INNATENESS

4.1 Introduction

The previous chapter argued for a particular analysis of *innateness* in cognitive science. This chapter explores whether this analysis can be subsumed under a more general notion of innateness, one according to which the notion of innateness employed in cognitive science is the same notion employed in biology. As hinted at in previous chapters, I do not believe it is possible to identify a singular analysis of innateness that can do all the explanatory work required of it in both of these domains. The remainder of this chapter is dedicated to offering arguments for the position that whatever analysis we might give for innateness in biology, the primitivist analysis of innateness in cognitive science is an independent explanatory notion.

Let us begin by spelling out—in an intuitive and preliminary fashion—the view I aim to argue against, namely, that "innateness" means in cognitive science the same thing it means in biology. In order for this idea to work, we would need to identify some property of biological systems that distinguishes innate from non-innate traits, and then we would need to show that the set of *psychologically primitive* traits also possess this further property. If this were accomplished, we would have shown that the primitivist analysis of innateness in cognitive science is in an important sense unnecessary: it picks out innate psychological traits, but it needlessly clutters the explanatory landscape, for these traits are *also* picked out by a more fundamental biological property.

What might this property be? The most obvious answer is that organismal traits—biological or psychological—are innate in virtue of the fact that their presence is in some way explained by the activities of the organism's genes. And in fact, some of the leading theorists of nativist cognitive science have espoused just such a view. Fodor, for example, says that nativist claims in cognitive science and in biology share the basis feature of positing that

> there is a characteristic human phenotype…that *can be attributed to a characteristic human genetic endowment*…the idea that some properties are *significantly genotypically determined* is now deeply scientifically entrenched; to that extent, biology seems to be in the process of constructing a concept of innateness that saves many of the rationalists' paradigms. (Fodor, 2001, p. 102, emphasis added)

Fodor's statements here seem to accord with a common-sense view of biological development, according to which genes count as a special kind of explanation for particular biological outcomes. It seems natural to suggest, therefore, that the notion of *innateness* can be characterized in such terms. Similarly, we hear from Chomsky that innateness (or as he puts it, "Platonic 'remembrance,' " 1990, p. 633) is to be thought of "in terms of the *genetic endowment*, which specifies the initial state of the language faculty, much as it determines that we will grow arms not wings" (1990, p. 633).

As we can see in the quotes from Chomsky and Fodor, a variety of locutions might be (and are) used to describe the kind of explanatory relation that genes are supposed to bear to innate traits: such traits are *attributed to* a *genetic endowment*, are *significantly genotypically determined*, and are *specified* by the genetic endowment. But we might ask just what these descriptions amount to. What does it mean to say that a trait is *significantly*

genotypically determined (as opposed to, say, *moderately* genotypically determined)? How do genes *specify* things? In the surrounding passages, Fodor acknowledges the considerable difficulty of answering such questions, but remains confident that answers are there for the discovering, if only because "it would be unlikely if the notion of innateness according to which [claims that traits are largely genetically determined] are true will turn out to be dispensable for the larger purpose of biology" (2001, 102).

The remainder of this chapter is dedicated to considering what kind of case can be built for the Fodor/Chomsky view, by exploring various ways we might attempt to make good on the idea that innate psychological traits can be identified via their special explanatory relationship to the genome. The arguments presented herein will build upon those offered in Chapter 2, where I provided reasons to doubt that genetic factors can, *in general,* be granted special explanatory status in the emergence of even non-psychological traits. My critique of the proposal to subsume psychological innateness under a more general biological account begins in Section 4.4 below; first, however, I set up that critique with a some general discussion of issues that arise when we begin to consider the question of whether our explanatory toolkit contains (or ought to contain) one analysis of innateness or many.

4.2 The scope of innateness

A question that arises immediately for any account of innateness concerns its scope of application. For any proposed definition of *innate*, we must ask whether the definition is meant to—or indeed, whether it *can*—apply to *all* scientific uses of the term, or merely

some of them. We must ask, that is, how *unified* are the phenomena that surround talk of innateness? This question is of particular importance given that innateness claims occur in what appears to be a broad range of contexts and in a variety of scientific disciplines and sub-disciplines. I have defended a notion of innateness in cognitive science with a limited scope of application, but as the issues here are rather complex and potentially confusing, it's worth taking the time to spell out the variety of options for views of varying breadth, as well as some of the reasons why one might favor one view over another.

The question of unity for analyses of innateness has the potential to cause considerable confusion, since there are multiple ways in which the role of innateness in different scientific contexts and disciplines might diverge. Specifically, there are two questions we must ask. One question is whether there is a single, universal *notion* of innateness—a single property or cluster of properties that everyone *intends* to pick out each time they use the term "innate," and that they go looking for when they look for innate things.[1] A separate question concerns whether there is a single *property* (or cluster of properties) that scientists investigate in the various contexts in which they study innateness. We can express this latter issue as a question about whether innateness is a *natural kind* in the sense that there is some core property (or property cluster[2]) the presence of which

[1] This question might be put in terms of whether the word "innate" picks out multiple *concepts*, but articulating this claim in a precise manner would require discussion of what exactly concepts are. Since I won't be engaging in that discussion here, I rely on the admittedly vaguer but more forgiving term "notion."

[2] Boyd (1991) offers a view of natural kinds that does not require natural kinds to be defined by the possession of a *single* property, but rather may be thought of as what he calls *homeostatic property clusters*. I do not take any substantive position on the nature of natural kinds here, but I leave Boyd's proposal open as a possibility.

causally explains the reliable co-occurrence of various further properties. (See Mameli, 2008; and Samuels, 2007 for related discussion of these issues.)

There are thus four possibilities when it comes to the unity of innateness, depending on whether we answer "yes" or "no" to each of the two questions posed above: (1) *One notion, one property;*[3] (2) *One notion, multiple properties*; (3) *Multiple notions, one property;* and (4) *Multiple notions, multiple properties*. We'll look at each possibility in turn.

The first possibility (*one notion, one property*) is simple enough: innateness is rather like femaleness. Everyone, presumably, has pretty much the same thing in mind when they talk about females, and there is a single naturally-occurring phenomenon that is investigated under the label of "femaleness" and that explains a wide range of other properties.[4] This possibility looks dubious from the start given the long history of still-unresolved debates about innateness that seem to be replete with cross-talk and confusion (to say nothing of the wide range of properties that have been suggested to be constitutive of innateness—see Mameli & Bateson, 2006, and below for an overview—and the recent charges that innateness talk is "fundamentally confused," Griffiths, 2002). Moreover, there is empirical evidence (Griffiths, Machery, & Linquist, 2009) that intuitions about innateness (among "the folk," at least) are hazy at best and likely inconsistent, and it is not implausible that such confusion enters into scientific practice at least some of the time. But

[3] For simplicity's sake, hereinafter I will drop the "or property cluster" qualification, but again, I do not mean to close off the possibility that innateness phenomena might be unified under a single property cluster.

[4] That there is a single property of femaleness is, I take it, consistent with the possibility that sex categories in many species—including perhaps our own—are not entirely binary.

nevertheless, "one notion, one property" is a live possibility that appears to have its supporters (e.g., Fodor, 2001).

The second possibility (*one notion, multiple properties*) is that there is a single thing we mean by "innateness" (though we may occasionally have inconsistent or confused beliefs about it), but there is no one property in the world that satisfies all and only the criteria included in our idea of innateness. We might all believe, for example, that innate traits are specified in the genome and are invariant in a population. But genetic specification and invariant emergence might end up picking out rather different classes of traits. The worry then is that we would be inclined to think the presence of one entails the presence of the other, even though these are distinct natural phenomena. In this vein Mameli (2008) suggests that "innateness" could turn out like Newton's "mass," which in modern thinking conflates rest mass and relativistic mass.

The third possibility (*many notions, one property*) is that our present situation with regard to innateness is not unlike that of scientists studying electromagnetism in the mid-18th century. For these scientists *thought* they were studying two different phenomena (electricity and magnetism), whereas the various phenomena they observed actually stemmed from the same underlying property.[5] If this is the case, then we can hope that as we continue to investigate the various phenomena associated with innateness, eventually we will come to see that phenomena we have treated as distinct (perhaps genetic encoding

[5] The analogy is not perfect, since there was, for a long time at least, little or no debate among scientists as to whether two distinct phenomena were being studied (they were mistaken in thinking there were two properties, but at least they all agreed), whereas in the present case there is disagreement and uncertainty about how many properties (if any) we are studying under the banner of "innateness."

and developmental invariance, for example) arise from the same underlying property (perhaps identical to one of the original options).

Finally, our fourth possibility is that there is little unity about innateness at all: perhaps it is the case that, not only are there multiple different notions of innateness at work in the sciences, but scientists who all say they are studying innate phenomena are actually studying different kinds of things entirely. This has the potential to make matters very confusing indeed, since it opens the possibility that scientists (and philosophers) are both constantly equivocating on their use of the "innate" (switching from one notion to another without realizing), and lumping together distinct phenomena that will turn out to have no unifying core of common causal properties. As we will see, my own estimation is that this fourth possibility describes the actual state of affairs regarding the study of innateness, though I will argue that there is (or at least can be) some order to the madness.

The array of possibilities described above raises the question of how we ought to proceed in studying innateness. That is, given the fact that there is widespread disagreement about what we should be looking for when we look for innate things, and about how many types of things we are likely to find, what should be our general attitude and orientation towards whether and how the notion of innateness should play a role in our scientific theorizing? Three such attitudes seem plausible and can be discerned in the literature:

The assumption of unity: One might argue that we should assume, until it is shown conclusively otherwise, that there is but one notion of innateness and but one natural-kind property that accounts for the various phenomena associated with that notion. On

this assumption, if we think we have found a good candidate for the one innateness property, we can chalk up the various instances of crosstalk and bickering to mere confusion (as does Fodor, 2001). Proponents of this view typically identify innateness with some biological property, and then argue (or assume) that psychologically innate traits are also innate in virtue of their possession of this biological property (see discussion of Fodor's view below).

Pessimistic disunity: On the opposite end of the spectrum, we might respond to the confusion regarding innateness notions and innateness properties by abandoning the concept altogether. Several philosophers of biology (Griffiths & Machery, 2008; Griffiths, 2002; Mameli & Bateson, 2006) have recently defended just this suggestion, on the grounds that the concept of innateness is "hopelessly confused." Proponents of this view tend to hold that there are both multiple notions of innateness and multiple, independent properties associated with it,[6] a state of affairs that yields only confusion and unwitting equivocation. According to the pessimistic disunity thesis, then, there is little or no hope of salvaging anything scientifically useful from the wreckage that comprises the notion(s) of innateness in its current state—for even if some of these notions actually do pick out natural biological or psychological kinds, the continued use of the term is simply an invitation to unwittingly equivocate among conceptually distinct notions.

[6] Samuels (2007) points out that the mere fact that multiple properties are "associated with" innateness need not be problematic, since "associated with" is a rather weak relation. The claim must therefore be a stronger one, namely that there are multiple independent, non-co-occurring properties that are taken to be *constitutive* of innateness.

Optimistic disunity: Finally, we can respond to the present quandary about the status of innateness by supposing that there may indeed be multiple distinct notions of innateness that should not be confused with one another, and multiple distinct properties that should be studied independently; however, at least in some instances, applying one of these notions within a restricted domain yields a scientifically legitimate concept, one that picks out an interesting natural property—we only get into trouble when we give in to the Procrustean urge to apply these restricted notions of innateness outside their proper scope of applicability. Samuels (2002, 2007) and Godfrey-Smith (2007) both take an optimistic disunity view, though they do not necessarily frame their own views as such. The position I defend in this essay is likewise an Optimistic Disunity view, since I argue—much as Samuels (2002) does—that there is a notion of innateness that serves the explanatory purposes of mainstream cognitive science but is inapplicable outside this domain.

It's worth noting that the optimistic and pessimistic versions of the disunity thesis are much closer to one another than either is to the unity thesis, since the two disunity positions are generally in agreement about what the facts of the world are: there are multiple distinct properties that get referred to by "innateness." The difference here is simply in the attitude one takes towards this fact. To endorse an optimistic disunity position is to consider it a mistake to respond to the disunity of innateness phenomena by abandoning the idea altogether.

It should be obvious why a Primitivist analysis of innateness goes hand-in-hand with an Optimistic Disunity view about the scope of this analysis: if Primitivism constitutes the best analysis of innateness in cognitive science, then there is little hope for unifying this notion with a biological concept of innateness, since the idea of psychological primitiveness has no obvious analogue in biology.[7]

My endorsement of Pessimistic Disunity means that I will be arguing *against* the idea that the notion of innateness in cognitive science defended in Chapter 3 can be subsumed under a more general biological notion of innateness. The rest of this chapter considers proposals for unitary notions of innateness—proposals that seek to identify a single notion of innateness that is useful in both the biological and psychological sciences. I demonstrate that each of these proposals fails to capture what is most central to psychological uses of the concept of innateness; hence I conclude that psychological innateness is distinct from biological innateness.

4.3 Theories of innateness: an overview

There are a variety of proposals in the philosophical and scientific literature for how innateness should be understood. Here I provide brief descriptions of some of the options, which will be followed by more detailed analyses of each, and assessments of the prospects for subsuming the primitivist account under them. I will argue that even if these approaches

[7] Gross and Rey (2012) do suggest that Jerne's (1985) work in immunology may offer a way to construct a biological notion of innateness based on the idea of *primitiveness*. Jerne discovered that all the antibody types one will ever have are present at birth; these antibody types thus comprise a set of biological primitives that are the building-blocks of the body's defenses against antigens. While the suggestion is intriguing, it does not seem particularly promising as a basis for a *general* account of innateness in biology.

constitute useful notions of innateness in biology, they cannot be expanded so as to do the explanatory work required of a notion of innateness in cognitive science. Thus I will not be seeking to show that these are bad notions of innateness *in general*, only that their scope must be limited (as is the case with my preferred notion of Primitivism).

Invariance/canalization views: According to this family of views, innateness can be characterized in terms of a trait's propensity to manifest in a wide range of environments. Traits whose emergence in development is insensitive to certain kinds of environmental variation are innate according to this view. Sober (1998) and Ariew (1999, 2006) have both offered accounts along these lines.

Genetic views: Genetic views tap into the widely-held—if fuzzy—assumption that innate traits are somehow genetic in nature. Genetic views come in two varieties. First, *genetic information* theories hold that genes in some way "code for" or "represent" certain phenotypic traits, and these traits are the innate ones; *genetic causation/determination* theories posit that some traits—i.e., the innate ones—are *caused by* (perhaps *determined by*) genes in a way that others are not.

4.4 Invariance, canalization, and innateness

Several philosophers of biology have proposed that innateness can be characterized in terms of a trait's propensity to manifest in a wide range of environments (Ariew, 1999, 2006; Sober, 1998). Sober (1998) casts this propensity in terms of *invariant emergence*.

Thus, "*a phenotypic trait is innate for a given genotype if and only if that phenotype will emerge in a range of developmental environments*" (p. 795, original emphasis). In similar spirit, Ariew contends that innateness is *canalization*. The canalization account holds that innate traits are those that are *insensitive* to a variety of environmental perturbations. The term *canalization* comes from Waddington (1957; 1975), who likens development to a ball rolling through a contoured landscape of possible phenotypes with hills and valleys (or "canals"). Each point in the ball's trajectory therefore represents the phenotype of an organism at a point in its lifecycle. Innate traits, according to this account, are the ones that arise from developmental pathways represented by the deepest grooves in the landscape—the areas of the landscape where the ball is likely to end up even if it is buffeted about by environmental forces.

These accounts are not without intuitive appeal—it seems quite natural to suggest that those traits that are liable to emerge even in the face of substantial variation in the nature of the environment are innate. For example, the fact that humans develop ten fingers and ten toes in all but the most inhospitable of developmental (e.g., pre-natal) environments is a plausible reason for thinking that number of fingers and toes is an innate trait in humans. We might further suppose that a similar story can be told for psychological traits: when a psychological trait seems to emerge invariantly despite environmental differences, then it counts as innate. Indeed, a proponent of an *invariance* view of biological innateness might extend this view to psychological contexts by arguing that Lidz et al. are right to classify syntactic competence as innate, but that what *makes* this competence innate is the fact that children exhibit this competence despite substantial variation in a range of

linguistic and non-linguistic environmental factors (e.g., diet, climate, parental nurturance, particular grammatical constructions prevalent in the linguistic input, and so forth).

To see why this approach fails, however, let us consider an objection to Lidz et al.'s conclusion that competence with NP is innate, one that comes from an *invariance* perspective. The objection comes from Griffiths and Machery, who write:

> [Lidz et al.'s data] is perfectly consistent with the development of children's knowledge of the constituent structure of noun phrases being contingent upon the presence of very specific stimuli in the environment of children. . . Furthermore, the stimuli required. . .might well be non-obvious. (Griffiths & Machery, 2008, p. 407)

Griffiths and Machery's contention is that Lidz et al. have not demonstrated the innateness of this piece of grammatical competence, because showing that knowledge of NP is innate would require demonstrating that the emergence of the knowledge does not depend upon (i.e., that it is invariant with respect to) some hitherto unexplored and potentially non-obvious environmental stimuli.

However, *pace* Griffiths and Machery, even if the emergence of syntactic competence *were* sensitive to some obscure environmental variable, this would constitute no objection to Lidz et al.'s argument. Suppose it did turn out that, by dint of some quirky neurodevelopmental mechanism, knowledge of NP structure fails to emerge in children whose neonatal diet is deficient in some obscure trace nutrient (cesium, let's say). There is no reason to suppose that the discovery of such a fact would shake Lidz et al.'s confidence in the innateness of this psychological trait. To all appearances, Lidz et al. take the innateness of knowledge of NP to be established by the fact that the kids didn't *learn it*. Even if the neurodevelopmental facts are such that knowledge of NP structure covaries in

some surprising way with an environmental variable, the core of Lidz et al.'s argument is still intact. Insofar as the potential existence of the kinds of environmental dependencies Griffiths and Machery worry about appear to be *irrelevant* to Lidz et al.'s argument and to the explanatory project in which it is situated, we have every reason to think that the notion of psychological innateness assumed by Lidz et al. is independent of the biological properties of *invariance/canalization*. Hence, we have no reason at this point to revise the view that the psychological notion of innateness is independent of any biological notion.

4.5 Genes and innateness

The idea that innateness—whether biological, psychological, or both—is somehow related to genetic properties is both ubiquitous and intuitively appealing. It is therefore a natural place for proponents of the Unity thesis concerning innateness to set up camp, as both Fodor (2001) and Chomsky (1988, 1990) do. Unfortunately, however, advocates of this approach are not always clear about how exactly the account is supposed to work. The idea is supposed to be that certain phenotypic traits are innate in virtue of some crucial relationship they bear to particular genotypic properties, but often we are given little insight into which genotypic properties are relevant or what the relationship between these genotypic properties and phenotypic traits is. Instead we are given rather vague gestures, such as the idea that innate traits are those that are "attributable to" the genome (Fodor, 2001, p. 102) or that phenotypic traits are innate in virtue of the fact that "the genetic endowment…specifies" them (Chomsky, 1990, 633).

In spite of this dearth of specificity, we can identify two strategies for making good on the idea that innate traits are genetic traits, each comprising a cluster of similar proposals. The first we can label the *informational approach*, according to which genes "code for" (or "represent" or "carry a program for constructing") certain phenotypic traits, and all and only those traits that are genetically encoded/represented/programmed are innate. The second we can call the *causal approach*, which seeks to identify some special causal relation (usually that of "determining") that holds between genes and certain traits. All and only the traits that are "genetically determined" in this way are innate, according to this view. I will argue that both of these accounts fail to deliver what the Unity theorist needs, however.

4.5.1 Innateness and genetic information

Let us begin with the informational approach. The idea that genes carry information is perhaps not entirely unproblematic (see Oyama, 2000), but let us assume for the moment that something like the following is right: genes utilize a code to store and convey information about biological structures (such as proteins), such that they can be said to *represent* those structures. If this is right, then we have the potential for precisely the sort of distinctive relationship between genes and traits needed to ground a genetic account of innateness: if genes hold coded representations of phenotypic traits like proteins, fingernails, blue eyes, and language acquisition devices, then there is a biological property that can be used to pick out innate traits, from the sub-cellular to the psychological.

However, as several philosophers have pointed out (Godfrey-Smith, 2007; Mameli & Bateson, 2006), the problem is that there does not appear to be any reason to suppose that genes code for or represent *any* of these things, except perhaps for proteins. The reason is that while it makes sense to speak of a genetic code that maps nucleotide sequences onto amino acid sequences in the process of protein synthesis, it is not at all obvious how one could generalize this story into an account of how genes code for or represent anything developmentally downstream of protein synthesis (including fingernails, eye color, or a recursive computational system). The reason we can speak of genes "coding for" proteins is that we can specify a sort of semantic mapping from one to the other, much as we might specify a semantic mapping from the dots and dashes of a string of Morse Code to sentences of English. But there is simply no way to specify such a semantic mapping from genes to fingernails, for example, much less from genes to a hypothesis space of grammatical principles. Hence, genes do not "code for" these higher-level traits.

The implications for the proposal that genetic encoding can provide a unitary concept of innateness should be clear: this proposal cannot countenance any innate psychological traits, so if there *are* innate psychological traits (and of course there are), a notion of innateness that appeals to genetic encoding is one with limited, rather than universal, scope.

Now, one might object that my dismissal of the prospects for genetic encoding to undergird a unitary notion of innateness has been too hasty. For one could point out that although genes do not, strictly speaking, "code for" anything beyond proteins, these very

proteins that the genes code for will go on to play a crucial causal role in the development of everything from fingernails to the assemblages of brain cells that implement the psychological structures that are of interest to cognitive scientists. So we can, in principle, trace a continuous causal path from genes to the very psychological traits that nativists in cognitive science count as innate.

However, it should be clear that the existence of such causal pathways does not vindicate the proposal that genetic *encoding* is the biological property in virtue of which traits of all kinds are innate. This is because the existence of such causal pathways does not change the fact that organism-level psychological traits are not genetically encoded. Simply put, encoding is one thing, and causation is another. Godfrey-Smith sums this up nicely: although we can perhaps unproblematically count *some* of an organism's traits (e.g., its proteins) as being genetically encoded,

> [o]nce we are asking questions about levels of aggression, about Universal Grammar, about sexual orientation or basic numerical cognition, we have far outrun the kinds of entities that can be said, on the basis of biological theory, to be coded for. When understanding those sorts of traits, we are back in the land of ordinary causal explanation." (Godfrey-Smith, 2007, p. 61)

Since following the trail of genetic encoding has led us to the land of ordinary causal explanation, let us consider the proposal that there is a unique kind of *causal* relationship between genes and certain phenotypic traits (rather than a semantic/informational relationship), and it is in virtue of this relationship that traits—physiological or psychological—are innate.

4.5.2 Innateness and genetic causation

The idea that underlying the notion of innateness is the fact that some phenotypic traits—cognitive and otherwise—are "attributable to" genetic factors in ways that other traits are not is perhaps the most widely-held position on innateness. Unfortunately, it is also in many ways the least well-defined of the available theoretical treatments of innateness, so part of our task will be to spell out what exactly the proposal is. This will be the last challenger to the view I have espoused according to which innateness in cognitive science cannot be identified with any biological property. If this challenge fails, we will be left with Primitivism as our only viable candidate for a notion of innateness in cognitive science.

Let us begin by attempting to specify in a bit more detail just what the genetic causation view of innateness amounts to. It is widely acknowledged that the view cannot be that innate traits are those that have *only* genetic causes, for there are no such traits—all traits will have both genetic and environmental causes (Kitcher, 2001, calls this the "interactionist consensus"). It is similarly widely acknowledged that it will not help to suggest that what makes a trait innate is that it owes *more* to the influence of genes than to the influence of the environment. Genetic and environmental factors *interact* to give rise to any particular trait, so their contributions cannot be separated out from one another and the quantities compared (Sober, 1988; Keller, 2010).

So we cannot make sense of innateness-as-genetic-causation by appealing to some quantitative measure of the causal contribution to the trait made by genes (e.g., *all* or *most*). Rather, it appears we must find some way of affording genes a special *kind* of role in the

emergence of certain traits, a role not also played by environmental causes. But then we are left with the question of what this special causal role could be. One might be tempted at this point to reach for the notion of genetic encoding, since coding for biological structures is a special kind of causal influence that genes can exert but that environmental factors, plausibly, cannot. But clearly this will just take us in a circle: we abandoned the genetic encoding account precisely because it only forestalls momentarily the need to sort through the many causal factors contributing to the emergence of any particular trait.

There is one other move open to the proponent of the genetic causation account. I said a moment ago that there is no way to "single out" one cause among many (e.g., a particular genetic factor) of a phenotypic trait and assign it a (quantitatively or qualitatively) special causal role. This will remain true so long as we are considering the causal etiology of a particular trait in a particular individual: questions about the relative causal contributions of genetic and non-genetic factors to that trait will be senseless, and all we will be able to say about the trait's causal etiology will be to describe the particular causal interactions that took place in the trait's development. However, if we shift our focus beyond the individual to the population the individual inhabits, we can potentially begin to ask different sorts of questions, some of which seem like they might actually help us identify a sense in which a trait could be "more genetic" than it is environmental. So a defender of the genetic causation account might advocate calling upon population-level considerations in an attempt to salvage the idea that some traits are more attributable to genes than to the environment. If all this can be accomplished, then we should, in principle, be able to identify an organism's innate traits with these "more genetic" ones, and voilà: a

vindication of the idea that innate traits are those that are significantly genetically determined.

Needless to say, I don't think this will work, but let's examine how the story would go.[8] Again, questions about the relative causal contributions of genes and environment to a particular trait are meaningless, but we can begin with such a meaningless question—"How much of Jane's hair color is due to her genes and how much to her environment?"—and, taking into consideration facts about Jane's population, pose instead a similar (but not identical) question that can be coherently asked and coherently answered. If we have information about the hair color of individuals in Jane's population and about degrees of genetic relatedness among individuals in the population, an analysis of variance (ANOVA) can tell us the relative contributions of genetic differences and environmental differences to differences in hair color across the population.

This procedure is common in the field of population genetics (and the subfield of *behavioral* genetics in particular), where the resulting value for the amount of phenotypic variation in a trait that is "attributable to" genetic variation[9] is called the *heritability* of the trait. Note that in the domain of behavioral genetics, "heritability" is a technical term that differs in meaning from the term's vernacular use. Heritability for a trait is here defined as the proportion of total variation in that trait in a population that can be attributed to genetic variation. Heritability is therefore represented as a number between zero and one, with values near zero counting as "low heritability" and values near 1 counting as "high

[8] My discussion here draws heavily on Sober's (1988)—see also Lewontin (1974) for a similar take.

[9] As we will see, there is considerable disagreement about how this notion of "attributing" phenotypic variance to different sources (also expressed as genetic or environmental variance *explaining* phenotypic variance) should be interpreted—and in particular whether it should be given a *causal* interpretation.

heritability." Debates about the notion of heritability will figure prominently in the discussion in Part II of this dissertation, and so it is worth taking some time to explain familiarize ourselves with the basic methodological and theoretical issues relating to heritability.

A heritability statistic is generated by collecting data about the variation in a trait within a population, and about the degree of genetic similarity among different pairings of individuals in the sample. The analysis of variance (ANOVA) is then performed to determine how much of this variation can be "explained" by variation in genotypes between individuals. In practical terms, there are two primary methodologies for determining the heritability of a trait in a population. One is to measure trait differences among genetically related individuals who have been separated (i.e., reared in different environments). Typically this is done by comparing the similarities of monozygotic (MZ, or "identical") twins reared apart with those of dizygotic (DZ, or "fraternal") twins reared apart. Since MZ twins share 100% of their genes and DZ twins (on average) share 50%, we can infer how much genotypic differences contribute to phenotypic differences. The second, more indirect methodology is to compare trait correlations among people with varying levels of relatedness in the population (i.e., asking the question: how much more similar to one another are siblings than are first cousins? Second cousins?).

The basic intuition behind the heritability statistic is this: when individuals who are more genetically similar to one another tend to be similar to one another in their manifestations of some trait, and when genetically dissimilar individuals are more different

in their phenotypes, then it would seem that the similarities and differences in traits are best "explained" by similarities and differences in genes.

Now, returning to our question about Jane's hair, suppose we conduct a heritability analysis and we discover that genetic variation makes a much greater contribution to variation in hair color than does environmental variation. Can we now conclude that Jane's genes make a bigger contribution to her hair color than her environment (and hence, that her hair color is innate)?

No. The question of whether Jane's genes or her environment made a bigger contribution to her hair color is just as meaningless as it was before. However, with the information provided by the ANOVA, we *can* ask what Jane's hair color would have been like if she'd had the same genes but inhabited a different environment, or if she lived in the same environment but had different genes. And we can further ask: in which of these counterfactual scenarios would her hair color have been *more* different from its actual color? By drawing on population-level facts we can eventually generate an answer to the question of whether Jane's genes or her environment made a bigger *difference* to her hair color (importantly, this is *not* the same as the question of which made the bigger *contribution*). It might appear that at long last we have found our biological notion of innateness: if Jane's genes made a bigger difference to her hair color, then it is innate. If her environment did, it's not. Moreover, there is no reason, in principle, why we couldn't apply this same procedure to psychological traits to determine whether they meet this criterion for innateness, which would yield a single, unified notion of innateness.

There are two reasons why this approach fails, however. First, note that this proposal's rather uncritical use of the question, "would Jean's hair color have been more different if she'd had different genes or a different environment?" is deeply problematic. For the answer to this question will be different depending on exactly what properties we assign to her genotype and to her environment when we consider the counterfactual scenarios in which these factors would have been different (i.e., if we are asking what her hair color would have been if she'd had the same genes but lived in a different environment, how do we know which of the infinitely many possible alternative environments we should consider?). And as Sober (1988) points out, there do not appear to be any principled theoretical guidelines for making these decisions.

Second, and more importantly, this approach seems to characterize innateness in terms of considerations that are simply *irrelevant* to what I have argued (in Chapter 3) that innateness is all about. I have argued, again, that whether a particular trait is innate for an individual should be a matter of the facts of *that individual's development*. But under this latest construal of the genetic causation account, we can only determine whether some trait of Jane's is innate by considering facts about the distribution of various genetic and environmental variables in Jane's population. It seems odd, to say the least, that whether Jane's hair color (or her fingernails, or her hypothesis space of possible grammars) is innate should depend on such a diffuse set of facts that are completely extrinsic to the developmental processes through which Jane's traits emerge. Indeed, if we take this take this population-based approach to defining innateness, then an organism in a population of

one could have no innate traits, no matter what sorts of unlearned cognitive capacities (for example) it displayed.

Taking these matters into consideration also helps us to see why a recent view from Waters (2007), which seeks to afford genes a unique causal role as causal *difference makers*, is of no help to the Fodorian/Chomskyan view of innateness as genetic causation.[10] For although Waters aims to establish that there is an "ontological feature that distinguishes DNA as the actual difference maker" (Waters, 2007, p. 3), his approach features the same drawback as do heritability analyses. That is, Waters' notion of *causal difference-making* is designed to account for the causal role of *genetic differences* in generating *trait differences*, rather than in the causal role of *genes* in generating *traits*.

Waters provides a formal description of his notion of *difference maker*, but I think the idea is best seen in his example of an experiment by T.H. Morgan, who cross-bred two strains of *Drosophila*—females that were homozygous dominant for a gene controlling eye color ("the purple gene"), and males that were homozygous recessive for the same gene. Thus, there were three populations (mothers, fathers, and offspring) exhibiting two different traits. What is the cause of the difference in eye color? As a matter of fact, a Woodward-style intervention on several other genes would *also* have yielded differences in eye color, and so, Waters concedes, these genes all count as "causes" of the flies' eye color in some sense. But, Waters argues, it was at the site of the purple gene that the populations *actually* differed, and so it was the purple gene that "made the difference."

[10] Note that Waters does not defend his view as a genetic account of *innateness*, though he does seem to think that his account constitutes a response to "causal parity" theorists (e.g., Oyama, 2000), whose work generally centers among the factors contributing to *the development of an individual*

Waters may be correct that the notion of "making a difference" grounds an ontological distinction among causes when those causes are causes of *differences within a population*. But this again, this won't help to formulate a genetic notion of innateness. For if we were to attempt to apply Waters' analysis to a case of individual development, we would be forced to appeal not to an existent difference between two individuals but between an individual and a *counterfactual* version of that individual. And a Woodward-style counterfactual account of causation, when applied to such circumstances, yields causal parity (i.e., does not ontologically privilege one set of counterfactual circumstances over another). So if our aim is to develop an account of the role of genes in development, then the sort of question to which Morgan's experiment was directed cannot help us.

Putting these critiques together, it appears the proponent of a genetic causation approach to innateness faces a choice between two unappealing options: either try to develop a notion of genetic causation/determination that operates at the individual level, and face the difficulties of identifying a unique causal role for genetic factors, or else draw on population-level facts that can potentially identify statistical asymmetries between genetic and environmental factors, and risk ending up with an account that appeals to intuitively irrelevant considerations in determining whether or not a given trait is innate.

I conclude, therefore, that the prospects for the genetic causation (or genetic determination) account of innateness are looking dim. All told, our consideration of three different proposals for unitary notions of innateness (*invariance/canalization*, *genetic information*, and *genetic causation*) has given us no reason to reconsider the judgment

reached earlier that psychological innateness is best construed along the lines of the Primitivist account.

4.6 From populations to individuals (again)? And a preview of things to come

This concludes my consideration of the meaning of *innateness* in contemporary cognitive science and my defense of a primitivist account, which has been the focus of Part I of this dissertation. But before closing this chapter, and setting aside (for a while at least) the questions about heritability, populations, and individuals I considered in the previous section, I would like to consider one further argument for the view that measures of genetic contributions to population-level variability (e.g., heritability analyses) can justify privileging genetic over environmental factors in individual development.[11] This argument comes from philosopher Neven Sesardic, who, as we will see, is a staunch defender of the hereditarian tradition in the "nature-nurture" debates (and race-IQ debates in particular) that I'll be examining in Part II. Analyzing Sesardic's arguments in detail will serve as a useful transition to the second part of this dissertation for several reasons. First, Sesardic is one of the few authors to attempt a full-throated defense of the value of heritability statistics in drawing inferences about *individual-level* phenomena. If I am correct in defending the received view, then we should be able to identify where Sesardic goes wrong in rejecting it. Second, examining Sesardic's mistakes will help us to see why confusion and cross-talk are so common and so deeply ingrained in nature-nurture debates.

[11] While I hope that the arguments in this section will prove illuminating, they are ultimately outside the main thread of the argument for either Part I or Part II; the reader may proceed to Chapter 5 without missing any essential steps.

As we saw in the previous section, the received view about heritability is that one cannot infer anything about individual-level developmental processes from population-level statistics like heritability (Northcott, 2006; Sober, 1988). Sesardic (2005), however, takes exception to the received view. He discusses a series of examples and statements offered by both critics of heritability and by behavioral geneticists themselves expressing the received view, and argues that they are wrong.

Sesardic criticizes several expressions of the received view. One is from a 1998 article by behavioral geneticist David Lykken: "It is meaningless to ask whether Isaac Newton's genius was due more to his genes or his environment, as meaningless as asking whether the area of a rectangle is due more to its length or its width" (1998, p. 24). A second comes from Cosmides and Tooby (1997), and the passage there that Sesardic attacks is worth reproducing in full:

> A heritability coefficient measures sources of *variance* in a *population* (for example, in a forest of oaks, to what extent are differences in height correlated with differences in sunlight, all else equal?). It tells you nothing about what caused the development of an *individual*. Let's say that for height, 80% of the variance in a forest of oaks is caused by variation in their genes. This does not mean that the height of the oak tree in your yard is "80% genetic." (What could this possibly mean? Did genes contribute more to your oak's height than sunlight? What percent of its height was caused by nitrogen in the soil? By rainfall? By the partial pressure of CO2?) When applied to an individual, such percents are meaningless, because all of these factors are necessary for a tree to grow. Remove any one, and the height will be zero. (qtd. in Sesardic, 2005, pp. 55–56)

Against these claims, Sesardic argues that there are "perfectly meaningful answer[s]" to questions about the quantitative causal contributions made by various factors

that contribute to a phenotype, and that "population-to-individual [inferences are] not 'meaningless' at all" (2005, p. 56).

However, it is not entirely clear what proposition Sesardic means to establish. He says that Cosmides and Tooby "are wrong that heritability tells us *nothing* about what caused the development of an individual," (2005, p. 56, emphasis preserved) so clearly we are to understand that Sesardic thinks heritability tells us *something* about what caused the development of an individual. But what is this something? In the next sentence, Sesardic notes that if a trait is highly heritable, then we can infer that an individual's deviation from the mean for that trait is "in all likelihood caused more by its deviation from the genetic mean than by its deviation from the environmental mean" (2005, p. 56). This is plausible enough, but Sesardic doesn't say whether this is the "something" that heritability ostensibly tells us about causal etiologies in individual development, and if so, what exactly we're being told.

Sesardic's next move is to observe that "someone might object" (again, since it's not clear what claim Sesardic is advancing, neither is it clear what proposition this is supposed to be an objection *to*) that deviations from the group mean are *relational* facts and therefore not genuine facts about individuals. In response, Sesardic points out that "many population-relative facts are usually regarded as giving *important information about individuals*, like being an Olympic champion, having an IQ above 140, being a best-selling author, being a surgeon with an unusually high patient mortality, etc." (2005, p. 56, emphasis added). So perhaps the claim Sesardic wishes to establish here is simply that we can acquire "important information about individuals" from heritability analyses. The

notion of "important information" is perhaps on the vague side. Yet even if we give Sesardic wide berth on this, it's unclear why he would take this quite general claim to be important in the present context: what bearing does the rather generic notion of "important information about individuals" have on questions about "what caused the development of an individual"? Again, we are left unsure of the thesis being advanced and why it matters.

Here is what I take to be the most plausible construal of Sesardic's argument. Broadly speaking, we can distinguish questions about individual-level phenomena (such as heights of individuals) and questions about population-level phenomena (such as the mean height of a *group* of individuals). It is a well-worn point in debates about heritability and development that different "nature-nurture questions" arise at these different levels; the explananda at one level are different from the explananda at the other, and different investigative and explanatory tools must be applied with respect to each sort of phenomenon (i.e., population-level questions are pursued using heritability/ANOVA and individual-level questions are pursued with analyses of developmental mechanisms). This might be summed up in rough-and-ready fashion with the following dictum*: One can't infer anything about individual traits from population-level tools like heritability*. This dictum might be seen as positing an unbridgeable chasm between the individual and population levels. It is such a dictum, with such a reading, that I think Sesardic has in mind as his target. His aim, then, is to show that the divide between the population and individual levels can be bridged: it can be bridged by using population-level analyses like heritability to make inferences regarding *relational facts* about individuals. If the argument goes through, Sesardic will have succeeded in rebutting one of the most prominent criticisms

advanced by developmentalists against heritability, by showing that we can, despite developmentalists' insistence to the contrary, unproblematically move from the population to the individual level.

But Sesardic's conception of the relation between the individual and population levels is too course-grained, and in consequence his rebuttal is aimed at the wrong target. Sesardic is right that when talking about biological traits we can oftentimes move from the population to the individual level by considering relational facts (i.e., we can use heritability to say what caused an individual's deviation from the mean). But the cautionary notes expressed by Lykken and Cosmides and Tooby were never intended to deny that *any* information *whatsoever* about individuals can be adduced from heritability. The scope of these admonitions, rather, is limited to the practice of making inferences about *developmental processes* or the *causes of particular developmental outcomes*.

To see the mismatch between what Sesardic seeks to establish and what the challenge to heritability is supposed to be, notice the slippage in the language he uses when characterizing his opponents' views. In introducing the quotation from Cosmides and Tooby, for example, he says, "Cosmides and Tooby also argue that a heritability value tells us *nothing about an individual"* (2005, p. 55, emphasis added). Note, however, that what Cosmides & Tooby actually say is that a high heritability score for a trait in a population "tells you nothing about what *caused the development of* an individual". Cosmides and Tooby's claim is clearly about what heritability can tell us about a *particular aspect* of individuals (viz., the causal processes underlying their development), not about what heritability can tell us about individuals *tout court*. And, oddly enough, a few sentences on

(and as quoted above), Sesardic, having just characterized Cosmides and Tooby's claim as about the informativeness of heritability regarding *individuals* (*tout court*), lapses back to talking about the relevance of heritability to causal processes in individual *development*: Cosmides and Tooby, he says, "are wrong that heritability tells us nothing about *what caused the development of an individual"* (2005, p. 56, emphasis added).

The vacillation in Sesardic's characterization of his target notwithstanding, his core suggestion—that relational facts are all the bridge we need between the population level and the individual level—fails to address the key issue raised by the developmentalist's insistence that heritability is silent about the causes of individual development. For knowing some relational fact *about an individual oak tree* (say, that it is one standard deviation taller than average) is not, *pace* Sesardic, the same as having information about the causal mechanisms by which the tree came to be (say) eighty feet tall. Sesardic is quite correct that it is perfectly sensible to say that it is a fact (specifically, a relational fact) about the individual tree that it is taller than average. But the fact that an individual tree is taller than average is *not* a fact about the mechanisms at work in the development of the tree which are causally responsible for its being eighty feet tall.

It would seem rather obvious that relational facts about an individual's standing within a population and causal facts about its development are two quite different things. But should we need confirmation that these are distinct phenomena, we need only observe that the intrinsic causal character of the individual's developmental mechanisms would be the same no matter what the population mean; that is, identical epigenetic sequences could underlie the development of a tree that deviates in height from the population mean by -

2SD or +2SD—it simply depends what population the tree is part of. So while the heritability of height in the population may be informative about the causes of an individual's deviation from the mean (and one might be interested in this information for a variety of reasons), the causal-mechanistic processes underlying development are a distinct class of phenomena, and so must be explored and explained in different ways.

To sum up: the objection to heritability that Sesardic *should* be concerned with is the following: the causal-mechanistic etiology of a trait value is one thing, the deviation of that value from the population mean is another; just because heritability provides information about the latter does not mean it provides any substantive knowledge about the former. Sesardic's (again, perfectly reasonable) observation that *relational* facts about individuals are nonetheless still facts about individuals, and that heritability can provide information about relational facts, does nothing to disarm this objection.

The foregoing discussion undercuts Sesardic's attempt to collapse questions about the causal mechanisms at work in individual development into questions about individual and/or group differences. In spite of Sesardic's attempt to demonstrate that heritability analyses are informative about the causes of individual development, the received view (that heritability analyses are, for all practical purposes, silent about mechanisms of individual development) still stands. To put it bluntly: where developmentalists talk about analyzing the *causes of individual development* (and the irrelevance of heritability of such analyses), introducing the idea of individual deviations from the population mean, as Sesardic does, is simply changing the subject.

However, although Sesardic's attempt to close the chasm between the population level and the individual level is not successful, the above discussion—in particular the distinction between causes of individual development and population-relative facts about individuals—can help us make sense of cases like Isaac Newton's genius, as well as disagreements about them. To review the disagreement: Lykken takes the standard line on heritability and individuals: even if intelligence is highly heritable, it is "meaningless to ask whether Newton's genius was due more to his genes or his environment" (1998, p. 24); Sesardic disagrees, saying once again, "it makes perfect sense to inquire whether Newton's extraordinary contributions were due more to his above-average inherited intellectual ability or to his being exposed to an above-average stimulating intellectual environment" (2005, p. 55). Evelyn Fox Keller (2010, p. 42) also weighs in on this particular case, defending Lykken against Sesardic and citing the latter's treatment of the Newton case as a clear example of conflating questions about individual traits with questions about trait differences.[12]

As should be clear from the above discussion, my own position ultimately aligns with Lykken's and Keller's, but Sesardic's mistake is understandable. This is because a case like "Newton's genius" is particularly amendable to engendering confusion about the relevance of heritability to individuals, because the example exhibits an ambiguity between the causal and the relational. That is, either side (Sesardic or Lykken/Keller) can draw support from the example depending on whether "Newton's genius" is conceived as an

[12] The difficulty of keeping separate questions about individual traits and questions about trait differences—even for careful thinkers—is the main theme of Keller's (2010) book, and the analysis I provide here is in the same general spirit.

intrinsic or a *relational* fact about him. On the one hand, Newton's particular assortment of cognitive capacities (for abstract reasoning and so forth) were *intrinsic* properties of him as an individual: somehow or other Newton came to be possessed of a mind/brain with certain properties, among these being the ability to generate an ingenious physical theory that unifies a wide array of phenomena. On the other hand, a very commonsensical way of defining a "genius" is simply as "someone who's a lot smarter than the rest of us."

The first of these construals (Newton's genius as a set of cognitive capacities) takes a causal process as the relevant explanandum: by what mechanisms did Newton come to have a brain that could do the things that Newton's brain did? This is the question Lykken and Keller have in mind, and they are correct that the causal contributions of Newton's genes and his environment to this outcome cannot be partitioned and quantified. The second construal (Newton's genius as relational intellectual superiority) takes the relevant explanandum to be a relational fact: why was Newton so much smarter than other people with whom he shared varying degrees of genetic and environmental similarity? This is how Sesardic interprets the question about Newton's genius, and he is right that we can use information about the heritability of intelligence in Newton's population to make an informed judgment about whether atypical genes or an atypical environment made the greater contribution to Newton's large deviation from the population mean.

Again, Sesardic is not wrong that there are some individual-level facts (i.e., relational ones) about Newton to which heritability is relevant, and moreover, there might certainly be contexts in which these are the individual-level questions about Newton that most concern us. The point, then, is simply that facts about the intrinsic character of

Newton's cognitive capacities (and how they got to be that way), and relational facts about his capacities as compared with the rest of his population, are simply different phenomena, to be explained in different ways. In other words, Sesardic's mistake is not in taking heritability to provide information about Isaac Newton (for it does do that), but in apparently assuming that *relational* facts about Newton are the only ones we ought to care about.

The general lesson to be drawn from our discussion of the Isaac Newton case is that specificity in characterizing the phenomenon of interest is especially crucial in the context of nature-nurture debates. Relatedly, the foregoing discussion also helps to highlight one of the central themes to come in Chapter 7, namely the *divergence between the explanatory/pragmatic interests of hereditarians and their opponents in the nature-nurture debates* (one manifestation of which is the kinds of phenomena towards which the different camps direct their explanatory energies).

To see this divergence of explanatory interests, note that despite the confusion in his argument, Sesardic ultimately does seem to acknowledge the gap between individual-level and population-level questions; however, he argues that his focus on the latter is justified because considering individual developmental processes "amounts to abandoning a *population* perspective. With this move the connection with the nature-nurture controversy becomes tenuous because that debate makes most sense as looking for answers at the level of population differences" (2005, p. 54).

Now, Sesardic is correct that *if* we are merely concerned with *describing* population differences, or perhaps *predicting* them, then nature-nurture debates are indeed about

"answers at the level of population differences." But, as I will demonstrate later on, if we are instead concerned with exerting *control* over developmental processes in order to *change* the distribution of phenotypes in the population (and in various sub-populations), then it is entirely appropriate to focus on individual-level processes, and to point out that heritability analyses will not tell us what we want to know. In other words, Sesardic's stipulation that nature-nurture debates must be confined to population-level questions (i.e., heritability of the relevant traits) merely reflects the fact that his interests are limited to description rather than manipulation. Since others do not (and, I have argued, should not) share these same interests, there is no need for us to maintain focus exclusively on population-level questions, as Sesardic insists.

CHAPTER 5: INTRODUCTION TO PART II: SCIENCE, MORALITY, AND RACE

5.1 Introduction

Racial and ethnic groups continue to enjoy vastly different levels of political, social, economic, and cultural power, both in the United States and around the world. It is typically thought that such inequality is a morally unacceptable state of affairs, and that efforts should be made to remedy it. However, a number of thinkers in what is known as the *hereditarian* tradition in the behavioral sciences have argued that racial inequality is a natural and inevitable outcome of "innate" or "genetic" differences between racial groups (e.g., in intelligence and other socially-relevant psychological characteristics) and need not raise any particular moral concerns. In contrast to the prevailing post-Civil-Rights-Era view, these hereditarians argue that we should assume psychological differences among racial groups to be fixed and irremediable, that we should take such differences as empirical premises in our collective policy deliberations, and that we should accept permanent racial inequality as simply a fact of life.

This chapter and the three that follow are dedicated to arguing against these claims. My arguments draw on empirical, conceptual, and moral considerations. Many scientists and philosophers have also mounted critiques of hereditarianism's scientific and social agenda on these three (empirical, conceptual, moral) grounds, but my arguments combine these considerations in a novel way. I argue that our scientific, social, and moral interests

are best served by adopting a particular theoretical orientation towards the investigation and explanation of biological and psychological development. The approach I favor privileges *mechanistic* understandings of organismal development—of the causal interactions between genes and environments in developmental processes—over generalizations about the statistical relevance of genotypic variation to phenotypic variation. As will be explored in the next chapter, mechanistic understandings of development afford greater opportunities for intervention and control of developing biological systems than do behavior-genetic alternatives. Insofar as we wish to intervene to eliminate racial differences, then, mechanistic strategies are to be preferred.

The project pursued in this second part of the dissertation is largely distinct from that completed in Part I, but there are some important connections between the arguments of Part I and the goals of Part II as well. For example, I argued in previous chapters that the acceptability of different explanatory strategies in science depends upon our pragmatic interests (including moral interests)[1] and goals. Therefore, assessing the scientific worth of different strategies in these nature-nurture debates[2] will require that we attend to the broader interests and goals we bring to bear on the scientific investigations at hand. Whether terms like "innate" and "genetic" earn their explanatory keep in the domain of debates about racial differences will depend, in large part, on the broader explanatory context in which nature-nurture battles are engaged.

[1] Recall our discussion of *omissions* in Ch. 2: my failure to water your plants is relevant in explaining the plants' death if I have made a *promise* to water them

[2] Throughout Part II, I use the term "nature-nurture debates" interchangeably with "racial-difference debates." Both of these are, I believe, distinct from the *nativist* debates discussed in, e.g., Chapter 3.

Understanding the full social and moral context of these debates will require exploring the dynamics of nature-nurture debates—claims and counterclaims, polemical strategies, and so forth—in some detail. But my overarching aims in these chapters will be to show that we have substantive *moral* reasons to disfavor hereditarian explanatory strategies, and to show that mechanistic, interactionist strategies constitute a viable alternative that comports with our collective values.[3] A key step in establishing these claims will be my enumeration of the moral costs (i.e., the morally bad outcomes) of adopting hereditarian strategies (Section 6.2). The rest of the present chapter, however, is dedicated to setting the stage for these arguments. I begin in Section 5.2 with some general comments about how I will approach the rather difficult and sensitive topic of scientific claims about racial inferiority. Section 5.3 specifies the core claims of hereditarianism (i.e., the claims I will be arguing against), while Section 5.4 provides an overview of my overall strategy—pursued throughout Part II—for resisting these core claims. Finally, Sections 5.5 and 5.6 establish some fundamentals of the current empirical status of questions about racial differences, by discussing the biological nature of "race" and providing an overview of some of the main empirical matters under dispute in the race-IQ debate.

[3] In other words, I am offering a version of what has been called the "underdetermination argument": the choice between different research agendas and explanatory strategies is underdetermined by the totality of empirical evidence that could be gathered; thus some set of non-scientific criteria or other must therefore be used to make such choices, and moral and social values are the best criteria available (Anderson, 2004).

5.2 Some general comments about racial difference debates

Given the long and fraught history of scientific debates concerning potential racial differences, it is worth beginning with some general comments on the nature of the debate. These are highly sensitive issues, and as will emerge in Chapter 8, I believe there is a duty to talk about them in the right way. Therefore, my aim in this section is to offer some context to the discussion that follows, including situating the specific debates I'll be looking at within the broader context of historical and current conversations about the possibility of racial differences.

Claims have been put forth that races differ in their "innate" or "genetic"[4] propensities with regard a wide variety of socially-important psychological traits, and such claims are typically followed by heated dispute. The most famous of these debates concerns intelligence, but similar controversies surround traits such as personality (extrovertedness, openness to experience, conscientiousness, and so forth), aggressiveness, and (lack of) sexual restraint.[5] For reasons of space and of tractability, my discussion will, for the most part, be limited to what is by far the most well-known of these debates, that centering on differences in intelligence or IQ[6] between White and Black racial groups.[7]

---

[4] In Part II of the dissertation, I'll mostly be using the terms "innate" and "genetic" when referring to hereditarians' positions. Hence, the words will mostly appear in scare quotes to indicate that I am endorsing neither the hereditarian's definitions of these terms, nor their application to the traits in question.

[5] It is common in these debates for there to be disagreement not only about the extent to which differences are "innate" or "genetic," but also how differences should be defined and measured. For a helpful discussion of such issues in the case of aggression, see Longino (2001).

[6] Strictly speaking, there is a threefold distinction between *intelligence* (a mental ability), *IQ* (a score on a test) and *g* (symbolizing *general intelligence*; a latent factor computed from batteries of psychometric tests). Since the technical distinction between intelligence and IQ is immaterial for my purposes, I use these two terms interchangeably.

[7] "Blacks" here refers to both native Africans and members of the African Diaspora throughout the world, including African-Americans. I capitalize names of racial/ethnic groups for the same reasons given by Sally Haslanger: "Doing so is warranted, I believe, in order to be consistent between races that are referred

However, it is with some hesitation that I adopt this restricted scope. The reason for this hesitation is that I believe the full depth and breadth of the moral issues raised by claims about racial differences cannot be truly understood when these claims are approached as atomistic empirical statements—he full moral weight of these statements emerges when they are considered collectively. For example, examining a range of claims about racial differences allows us to see that it is not only IQ in which Blacks are allegedly deficient, but a host of other traits as well—they are consistently portrayed as all-around inferior human specimens (Levin, 1997b; Rushton, 1997). Thus, pretty much without exception, the hereditarian position happens to be that wherever racial/ethnic differences in socially-relevant traits are found (and "genetic" explanations offered), the prevalence of desirable traits will be lowest, and the prevalence of undesirable traits highest, among Blacks.[8]

Similarly, although hereditarians do typically attribute higher IQs to Northeast Asians as compared to Whites, very little is made of this difference as compared to the amount of attention hereditarians dedicate to the gap between Whites and Blacks. The gap between Whites and Northeast Asians is thought to be only a fraction of the size of the gap

---

to using color terms and those referred to using names of continents, to highlight the difference between ordinary color words and the homonymous use of such words as names for some race, and highlight the artificiality of race in contrast to the apparent naturalness of color. . .making this distinction between "color" and race explicit is, I believe, theoretically important" (2004, n. 1).

[8] The supposed superiority of Blacks in athletic endeavors (e.g., Entine, 2008) is frequently cited as a counterexample, but it's actually not clear how favorable the stereotype really is for Blacks: as a number of commenters have pointed out (Miller, 1998; Sheldon, Jayaratne, & Petty, 2007), a sort of implicit assumption prevails in our society that there are people who are smart and people who are good at sports—people with good brains and people with good bodies—and that these categories are to some degree exclusive and compensatory (i.e., an enhanced prowess in one entails a lack of prowess in the other). The idea that Blacks are superior athletes may just be the flip side of the belief that they aren't smart. The suggestion is speculative and fully assessing it would take some work, but the point, again, is that how stereotypes of athleticism actually play out for Blacks is not as clear-cut as it is sometimes made out to be.

between either of these groups and Blacks (i.e., the White-Asian IQ gap is standardly reported at about 3-5 points, whereas the gap between Whites and Blacks is claimed to be at least 15 points).[9] Thus, Blacks are portrayed as inferior to Whites and Northeast Asians by a wide margin and across *all* traits that are thought to contribute to a person's being a successful, responsible, and morally upright member of society. While Black-White IQ gaps and White-Asian IQ gaps are both part of hereditarian theory, only one of these is presented as a creating a genuinely qualitative difference among human beings; hence Rushton makes a point of emphasizing that "While Orientals developed complex societies in Asia, and Whites produced complex civilizations in Europe, Black Africans did not" (2000, p. 22).[10]

To the extent that many observers are suspicious of hereditarian claims about racial differences, much of this suspicion likely stems from the fact that it seems an odd coincidence that a single group within a species, a species that exhibits an extremely small amount of overall genetic variation (relative to other animals—and indeed, other primates), would happen to differ so substantially—with genetic differences as the primary cause—on such a wide range of traits that happen to be important for professional and financial

[9] Note that hereditarians take different positions about the genetic characteristics of various peoples subsumed under the commonsense racial term "Asian." People from the Indian Subcontinent are thought to be the same as Europeans in their genetic psychological dispositions. Peoples of Southeast Asia are considered a distinct racial group, with levels of intelligence and other socially-desirable traits below those of either Northeast Asians or Whites (Lynn & Vanhanen, 2002).

[10] Comments like these also highlight the disingenuousness of a common hereditarian defense against charges of racism, namely that they (hereditarians) hold that Asians, not Whites, have the highest average IQs of any racial group (i.e., "If we were racist, wouldn't we say Whites are smartest? But we don't, so we're not racist."). But as Rushton makes clear, the differences between Whites and Asians are to be regarded as negligible, and not the sort of thing that makes a difference to the capacity to develop and participate in "complex societies," whereas Blacks are implied to be incapable of such accomplishments.

success and harmonious coexistence in modern human societies.[11] Again, I cannot fully explore the plethora of issues raised by the observation of this "odd coincidence," I merely wish to highlight that, while the empirical and epistemological (and moral) facets of the IQ debate *can* be assessed individually, we should bear in mind that hereditarians' claims about IQ are but one component in an interconnected series of claims that paint a comprehensive picture of stark racial differences in socially desirable characteristics.

Finally, I should also note my reservations about confining the discussion to debates about alleged differences between just two racial groups, White and Black. One reason for this narrow focus should be clear from the above: differences between Blacks and the remaining groups are thought to be the largest and most socially significant. Another reason is that the issues raised by debates about Black-White differences will, I surmise, be most familiar and salient to the majority of readers. Nevertheless, hereditarian claims about the characteristics of Asians (and various subgroups thereof), American Indians, Latinas/os, Jews, and Aboriginal Australians,[12] and the ways in which these claims are intertwined with the history of relations between these groups and White Europeans, and with popular stereotypes about each group, raise a number of issues that are, again, ultimately vital to a

[11] Some researchers, most notably Rushton (1997), claim to have developed an overarching theory of human evolution according to which patterns of racial differences are not coincidental, but rather reflect different evolutionary strategies taken by different ancestral human populations. However, Rushton's grasp of the complexity of human evolutionary history and biological diversity is suspect (when challenged on his division of the human species into three monolithic groups—Black, White, and Asian—he defended his racial classification scheme with the claim that it was obvious that aliens visiting Earth would immediately observe these groups and make the same tripartite division; see Rushton, 1991), and evolutionary biologists have argued that his use of concepts in evolutionary theory to support his model is entirely out of step with mainstream understandings of these notions (Graves, 2002).

[12] The distinction between "races" and "ethnic groups" is of course fuzzy, and while those of us in the United States are perhaps accustomed to thinking of "Jewish" and "Latino/a" as ethnicities rather than races, these groups very much have been and continue to be "racialized," both in popular culture and among hereditarian scientists.

complete understanding of the moral dimensions of racial science. To take one example, it is notable that even while hereditarians acknowledge higher Northeast Asian IQ, there is a pronounced tendency in the writings of certain hereditarians to characterize Northeast Asians as meek, submissive, and sexually inadequate (Lynn, 2013; Rushton, 1997). Again, I think when we consider claims about racial differences in intelligence in the context of these larger claims, we gain valuable perspective about the broader context of the debates. But there is insufficient space to consider all of these issues here.

The takeaway point for this section is that considering only the Black-White IQ dimension of debates about racial differences necessarily leaves out much that is relevant to the empirical and—especially—the moral assessment of hereditarian claims. Going forward I will occasionally make reference to these other alleged areas of difference, where doing so is particularly relevant. Again, however, my main concern will be with debates about the Black-White IQ gap. The next section sets up my critique of the hereditarian position by providing a precise specification of what the core claims of hereditarianism amount to.

5.3 The claims of hereditarianism

Let me begin this section by stating what I consider to be my main target—that is, the hereditarian claim that I am most concerned to reject. The claim, as I express it, will require some explanation, much of which I will put off for later chapters. For now, I'll simply present what I take to be the "core hereditarian claim," or HED:

HED: The genetic differences between racial groups are such that currently-observed psychological and behavioral differences will manifest across the full range of relevant environmental conditions.

While I'll be forestalling full discussion of this "core" claim for the time being, there are a number of elements to the hereditarian position that are worth spelling out. Below I list a series of claims about intelligence and heritability that are key components of the hereditarian position on racial differences. Spelling out these claims in this way will serve the dual purpose of summarizing the hereditarian argument and allowing me to make clear which of the hereditarian's claims I will and will not be arguing against. Here are the claims, listed roughly in order of the degree to which they are accepted in mainstream scholarship on intelligence:

1. General or fluid intelligence (known as *g*) is a stable property of individuals, consisting, roughly, of the ability to process and manipulate complex information efficiently across a wide variety of contexts.

2. Phenotypic intelligence differs among individuals.

3. IQ tests accurately measure the phenotypic intelligence of individuals; that is, an individual's score on an IQ test accurately reflects a stable property of that individual, namely her ability to process complex information. IQ tests are equally accurate measures of *g* in individuals across genders, races, cultures, and socioeconomic classes.

4. IQ becomes fixed in adolescence. While it may be possible to change IQ early in development, adults' IQ is essentially fixed and cannot be significantly modified.

5. Intelligence is a prime determinant of success in modern society. On average, more intelligent people will accumulate more wealth, prestige, and political influence than less intelligent people. Moreover, IQ is a better predictor of professional competence and economic status than any other variable.

6. The mean IQ score among African Americans is about 1 SD (15 points) lower than the mean among Whites.

7. Intelligence shows high *heritability*, meaning that individual differences in intelligence are largely attributable to genetic (as opposed to environmental) differences.

8. Just as individual IQ differences can be attributed to genetic differences between individuals, *group* IQ differences (i.e., the Black-White gap) are attributable to genetic differences between groups.

9. The high heritability of IQ places severe limits on the potential for environmental manipulations to increase an individual's (or group's) IQ.

Most of these claims are have been the subject of significant controversy. Critics of hereditarianism have argued that there is no such thing as general intelligence (Gould, 1983), that IQ tests are inadequate measures of general cognitive ability (Sternberg, 1985), that IQ tests are biased against certain cultural groups (Sandoval, 1979), that IQ is less important than other factors for job performance and success (Ceci & Liker, 1986; Chomsky, 1972), that the Black-White IQ gap has closed to less than 1 SD (Dickens & Flynn, 2006; Nisbett, 2009), that heritability is a misleading measure of genetic influence (Lewontin, 1974; Wahlsten, 1990), and that it is illegitimate to infer from the fact that *individual* differences are genetic that *group* differences are also genetic (N. Block & Dworkin, 1976).

For the sake of the argument, however, I propose to grant all of the above propositions save for the final two. That is, I will not be challenging the claims that IQ tests measure general cognitive ability, that IQ is an important determinant of academic and economic success, that IQ is a fixed quantity in adulthood, or that *individual* differences in

IQ owe largely to genetic differences among individuals. All of these claims appear to enjoy mainstream support among experts in the field, even those who disagree with claims about "innate" racial differences. As we will see, however, I think there are good reasons to doubt that group (i.e., racial) differences in IQ are best explained by appeal to genetic differences among racial groups, and I will argue that high heritability for a trait like IQ need not place any significant limits on the effectiveness of well-designed interventions.

There are as number of consequences that follow from the above concessions, and these are worth spelling out explicitly, as they will help to frame the aims and scope of my argument (of which I will present an overview in the next section). First, if we are to achieve racial equality, we must eliminate or at least minimize the IQ gap. While there are a number of characteristics beyond IQ that contribute to success, and while effort and luck will be sufficient in many *individual* cases to overcome low IQ, it seems unavoidable that *in the aggregate* a racial IQ gap of any significant size will yield differential socioeconomic outcomes in a society that increasingly values intellectual labor and cognitive efficiency.

Second, improving IQ among disadvantaged groups will require intervening early in development. There is some evidence that certain kinds of training (specifically, working memory training) can induce moderate IQ gains in adults, but these gains are slight compared to the size of the Black-White IQ gap. IQ appears to be much more malleable in childhood (early childhood especially) than in adulthood; hence, the best prospects for raising IQ among African-Americans will be found in early interventions.

Combining these two points, I will be assuming that given our present state of knowledge, early-childhood interventions to improve intellectual development ultimately constitute the best (and perhaps the only) route to promoting racial equality.

5.4 A sketch of the argument against hereditarianism

I argue that we ought to reject the hereditarians' core claim, but I present a novel argument for this conclusion. The argument will be developed over the course of this chapter and the two that follow, but the overall outline is as follows. Previous chapters established that our practices of scientific explanation cannot be divorced from our interests; explanation therefore involves the interplay of empirical and pragmatic considerations, where "pragmatic considerations" includes our *moral* concerns. Thus there is a role for values in assessing the worth of a proposed explanation or explanatory strategy. (This is especially true when it comes to decisions about whether a particular phenomenon has been given a "*satisfactory* explanation.") This much, again, has been established in previous chapters.

In keeping with the methodological and explanatory pluralism advocated in previous chapters, I suggest that the hereditarian and mechanistic/developmentalist strategies each constitute legitimate explanatory approaches in the sense of offering explanations that will be sufficient in at least some explanatory contexts. But I argue that we have practical and moral reasons to disfavor hereditarian strategies. This is not to say that the methodological and explanatory strategies employed by hereditarians are *wrong*, per se; rather, I aim to demonstrate that adopting hereditarian explanatory strategies will

compromise important moral values by closing off possibilities for closing racial gaps in achievement and in social and economic capital. Mechanistic strategies, on the other hand, comport with our values because they offer the possibility of discovering interventions that will serve to promote racial equality, and so we are justified in preferring them to the hereditarian alternatives.

Hereditarians will be inclined to call foul on this move, seeing it as evincing an Orwellian mindset that grants license for egalitarian ideals—rather than objective scientific inquiry—to dictate truth on what are clearly matters of empirical fact. That is, proponents of hereditarian scientific positions on the nature of racial differences in social outcomes strenuously resist the encroachment of social or moral concerns into "purely scientific" aspects of nature-nurture debates. It should be clear from the arguments presented throughout this dissertation that I do not think such interest-immunity for scientific practice is possible.

Nevertheless, I acknowledge that there are legitimate and illegitimate ways for our practical and moral concerns to inform our explanatory practices (as expressed in earlier chapters, it's not the case that "anything goes"). In particular, practical and moral considerations cannot give us legitimate grounds for rejecting propositions with overwhelming, unqualified empirical support, nor for accepting propositions that are clearly empirically inadequate. Human communities have, at various points, seen considerable value in a role for humankind as the center of divine creation, but the empirical evidence against geocentrism and creationism is so overwhelming as to swamp

out any value-based considerations: these theories have rightly been rejected, and we have had to learn to live with the implications.

Thus it remains to be shown that HED is not like heliocentrism or evolution, so far beyond any serious doubt that any reasonable person or society must accept it as true and take it as a premise in all relevant practical deliberations. If it were the case, as some hereditarians have suggested, that HED cannot be reasonably doubted, then the value-based considerations I offer against it would be entirely moot. Thus the first step in my argument (Section 5.6) will be to establish that the empirical case for HED—for genetically-determined, permanent racial differences—is considerably weaker than its proponents have supposed; while I do not take a stand on whether HED is, given the available empirical evidence, more or less plausible than its negation, I do argue that the evidence does not rationally require that we accept HED.

Having shown that the empirical status of HED is far from determinate, I go on, in Chapter 6, to demonstrate the moral costs that would ensue from taking HED as true, by looking at what a society that adopted hereditarian conclusions and implemented them wholeheartedly into its policies might look like (where the policy proposals I consider have mostly been offered by hereditarian scientists and philosophers themselves). Again, my argument in favor of deemphasizing, if not discarding, hereditarian explanatory strategies would go nowhere if it were not the case that there are genuine moral costs to adopting hereditarian thinking on a broad scale (that there are no such costs is the position of most hereditarians who have commented on the matter). Thus, having described some of the contours of the "hereditarian society," I provide arguments that this state of affairs is deeply

morally unfortunate (contra hereditarians who argue that a racially stratified society is either evaluatively neutral or indeed a paragon of justice).

I should clarify here what I am *not* saying about the relationship between morality and scientific inquiry. I do not mean to claim that when we are faced with alternative theories that purport to explain the same phenomena, the alignment of each competing theory with our moral commitments (for lack of a better term, the theory's "moral goodness") should be counted among the "theoretical virtues" (e.g., simplicity, explanatory power, coherence, and so forth) that we use to adjudicate theory choice. In other words, I am not claiming that the moral goodness of a theory counts in favor of the theory being *true,* in the sense that it accurately describes the causal relations among phenomena in its domain—alas, the causal structure of the world pays no heed to human moral strictures.

Rather, my claim is that in the case of debates about racial differences, we are not necessarily considering two competing accounts of the causal relations underlying disparate behavioral and social outcomes among racial groups. Rather, we are deciding between two explanatory frameworks that impose different criteria for assessing the adequacy of explanatory statements; that is, hereditarianism treats certain explanations for racial differences as adequate, while the opposing (mechanistic) framework rejects the adequacy of these explanations. No disagreement about the underlying causal structure is implied by this contrast.

5.5 The biological status of "race"

5.5.1 Is discussion of "racial differences" based on a confusion?

Before turning to the substantive empirical matters that divide hereditarians and environmentalists, it is worth considering an objection to the entire enterprise of looking for racial differences. It is sometimes argued that there can be no such thing as "racial differences in intelligence" (or racial differences in anything, for that matter) because there are no such things as races. That is, it is argued that race is at best a social category rather than a biological one, and so insofar as claims about psychological differences between "races" appear to rest upon the reality of biological distinctions between groups, such claims must be false.[13] Hence Bob Richardson, in a review of Michael Levin's (1997b) book defending racial differences in intelligence and criminal proclivities, remarks,

> Biologically, human races do not exist, however important race is as a social category and however much it features in our lives. Consequently, there is no biological basis for racial differences and no point in writing a book on the biological basis of racial differences (Richardson, 2000, p. 847).

Indeed, in the last half-century or so, it has become a point of scientific and philosophical orthodoxy that human "races" do not constitute valid biological categories. The claim has various formulations: races are not natural kinds, there are no biologically interesting subdivisions (i.e., subspecies) within the human species, and so forth. The key

[13] The claim that race is "not real" is sometimes identified with the claim that race is "socially constructed." However, we need not dwell on the tricky matter of what social construction is, and whether and how it relates to the "reality" of categories like race (see Hacking, 1999; Mills, 1998)—we can here restrict our attention to the status of race as a biological category.

move for our purposes is the inference from the non-reality of race to the implausibility or impossibility of psychological differences among races.

Richardson's seems to be a common view (see also Antony, 1993, pp. 143–144; Loury, 2002, p. 111). But I believe this position misrepresents the upshot of the claim (which, broadly speaking, I endorse—see Section 5.5.2) that race is "not real." The argument that race does not exist usually proceeds by way of giving an "error theory" about race (Appiah, 1996): our common-sense talk about race commits us to the existence of a small number of discrete human subspecies, but since there is greater genetic diversity *within* human populations than *between* them (Lewontin, 1972), and because human biological variation is clinal rather than discontinuous, it follows that our discourse about race fails to pick out any real categories in the world; hence race (as the word is commonly used and understood) is an illusion.

Thus, the "error theory" account of race depends crucially on a premise about what our common-sense talk about race commits us to. But racial anti-realism need not (indeed, should not) deny that we can use biological criteria (i.e., morphological and genetic differences) to draw distinctions among human population groups; it merely denies that these population groups meet the standards set by our use of the word "race." For our purposes here, we do not need to settle the issue of what exactly our racial discourse commits us to, or whether the error theory of race is correct. For a defender of group differences like Levin could concede that human populations do not live up to our use of the word "race" (and hence agree that there are no *racial* differences in intelligence, aggression, and so forth), but still maintain that genetic differences among human

population groups are causally responsible for observed differences in psychological characteristics among these groups. Indeed, Levin makes precisely this move early in his book, when he acknowledges the uncertainty of the meaning of "race," and suggests that readers may take the book "not as a discussion of *race* differences at all, but of *differences between descendants of Africans and Eurasians*. Nothing is lost but a word" (1997, p. 22, emphasis in original).

Thus, hereditarian claims have much more to do with the genetic structure of human populations than with "race," per se. And it turns out, unsurprisingly, that with enough genomic data and sophisticated statistical software, it is quite easy for researchers to distinguish genetically among human populations across continents (Bamshad et al., 2003; N. Rosenberg et al., 2002; Tang et al., 2005) and even across countries within a single continent (Novembre et al., 2008).

The upshot of this discussion is simply that insofar as the possibility of "genetically-determined racial differences" with regard to socially-relevant psychological traits has captured the attention of scientific researchers and the general public, the debate does not, *contra* philosophers like Richardson, really turn on whether it is legitimate to apply the term "races" to the populations studied; the debate could proceed in essentially the same way, and with the same momentous social import, even if less semantically loaded terms like "population" are substituted for "race." The non-reality of race does not obviate discussions of group differences in socially-relevant traits.[14] Claims about group

[14] At least, not in the way that many commenters seem to think. One might attempt to undermine investigations of group differences by inquiring as to why, given that race is not a genuine biological category, the scientific establishment nevertheless considers it so *damned important* to investigate whether there are "racial differences." I will offer some discussion of this strategy in Ch. 8.

differences must stand and fall based on what interpretation is given to the idea of "group differences in genetic endowments for a trait," and the empirical evidence that can be adduced for such differences.

5.5.2 "Race" and contemporary genetics

In spite of the non-relevance of the biological reality (or non-reality) of race as a biological concept to debates about innate racial differences, for the sake of completeness it is worth evaluating some modern attempts to vindicate common-sense notions of race using recent advances in population genetics. As noted above, the received view among contemporary biologists and philosophers is that race is not a legitimate biological category. However, a number of scientists, and at least one philosopher, have recently challenged this view. According to this latter camp, new developments in genetics have more or less vindicated the common-sense notion of race. As explored below, the details of human genetic variation are turning out to be quite complex, and little is known for certain. But I take it that a large part of what philosophers can contribute to the debate is to find a good answer (or at least some plausible answers) to the question, *what would genomics researchers have to find in order for them to have found "races"?* While I cannot provide an exhaustive discussion of this question here, this section aims to provide some context to discussions of race and biology by touching on some of the relevant empirical findings and some of the extant debate about what "races" would have to be in order for them to exist. The discussion here is, however, auxiliary to the main thrust of this chapter, and the reader may progress to Section 5.6 without losing the thread of the arguments.

5.5.2.1 The received view and "Lewontin's Fallacy"

Let us call the emerging consensus—the view that human races are not valid biological categories—"racial anti-realism." Classic statements of the anti-realist view can be found in Montague (1964) and Lewontin (1972). Lewontin's argument is perhaps the most widely cited and, being grounded purely in analysis of human genetic variation, the most relevant here.

Lewontin's claim that there are no human races is grounded in his finding that there is more variation within human groups than between them. Using a sample of over 100 global populations and 17 genetic loci, Lewontin found that about 85% of genetic variation is found *within* populations, while 15% is found *between* populations. Moreover, the between-population percentage shrinks to 6% when the populations considered are common-sense racial groupings. Another way of expressing Lewontin's finding is as a *fixation index* (F_{ST}) value, which expresses the ratio of within-population to between-population genetic variation as a number between 0 and 1 (i.e., if the F_{ST} is close to 0, then between-population variation is small compared to within-population variation—that is, populations are not very genetically distinct from one another). Lewontin's analysis corresponds to an F_{ST} value of .15 across all populations or .06 across common-sense racial classifications, whereas a value of > .25 is generally considered necessary to justify division into subspecies (Templeton, 2002). The low degree of between-population variation is sometimes summarized by the claim "two people selected from different 'races' will be no more genetically different than two individuals chosen at random from the same race."

However, papers by Mitton (1977) and, more recently, Edwards (2003) have challenged Lewontin's analysis, to the point that the latter is now often referred to as "Lewontin's Fallacy." The alleged fallacy is Lewontin's use of a measure of variation (F_{ST}) that reflects the *average* degree of genetic difference between populations at multiple loci. Such an averaging procedure, critics argue, is capable of masking clear, systematic patterns of variation that distinguish populations from one another. In other words, Lewontin's procedure would be sensitive to *large* differences in the frequency of *one or a few* alleles in different populations (which critics agree are not found), but it is not sensitive to *small but systematic* differences in allele frequencies at *multiple* loci. Critics also point to Lewontin's use of only a few loci, whereas modern genome-sequencing techniques and statistical software allow scientists to look for population differences using many thousands of loci.

And indeed, when such advanced techniques are applied to human populations, small but systematic differences in allele frequencies can be detected that allow for individuals to be assigned to populations with a high degree of accuracy (whether this amounts to "racial" classification is controversial and is discussed below). Studies in this vein often rely on a statistical package called STRUCTURE, which, given a series of individual genomes, identifies "clusters" of individuals with similar allelic patterns.[15] Using this technique, Rosenberg et al. (2002) were able to identify 5 such clusters that roughly correspond to major geographic regions (Africa, Europe/Middle East, East Asia,

[15] It has been argued, based on simulation data, that STRUCTURE does not reliably identify population structure (Kalinowski, 2011). However, most researchers seem to accept the viability of STRUCTURE's algorithm for discerning genetically distinct populations.

Oceania, and America), while Bamshad et al. (2003) identified three such clusters (Europe, Africa, East Asia). Responding to the claim that genetic population clusters might not correlate with folk or self-applied racial categorizations, Tang et al., (2005) identified four major genetic clusters using samples in the U.S. and Taiwan, which matched self-identified race/ethnicity (Asian, African-American, Hispanic, or White/Caucasian) almost perfectly.

Using a different measure of genetic distance, Witherspoon, et al. (2007) found that if sufficiently many loci are examined from individuals of geographically distant origins, it is—contrary to aphoristic summaries of Lewontin's argument—never the case that two individuals from *different* populations are more similar to one another than two individuals from the *same* population. Some of these scientists (e.g., Risch, Burchard, Ziv, & Tang, 2002; Tang et al., 2005) have argued explicitly that these findings establish the biological legitimacy of traditional racial categories, while others (N. Rosenberg et al., 2002; Witherspoon et al., 2007) caution against leaping to this conclusion. Sesardic (2010) has brought philosophers' attention to these findings by stepping in on the side of the "racial realists," and arguing that "in view of these new studies it becomes harder to accept the widespread but often unsubstantiated claim about the biological meaninglessness of race" (2010, p. 154).

5.5.2.2 But what are races *anyway?*

What are we to make of these claims that, contrary to what seems to be the prevailing opinion of biologists, anthropologists, and philosophers, race is in fact a biological category? Again we come to the issue of the semantics of the term *race*, for

whether or not there are races clearly depends in large part on what we mean by "race." For starters, given that "race" appears to be a putative natural-kind term, we can expect some degree of "semantic deference" on the matter, with ordinary users of the term deferring to "experts" on the precise meaning of the term (just as I might defer to experts on the precise meaning of "acid" or "magnetic field"). So a scientific notion of race need not correspond perfectly with ordinary usage (i.e., scientific racial taxonomy may differ from common-sense taxonomy, or the role of morphology in making racial distinctions may be lesser in scientific usage of the term than in folk usage, and so forth). Importantly, however, defenders of the "new racial biology" claim that modern science *vindicates* the *common-sense* idea of race. And as Glasgow (2003) points out, races cannot be just *any* biologically interesting subdivisions within the human species—it would be absurd to claim that one had vindicated the common-sense idea of race by discovering a "clear genetic difference" between two human populations: those with chromosomes XX and those with chromosomes XY. So the possibilities for a scientific notion of "race" must be *somewhat* constrained by ordinary usage. So what does "common-sense" usage of the term "race" entail?

It is by no means obvious exactly what all "ordinary usage" amounts to, but I take it that contemporary folk ideas of race hold, minimally, that races are major biological subdivisions of the human species (i.e., relatively few in number, usually three to seven), including at least the categories of "African/Black," "European/White," and "Asian." Appiah (1996) has argued for the much stronger claim that the term "race" commits us to the existence of racial *essences*—combinations of somatic, psychological, moral, and

aesthetic characteristics that are inherited together and are biologically fixed—the possession of which is necessary and sufficient for an individual to be a member of a particular race[16] (see also Zack, 2002). Sesardic (2010), however, points out—correctly, I think—that this kind of essentialism as a straw-figure version of the racial realist's claims: such rigid taxonomical conditions (requiring that, in order for a taxonomic category to be legitimate, "all and only" members of that group must possess some characteristic) would rule out even species-level classifications.

So if racial essentialism is too strong a characterization of the race-realist's claim that modern biology vindicates the idea of race, what might the alternative be? According to Sesardic, all that is needed to establish the existence of human races is that there should be human subspecies, and subspecies are simply "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" (2010, p. 148). The existence of such populations seems quite clearly established by the new genetic studies.

However, Sesardic's criterion for subspecies seems too minimal to count as a definition of race that captures ordinary usage. For populations exhibiting this minimal condition (some difference in allele frequencies) are quite easy to find and come in all shapes and sizes. Notably, the number of genetic clusters generated by the STRUCTURE program (*K*) is not "mandated by the data" but rather specified by the researchers in advance. Thus, one can just as easily use STRUCTURE to identify two human populations

[16] Appiah constructs this argument using a causal-historical conception of linguistic meaning, according to which the meaning of terms like "race" for contemporary speakers is fixed by the intensional states of the speakers who have bequeathed the term to us.

("races" in Sesardic's terms) or 20, suggesting that the number of "races" identifiable in this fashion is arbitrary. Moreover, the number and nature of the clusters seems quite sensitive to the particular population samples used. For example, both Bamshad et al. (2003) at $K = 3$ and Rosenberg et al. (2002) at $K = 5$ obtained clusters roughly corresponding to major geographic regions. But at $K = 4$ and $K = 6$, respectively, STRUCTURE generated additional clusters corresponding to small ethnic groups (the Mbuti of Central Africa for Bamshad et al. and the Kalash of South Asia for Rosenberg et al.). Thus it is by no means obvious that common-sense racial categories represent the most salient "objective" subdivisions of the human species as identified by genetic clustering techniques.

Even more embarrassing for Sesardic's "race" criterion, Novembre et al. (2008) were able to separate the population of Europe into a number of distinct genetic clusters (e.g., Iberians, Italians/Slovakians) and could reliably assign the individuals in the study to a fairly precise region of origin (e.g., Poland, Belgium, and even French- or German-speaking areas of Switzerland). While these results were obtained using principal components analysis (PCA) rather than the STRUCTURE program, the point is the same: these groups clearly meet Sesardic's criterion of being "populations that differ among themselves with respect to alternative alleles at multiple loci," and thus according to his definition Poles, Germans, and Swedes are all members of distinct subspecies and hence distinct races. Such a construal of "race" seems completely divorced from ordinary understandings of the term.

There are, however, two moves available to the racial realist that might salvage the idea that the genetic structure of human populations vindicates "race" as a valid biological category. First, Risch et al. (2002, p. 3) offer a more stringent criterion than Sesardic's for "race," namely that race is "based on continental ancestry." They argue that this is the "classical definition of races" and makes the following classifications: "African, Caucasian (Europe/Middle East), Asian, Pacific Islander (for example, Australian, New Guinean, and Melanesian), and Native American" (Risch et al., 2002, p. 3). That is, Risch et al. suggest that races are not just *any* identifiable genetic clusters, but only those corresponding to major continental regions.[17] If Risch et al. have accurately characterized the definition of "race," then it could be argued that this notion is vindicated insofar as these populations are in fact genetically distinguishable from one another. This move, however, would seem to require giving up the idea (also arguably part of the "classical definition" of race) that races are *non-arbitrary* categories; for there is apparently nothing "in the data" that demands we stop at five clusters (corresponding to the "major continental regions") rather than six (and counting the Kalash as a distinct race).

Alternatively, a defender of biological races might suggest the following: although STRUCTURE will generate as many clusters as the experimenter asks for, there is a plausible non-arbitrary way to determine what is the "right" number of clusters warranted by the data. If the genetic data fed to STRUCTURE do not actually reflect *K* genetically distinguishable populations, then it will produce drastically different outputs for multiple

[17] One might challenge whether Risch et al. are entitled to their particular notion of what the "major continental regions" are, inasmuch as Europe and Asia are not distinct continents, nor is Oceania/Pacific Islands really a continent at all. But Risch et al. can probably be charitably interpreted as articulating what most people consider to be the Earth's major geographic regions.

independent runs of the algorithm. And indeed, this is precisely what happened when Rosenberg's research team ran the algorithm at $K \geq 7$ (Bolnik, 2008). A proponent of using genetic clustering to rehabilitate the race concept might therefore suggest that there are precisely as many races (major human subspecies) as can be reliably identified by genetic clustering algorithms applied to a global sample of populations, no more and no less. This makes the question of how many human races there are, and who belongs in which one, a straightforward empirical one, albeit one which we are not in a position to answer with any appreciable degree of confidence (again, the outputs of the algorithm are quite sensitive to the populations utilized—in contrast to Rosenberg et al.'s six clusters, Xing et al. (2010) obtained reliable results up to $K = 12$).

In sum, Sesardic (2010) is probably right that Lewontin's classic (1972) analysis is insufficient on its own to establish that "race" is not a valid taxonomic category. But whether there are, in fact, any races depends on both the subtle conceptual questions of what exactly races are supposed to be, and on complex empirical facts we are only just beginning to uncover. The prospects for developing a genetic theory of race seem to me quite dubious, but a definitive answer must await further developments in the field of genomics and, perhaps, greater conceptual clarity on what, exactly, "race" means.

5.5.2.3 A final possibility: cladistics and race

I will briefly address one final recent attempt from a philosopher to ground racial divisions in modern biological practice. In contrast to the genetic approaches described above, Robin Andreasen (2004, 2006) has defended a *cladistic* view of races. A *clade* or

monophyletic group is a taxonomic group comprising an ancestral population along with all (and only) its descendents. According to analysis from Cavalli-Sforza and colleagues (1994), a phylogenetic tree can be constructed for human populations that identifies monophyletic groups such as Africans, Northeast Asians, Europeans, and Amerindians (see Figure 5.1). Andreasen argues that *races* can be identified with these monophyletic groups. However, Andreasen's theory has garnered little attention (almost all of it critical—see Glasgow, 2003; Marks, 2010; Kaplan & Winther, 2012), as it faces three major difficulties.

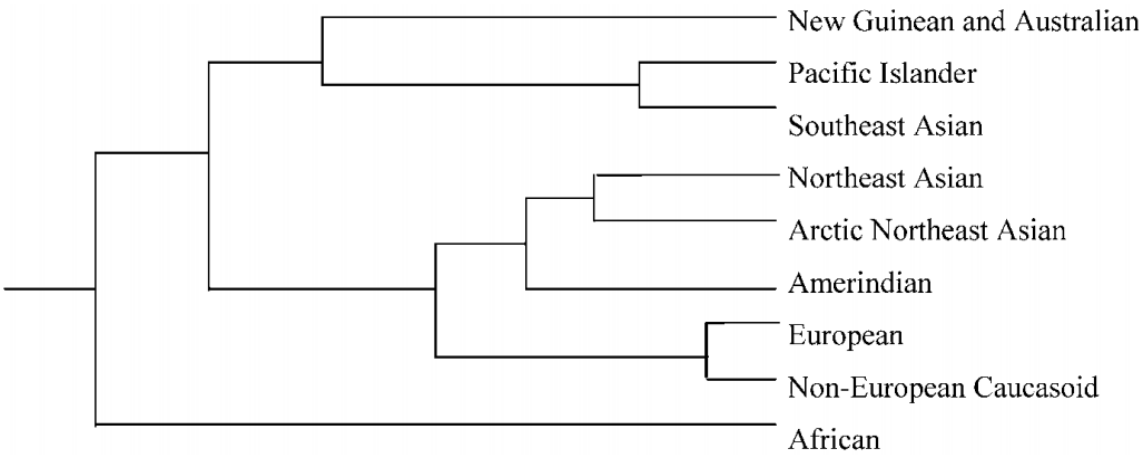

Figure 5.1: Proposed cladistic division of races (from Andreasen, 2004)

First, whether a tree structure with true monophyletic populations can be extracted from the molecular genetic data is contentious, with some (e.g., Templeton, 1999) arguing that the data do not fit a tree structure, and that there are in fact no human monophyletic groups of appreciable size. Second, the particulars of Cavalli-Sforza's tree do a poor job of matching up with ordinary usage of the term *race*, in that, for example, it counts people

from China and people from Thailand as members of separate races. And finally, the cladistic theory entails that people with ancestry from more than one monophyletic group do not belong to *any* race (for they themselves do not belong to any monophyletic human population); this yields, *inter alia*, the result—also entirely contrary to ordinary usage—that most African-Americans (who have some European or Native American ancestry) do not have a race.

5.6 Empirical challenges to hereditarianism

The debate over the causes of the Black-White gap in IQ and other outcomes is long-standing and shows little sign of being resolved any time soon. But it is crucial to my argument that the truth of HED is *not* (as hereditarians generally maintain) so firmly established that we ought to begin acting as though it is true. Thus in the remainder of this chapter I discuss three challenges to the hereditarian account of the race-IQ gap, which will establish that there are good reasons to doubt hereditarian explanations appealing to genetic causes. This is not meant to be an exhaustive review (for recent, extended discussions of the case for and against hereditarianism, see Rushton & Jensen, 2005; Nisbett et al., 2012); Rather my purpose is simply to establish that thoroughgoing agnosticism on this question is warranted. It is simply too early to tell.

5.6.1 Stereotype threat

A substantial challenge for the hereditarian perspective is the phenomenon of stereotype threat. Briefly, this is a well-established psychological effect whereby the

possibility of failure at a high-stakes task (such as an IQ test) generates more anxiety for groups who are subject to negative stereotypes about their abilities in the relevant domain ("Blacks are unintelligent," "women can't do math"). The additional anxiety stems from the perception among stigmatized groups that if they perform poorly on the task, it will be attributed to their identity (as Black or as a woman) and thus taken to confirm the stereotype, whereas no such threat is present for members of non-stigmatized groups (e.g., White males who perform poorly on a math test need not worry that their poor performance will be taken as evidence that "White men can't do math"). This stress reaction to high-stakes performance situations impedes cognitive processing and depresses performance. (For a comprehensive and accessible review of findings in this literature, see Steele, 2010.)

Given the prevalence of negative stereotypes about Black intelligence, and the fact that laboratory studies have established that interventions designed to reduce threat-related anxiety can result in substantial improvements in the performance of stigmatized groups, environmentalists suggest that some (perhaps significant) portion of Black-White IQ differences can be attributed to the additional anxiety Blacks experience in IQ testing circumstances relative to Whites, rather than to any genuine psychological difference (including genetically-caused differences) between Blacks and Whites.

Hereditarians who have addressed the putative role of stereotype threat in generating the Black-White IQ gap note that some psychologists (Sackett, Hardison, & Cullen, 2004) have expressed doubts about the ability of stereotype threat to account for all or most of the Black-White IQ gap. At issue is the way stereotype threat researchers have, in certain studies, statistically controlled for IQ differences among subjects. The

implied hereditarian position appears to be, then, that stereotype threat is a laboratory-only phenomenon, one that plays no role in actual performance situations and therefore poses no threat to the validity of IQ tests as measures of real psychological properties of either White or Black test-takers.

However, this skepticism about the ecological validity of stereotype threat is difficult to maintain in the face of the robust *theoretical* success of stereotype threat research. That is, the phenomenon of stereotype threat can be reliably induced and observed, is consistent with broader, explanatorily successful theories about the effects of anxiety on cognitive performance, and can be reliably manipulated in laboratory settings. In short, Sackett et al.'s misgivings notwithstanding, the potential effects of stereotype threat on Black IQ-test performance cannot be discounted.

5.6.2 Intelligence, working memory, and stress

Recent work on the neurocognitive underpinnings of intelligence have found robust correlations between measures of working memory (WM) capacity and traditional measures of psychometric intelligence (Alloway & Alloway, 2010; Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen, 2004; Conway, Kane, & Engle, 2003). While some studies suggest that WM and fluid intelligence are essentially identical (Colom et al., 2004), we do not here need to establish the precise nature of the relationship between WM and intelligence. The relevant point for our purposes is simply that, unsurprisingly, measures of intelligence and measures of WM seem to tap into a common cognitive

resource, one that is responsible for maintaining and manipulating active representations and suppressing irrelevant information.

These findings become relevant to the race-IQ debate when combined with other recent investigations of both the situational and long-term determinants of WM capacity. One such strain of research establishes that physiological responses associated with elevated stress significantly impair WM capacity. According to models developed by these researchers, elevated stress levels have both short- and long-term effects on WM capacity. In the short term, stressful life events have high cognitive salience and function as distractors that must be inhibited while directing attention to a task—in effect, stress increases cognitive load (Klein & Boals, 2001; Schoofs, Preuß, & Wolf, 2008). Meanwhile, the presence of physiological stress markers has been shown to have long-term adverse effects on the development of brain systems—e.g., the hippocampus and prefrontal cortex—that are known to be crucial in the operation of WM (Evans & Schamberg, 2009; Hackman, Farah, & Meaney, 2010).

Crucially, these physiological stress markers have been shown to be chronically elevated in both parents and children in positions of low socioeconomic status, and elevated to even greater levels among African-Americans—even after controlling for the effects of SES (Mays, Cochran, & Barnes, 2007; Troxel, Matthews, Bromberger, & Sutton-Tyrrell, 2003; Turner & Avison, 2003). It is hypothesized, specifically, that the persistent burden of slight or ambiguous instances of racial discrimination—often called "microaggressions"

(Sue, Capodilupo, & Holder, 2008)—is responsible for these chronically elevated stress responses among Blacks.[18]

From these findings it is reasonable to infer that the Black-White IQ gap may be mediated at least in part by a subtle but systematic difference in the environmental conditions of Blacks and Whites: the continual presence of stressors generated by racial stigma in Blacks' environments impedes the development and functioning of WM, thus depressing IQ scores.[19] And, consistent with the models proposed in the studies reviewed above, environmental interventions—specifically, early-childhood educational programs that provide additional support to low-SES children and their parents—can reduce stress (both parents' and children's) and improve children's attentional capacities and—crucially—intelligence test scores as well, at least in the short term (Fisher, Stoolmiller, Gunnar, & Burraston, 2007; Neville et al., 2013). While these interventions were directed at low-SES children across racial groups and have not yet explored the prospects of such manipulations for reducing the Black-White IQ gap, this line of research lays the foundation for a very promising and, importantly, theoretically well-grounded strategy for addressing the IQ gap.

[18] Another, related line of research suggests that training on WM tasks may well improve general intelligence in children, at least for some individuals (Hsu, Novick, & Jaeggi, 2014; Jaeggi, Buschkuehl, Jonides, & Shah, 2011; Karbach & Unger, 2014). This research is in its preliminary stages, and it is not clear precisely what factors influence the effectiveness of WM training in raising intelligence test scores for an individual, or how durable the effects are over the long term, but these results clearly suggest that certain kinds of environmental variations—e.g., the presence or absence of experiences that promote the enhancement of WM capacity—can generate individual differences in intelligence during development, and that intelligence may not be as impervious to interventions as hereditarians suggest.

[19] Note, however, that the precise hypothesis offered to explain differing levels of chronic stress doesn't matter for the main point here: the hereditarian hypothesis is impugned so long as the stress differences are environmental in origin—and it would, I submit, be bizarre to suppose that Blacks are genetically predisposed to experience greater stress than Whites when exposed to identical environmental conditions.

To my knowledge, hereditarians have not responded directly to the evidence for the role of stress-inducing environmental conditions in generating the Black-White IQ gap. Two responses seem likely, however. First, hereditarians have long maintained that environmental differences—including those arising from racial prejudice—cannot explain the IQ gap because the gap remains even when environmental variables are taken into account, and because social changes since the Civil Rights Era have minimized the role of racial prejudice in American society (Herrnstein & Murray, 1996; Jensen, 1969; Rushton & Jensen, 2005). But the environmental factors for which previous analyses have controlled are only very course-grained measures of broad socioeconomic indicators (Glymour, 1998), and would be insensitive to the effects of differential levels of chronic stress revealed in recent studies. Moreover, the heightened levels of stress observed among African-Americans give the lie to the already dubious premise that racial stigma has ceased to affect minority groups in any way that might matter for the development of cognitive ability.

As for the second likely hereditarian response, it is true that previous interventions have also resulted in significant improvements in IQ scores among disadvantaged youth (as defined by race or by SES alone) in early childhood, but these effects have faded as the targeted cohorts have progressed through the school system, such that recipients of early childhood interventions show little or no advantage in IQ by mid to late adolescence (Herrnstein & Murray, 1996; Nisbett et al., 2012; Scarr & Weinberg, 1983).[20] Thus, it

[20] Hereditarians typically attribute the diminishing longitudinal effects of early interventions to the fact that children gain progressively greater control over their own environments as they age, which results in those with "genes for low IQ" seeking out less cognitively stimulating environments after graduating from

would be premature to assume that the IQ gains engendered by recent interventions will be permanent. It is worth noting, however, that recent interventions that target known environmental influences on WM and attentional control (e.g., physiological stress responses) are, plausibly, more likely to yield sustained results than were their predecessors. This is because these recent interventions are embedded in a comparatively sophisticated and well-supported theory of both the operations and the development of the neurocognitive underpinnings of intelligence (Buschkuehl & Jaeggi, 2010, p. 266), while earlier interventions took a less targeted approach and sought to raise IQ merely by providing a more all-around enriching environment. (I will return to this point in Chapter 7.)

5.6.3 The Flynn Effect

Finally, let us briefly consider what has come to be called the "Flynn Effect," which is the phenomenon of persistent rising mean IQ scores worldwide, stretching as far back as IQ tests have been used (named for psychologist James Flynn, who brought the phenomenon to the attention of the scientific community in the 1980s; Flynn, 1984, 1987). Because IQ tests are consistently re-normed (to maintain a mean of 100 and a standard deviation of 15), a particular score on an IQ test administered at one point in time does not necessarily indicate an identical level of cognitive performance as the same score on a test administered at another time. The surprising fact is that a score of 100 on an IQ test

intensive early childhood interventions, thus diminishing IQ. This would be an example of *gene-environment correlation* (rGE), which will be discussed in Section 7.2.2.1.

administered in 1900 would correspond to a score between 50 and 70 measured against current norms (Flynn, 2007, pp. 9–10). And, not only have these IQ gains been observed around the world (Flynn, 1987), but the gains have been most significant in less developed parts of the world (which tend to exhibit lower mean IQ), and, within the United States, the gains have been most significant among African-Americans, thus closing the IQ gap somewhat (Dickens & Flynn, 2006).

The Flynn Effect is deeply puzzling on the assumption that IQ tests are reliable, culture-and-context-invariant measures of real-world cognitive skills. Given that scores of around 70 are considered a bare minimum for normal cognitive functioning, taken at face value the Flynn Effect implies that generally speaking, our grandparents and great-grandparents lived in a society where the average person was so cognitively impaired as to be scarcely able to walk to the store for a gallon of milk. As this suggestion is clearly absurd, something else must be going on with IQ tests.

Just what this "something else" might be is beyond the scope of our discussion here (see Flynn, 2007 for extended discussion), but the most reasonable explanation is some combination of (a) inherent shortcomings in IQ tests (i.e., they are not as reliable indicators of functional intelligence as their most vigorous proponents suggest, and may be to some extent artifactual) and (b) actual changes in cognitive performance brought about by environmental changes over the last several generations (e.g., modern test takers' improved facility with abstract thinking and improved brain development as a result of better nutrition). The technical issues that arise in attempting to characterize the precise nature and causes of the Flynn Effect are significant (see te Nijenhuis & van der Flier, 2013;

Wicherts et al., 2004), and there are those who question the effect's validity (e.g., Rodgers, 1998), but nevertheless we can extract some important lessons from the Flynn Effect even without resolving these questions.

In my estimation, the relevance of the Flynn Effect to the race-IQ debate is that it may be seen as an "existence proof" for something that the hereditarian theory has a difficult time explaining: not only have environmental conditions exerted a very large effect on levels of phenotypic intelligence, but some groups have benefited more than others—that is, certain environmental changes have closed the gap between racial groups. This last point is significant because hereditarians have sometimes argued that *even if* we could raise mean IQ among the Black population by providing, say, more enriching school environments, this would have little or no effect on the *practical* outcomes of group inequality, for the enriching scholastic environments would simply exert a similar upward effect on the IQ of *White* students, thus holding the gap steady and leaving Whites with the same relative advantage in educational and professional environments as they had before (Herrnstein & Murray, 1996; Levin, 1997b).

CHAPTER 6: THE MORAL STATUS OF RACIAL-DIFFERENCE DEBATES

6.1 Introduction

In this chapter I detail what I take to be the moral costs that are likely to result from adopting HED. I begin, in this introduction, by distinguishing my own argument for disfavoring hereditarianism from another that is common in the literature and bears some similarities to my own, followed by some comments about how hereditarian researchers have presented (or failed to present) the policy proposals that they take to be supported by hereditarian science. Then, in Section 6.2, I describe what the "hereditarian society"—a society that adopted the idea of permanent racial differences as a premise in its policy deliberations—might look like. Section 6.3 provides extended arguments that the hereditarian society exhibits a number of morally problematic features.

6.1.1 On the "argument from risk"

The reader might note a similarity between my argument and a popular move utilized by moral critics of hereditarianism, which is to urge that the potential harms of affirming racial differences, when there are in fact no such differences, are much greater than the potential harms of failing to affirm such differences if they *do* exist (i.e., the costs of a "false alarm" are much greater than those of a "miss") (Boxill, 1978, p. 253; Kitcher, 2001, pp. 97–99). In other words, proponents of the risk argument have taken the view that if we believe there are racial differences but there are none, the consequences will be

disastrous, whereas if we believe there are *no* racial differences when in fact there are, very little will be lost. My own argument, while similar in spirit to the classic argument from risk, is more nuanced and, I believe, more effective in that it (a) acknowledges extant normative disagreement about the (un)desirability of various social arrangements, and *argues for*, rather than *assumes* key normative premises; and (b) shows how our normative commitments positively recommend a *substantive theoretical orientation* and correlative *research methodology*, rather than merely recommending a suspension of judgment.

Again, although overall I am sympathetic to something like the argument from risk, I think that its proponents have at times assumed that the case for moral asymmetry between mistaken hereditarianism and mistaken environmentalism is stronger than it actually is. This has occurred, I think, because proponents of the risk argument have generally not taken seriously hereditarian arguments for the *good* social consequences of acknowledging racial differences (if such differences actually existed), and the *bad* consequences of *failing* to acknowledge them (again, assuming they did exist) (i.e., the benefits of a "hit" and the high costs of a miss).

Hereditarians typically cite three types of negative outcomes that would result from failing to accept racial differences and taking account of them in our policymaking (Gottfredson, 2000, 2005a; Miele & Sarich, 2005). First, they argue that since racial differences are immutable, any public resources (money, time, and effort) directed towards eliminating them are entirely wasteful. This waste of public resources, they emphasize, makes everyone (including those the compensatory programs are aimed at helping) worse off.

Second, failing to acknowledge racial differences encourages us to unfairly assign blame for disparate racial outcomes. If racial groups differ in the ways hereditarians suggest, we can expect that organizations utilizing strictly merit-based selection criteria will enroll or hire significantly disproportionate members of different groups. And if we continue to take disproportional selection as evidence of discrimination, those decision-makers will be unfairly accused of moral failings they do not exhibit.

Third, failing to take account of racial differences has more direct negative effects on allegedly less-able groups: by assuming that all racial groups have equal intellectual and moral potential, we harm the less-able groups by (a) holding them to unrealistic standards, which will lead only to frustration on all sides and lower self-esteem for the cognitively disadvantaged; and (b) missed opportunities to provide members of certain groups with education and moral guidance that is more appropriate for their genetic endowments. In other words, educating Black children the same way we do White children (on the assumption that their intellectual and moral potentials are equal), leads to Blacks ending up *less* functional in society than they would be if we acknowledged the differences between groups and educated Blacks in a way that allowed them to reach their full potential (again, though, recognizing that this potential will be less than that of Whites; Gottfredson, 2005a; Jensen, 1969, 1998; Rushton & Jensen, 2005).

Fully assessing the plausibility of these suggestions is beyond the scope of this chapter, but they strike me as sufficiently reasonable concerns that they would have to be taken into account by any variant of the argument from risk, and they appear to weaken the

case for the claim that negative consequences of "getting it wrong" about the causes of racial differences is a problem *only* for hereditarians.

The key difference between the traditional argument from risk and my own is that the traditional argument assumes that since adopting hereditarian presumptions is "morally riskier" than adopting egalitarianism presumptions, we should demand a higher standard of evidence for accepting hereditarianism than would otherwise be required (much as we might demand a higher standard of evidence regarding the salt content of a meal if that meal is to be served to someone with severe hypertension). That is, the traditional argument from risk asserts merely that we should *suspend judgment* while we continue to collect evidence, and that we should demand *more of the same kinds* of evidence before accepting hereditarianism (e.g., performing more behavior-genetic studies with larger sample sizes or that control for more variables). In contrast, my argument proposes that we should entirely rethink our methodological and theoretical orientation to questions of racial differences.

6.1.2 Identifying hereditarian policy recommendations

As we will see, much of my argument in the latter part of this chapter will turn on my identification of a particular evaluative attitude as implicit in mainstream hereditarian arguments—specifically, an attitude towards the sort of society that is likely to result if we were to adopt HED. In other words, given that it is an oft-repeated hereditarian mantra that denying HED can only lead to bad policy and bad outcomes, we would like to know just what hereditarians take to be the normative status of policies and outcomes that might result

from *adopting* HED. Critics of hereditarianism, of course, are not shy about expounding upon the potential harms of disseminating hereditarian claims. But the hereditarian, evidently, would have us believe that the likely consequences of *adopting* HED are all-things-considered preferable to those of *denying* HED. Unfortunately, however, most hereditarian writers have neglected to acknowledge potential harms of accepting HED or engage in explicit comparison of potential harms and potential benefits.[1] Because direct textual evidence of hereditarians' attitudes towards these outcomes is lacking, I will begin by calling attention to some passages in which hereditarians draw near this issue but seem to skirt round it at the last minute, forcing us to try to fill in the gaps.

Let me be clear that this is not merely an exercise in armchair psychology, an attempt to discover what individual hereditarians hold in their heart of hearts. Nor is this an attempt to score a series of cheap points by catching hereditarian writers in moments of unclarity. Rather, my aim here is to motivate the endeavor taken up in Section 6.3, where I attempt to make explicit what mainstream hereditarian writers have left implicit, namely the sorts of social conditions we can expect to ensue from widespread adoption of HED.

Again, hereditarians are fond of pointing to the alleged harms caused by—as well as the sheer epistemic indecency of—our society's current state of collective denial about racial IQ differences. But to the extent that foreseeable *negative* consequences of adopting a hereditarian outlook and incorporating it into our social policy are acknowledged in hereditarian writings, they are referenced only obliquely. For instance, Herrnstein and

[1] As we will see in the next section, there are some notable exceptions to this trend—perhaps unsurprisingly, hereditarian philosophers tend to be more explicit about the normative implications of their views than their counterparts in the social sciences.

Murray make only occasional, vague references to the potential harms of a book like *The Bell Curve*, such as the following:

> We are not indifferent to the ways in which this book, wrongly construed, might do harm. We have worried about them from the day we set to work. But there can be no real progress in solving America's social problems when they are as misperceived as they are today. What good can come of understanding the relationship of intelligence to social structure and public policy? Little good can come without it. (Herrnstein & Murray, 1996, p. xxiii)

Similarly, hereditarian psychologist Linda Gottfredson's (2005b) commentary on Rushton and Jensen's (2005) hereditarian manifesto neatly captures many of the dynamics of discussion surrounding the social implications of hereditarianism. In her forceful rejoinder to critics who are wary of the social implications of hereditarian research, she critiques existing practices aimed at equalizing racial outcomes:

> Currently, racial parity in outcomes is often treated as the ultimate standard for fairness and lack of parity as a measure of White racism. For instance, disparate impact in hiring is prima facie evidence of illegal discrimination in the United States, with employers, if sued, then needing to prove themselves innocent. By undermining culture-only explanations of racial inequality, the "provisional truth" of Rushton and Jensen's (2005) hereditarian hypothesis thereby undermines the moral legitimacy of all rationales for racial equalization that posit White misbehavior as its cause. That it might persuade the public to temper or abandon its efforts to close all racial gaps in success and well-being is surely what inflames critics most. (Gottfredson, 2005b, p. 317)

Again, there appears to be a (perhaps understandable) hesitancy on Gottfredson's part to state in plain English what she clearly takes to be a straightforward implication of hereditarian scientific conclusions: Blacks will, forever and always, fail to attain full social

and economic equality with Whites. Rather than simply acknowledging that this permanent inequality is a consequence of the hereditarian position, Gottfredson quickly diverts the reader's attention to critics' "inflamed" responses and their accusations of "White racism" and "White misbehavior."[2] Nathan Glazer makes a similar observation regarding an article in which Herrnstein (1990) likewise argues for the inevitability of racial inequality but shies away from stating this claim directly: "I believe…[Herrnstein] was trying to avoid stating (though all the necessary evidence was in his article) an explanation of black-white differences that simply leaves men of good will helpless" (Glazer, 1994, p. 16)

Given hereditarians' reticence to name or discuss the potentially problematic social implications of their scientific positions, this job has been left to their critics, which they have undertaken vigorously—often, it must be admitted, with excessive zeal. Hereditarians have understandably taken umbrage at some of the more defamatory accusations, as Gottfredson illustrates:

> Rushton and Jensen (2005) themselves acknowledge that open discussion of genotypic ability differences between the races might harm race relations. Their most vocal critics predict far worse. Widespread acceptance of the hereditarian hypothesis would, they say, put us on the slippery slope to racial oppression or genocide…The critics' predictions of mass moral madness, like their frequent demonization of scientists who report

[2] I should note that Gottfredson here makes a very unfortunate error—exasperatingly common among hereditarians—in relying on an incredibly simplistic and outmoded conception of "racism" and racial bias. She assumes that the only two possible explanations for racial disparities in outcomes are Blacks' inferiority and "White misbehavior." There is an enormous psychological literature establishing that people who sincerely avow egalitarian racial attitudes can nevertheless exhibit strong implicit biases (typically against Blacks), and that this implicit bias can manifest in discriminatory behavior, without the subject being aware of or able to control the bias. Therefore, *pace* Gottfredson's implication, it is entirely possible that hereditarian explanations for racial disparities in outcomes are false, but that "White racism" and "White misbehavior" (in the sense of willful racial antipathy or prejudice) play only minimal roles in the perpetuation of these disparities.

> unwelcomes racial differences, seem mostly an attempt to stifle reasoned discussion. (Gottfredson, 2005b, p. 317)

It is true that critics of hereditarianism have frequently referenced Nazi scientists and genocidal policies, but I will not be doing so here. My aim, rather, is to answer Gottfredson's call for "reasoned discussion" by exploring the more reasonable policy implications of hereditarianism. No modern hereditarian scientist has or would endorse genocide against less intelligent racial groups. But hereditarians have, or reasonably might, endorse the policies discussed below. Since hereditarians often—justifiably, I think—complain that their motivations are unfairly impugned and that their views on the social implications of their scientific positions are caricatured, I will endeavor to steer clear of any inflammatory accusations or absurd extrapolations of hereditarian views. But at the same time, given that hereditarians themselves often seem reluctant to be fully transparent about certain implications of their views, part of my task will be to draw out these implications and set them under good, clear light, so that we may see how we like them.

6.2 The hereditarian society

I will discuss four areas of policy to which scientific claims about racial differences are relevant (roughly, from less to more controversial): (1) educational programs aimed at closing the Black-White IQ gap, (2) what to do about differential representation of various groups in positions of prestige, power, and wealth, (3) discrimination in employment, housing, and criminal justice, and (4) allocation of foreign aid. I describe, in effect, what a society that wholeheartedly embraced hereditarian science might look like. I should

emphasize, however, that I am not claiming the consequences described below are *inevitable* if hereditarian conclusions are accepted. The policies outlined below would all require justification not merely from the empirical claims of hereditarians, but also substantive normative premises—about the value of economic efficiency, the enormity of wasting taxpayer dollars, and so forth—that many would be inclined to reject in the first place. In other words, a society that accepted hereditarianism as a scientific matter might, after a process of deliberation, still find at least some of these policies in conflict with its values. And indeed, those who endorse hereditarian scientific conclusions are not univocal in supporting these policies (and it is quite possible that some hereditarians would object strenuously to some of the more radical proposals discussed). But the fact remains that once permanent, ineradicable racial differences are taken as empirical premises in policy deliberations, the proposals discussed below are no longer outside the bounds of reasonable social arrangements, given the political values currently represented in society. Thus, because I consider (for reasons given in Section 6.3) these policies to be deeply morally problematic, my overarching aim throughout this and the next chapter is to entirely *preclude* these deliberations by staking out a scientific stance that forestalls the acceptance of hereditarianism in the first place.

6.2.1 Compensatory education

If hereditarians are correct that Blacks will exhibit lower intelligence in the full range of environmental conditions likely to manifest in current or foreseeable human societies, then social and educational programs aimed directly at closing the Black-White

IQ gap are entirely without justification. The Head Start program, which provides free pre-kindergarten education to low-income children, is the focus of much discussion on this point, with certain hereditarians—especially those of a libertarian political bent (e.g., Herrnstein & Murray, 1996)—arguing for the cessation of the program on the grounds that it is a gigantic waste of money. Such an argument depends not just on hereditarian positions regarding the science of IQ, but also an assumption that no other benefits (e.g., improved social skills or improved access to adequate nutrition) can be expected to accrue to children participating in the program. The specific goals and outcomes of the Head Start program (or other similar programs) are beyond the scope of this discussion, but the relevant point is this. If hereditarians are right that environmental manipulations cannot eliminate racial differences in IQ, there is nothing to be gained, and much to lose, by directing resources to programs (or particular aspects of larger programs) that have as their explicit, sole purpose the elimination of racial IQ differences. If hereditarian claims are true, cutting such programs is a no-brainer, whatever one's values.[3]

[3] Note that I am setting aside the question of affirmative action, which is a rather more complicated case. Some arguments for affirmative action frame such programs as a form of compensation for past discrimination. These arguments, if otherwise successful, would not be impugned by hereditarianism, since no one disputes that unjustifiable racial discrimination has occurred in the past. Other arguments for affirmative action, however, view such programs as aimed at countering the effects of *current* discrimination, on the hypothesis that discrimination is what accounts for disparate social outcomes among racial groups. Hereditarian claims (e.g., that Blacks are genetically less intelligent) *do* undermine these latter arguments, since they suggest that disparate outcomes result from innate differences rather than from discrimination. Thus, from the hereditarian's perspective, if affirmative action programs succeed in their goal of equalizing outcomes between Blacks and Whites, this will have come only as a result of offering positions (admissions to universities, jobs, promotions, etc.) to less-qualified Blacks. Therefore, to the extent that the case for affirmative action rests on the proposition that unequal outcomes are due to discrimination, scientific claims of innate racial differences appear to offer a strong *prima facie* case for eliminating or scaling back affirmative action programs.

Importantly, and in fairness to hereditarians, the above arguments do not entirely preclude *all* programs aimed at reducing the Black-White IQ gap. For while it is a core hereditarian position that, owing to genetic differences between Blacks and Whites, this gap cannot be closed entirely, hereditarians may acknowledge (and many do) that manipulable environmental factors play *some* role in generating the current 15-point IQ gap, and therefore that certain manipulations (i.e., social programs) might shrink the gap somewhat (though hereditarians are typically duly agnostic on how much the gap might be closed and how it might be done). The upshot of hereditarian arguments on this score is that although we might well continue to implement manipulations that seem (on intuitive or evidential grounds) promising in closing the IQ gap, we should be prepared for the fact that we will, sooner or later, arrive at a point when virtually all of the remaining variance in Black-White IQ differences is genetic in origin, and subsequent attempts to close the gap further will be mere wastes of resources (and, in the interests of protecting public resources from misuse, we should always be on the lookout for evidence that we have reached such a point).

6.2.2 Under/overrepresentation

It is a well-known fact that in American society, different racial groups are differentially represented in various social positions of high and low value. For instance, African-Americans are underrepresented among students at prestigious universities and in prestigious well-paying professions, and overrepresented among the incarcerated and the chronically impoverished.

A note about the terms "underrepresentation" and "overrepresentation" is in order. These terms may be used in multiple senses, and authors are often unclear as to which sense they intend. In one sense, to say that a group is "underrepresented" simply means that there are fewer members of a particular group in a certain demographic category (e.g., in a particular profession or a particular tier of colleges) than would be expected if people were assigned to such categories at random (i.e., if women make up 50% of the population, and fewer than 50% of legislators are women, then women are underrepresented among legislators). In another sense, a group is underrepresented if fewer members of that group attain some outcome than *deserve* to or *ought* to attain it. In this sense, if Blacks make up 15% of the overall population, but only 5% of college professors, then Blacks are underrepresented among professors *only* if the factors accounting for their lower representation are in some way illegitimate (e.g., prejudice and bias), rather than being traceable to qualities that make Blacks, on average, less interested in or less suitable for academic careers. Since I do not wish to beg any questions about the legitimacy of various reasons for differential representation, I will use these terms exclusively in the first, normatively-neutral sense.

It is commonly assumed that underrepresentation of Blacks in highly-valued social positions, and their overrepresentation in disvalued positions, are social problems to be remediated, with much discussion and disagreement about what are the most effective strategies for addressing the problem. Under a hereditarian outlook, however, there need

be nothing inherently problematic about differential racial representation in social roles.[4] For if hereditarian conclusions about the causes of racial differences in intelligence are combined with the (independent, though commonly endorsed by hereditarians) thesis that intelligence is an important—perhaps *the* most important—determinant of success in current and foreseeable human societies, we can expect that unless aggressive policies of affirmative action and/or wealth redistribution are implemented, African-Americans will remain underrepresented at prestigious academic institutions, in prominent and intellectually demanding professions, and in the wealthiest income brackets in the United States. (Looking internationally, we can also expect that prospects for economic development in African nations will be substantially lower than those for European or Asian countries.)[5]

Hereditarians are fond of insisting that hereditarian conclusions are themselves morally neutral, because while their scientific research can "give us the facts," it cannot "determine our goals" (Gottfredson, 2005b; Rushton, 1997). Thus it is possible, in theory, for a society both to endorse hereditarian conclusions on the science of IQ and race *and* adopt policies that would serve to equalize the representation of different racial and ethnic groups in positions of power and prestige, as well as in various socioeconomic brackets.

---

[4] E.g., Sesardic: "Racial inequality in itself does not constitute a fact that is morally condemnable. Whether it is condemnable depends on the origin or source of this inequality" (2005, p. 214).

[5] The hereditarian position on IQ may also entail that—absent radical changes to the criminal justice system—Blacks will continue to be overrepresented among the incarcerated, since some (e.g., Herrnstein & Murray, 1996) have argued that lower intelligence leads—via direct or indirect means—to increased tendencies toward criminal activity. Other hereditarians, like Levin (1997b) and Rushton (1997) also suggest that we can expect that Blacks will always be overrepresented among criminals, but offer a more direct explanation, namely genetic tendencies for Blacks to be more aggressive and less pro-social and future-oriented (a hypothesis that also has implications for the prospects for stable democratic governance in sub-Saharan Africa, as argued by Lynn & Vanhanen, 2002; and Rushton, 1997).

For instance, there could be a policy mandating that the distribution of ethnic groups among employees at all companies of a certain size, and among the student bodies at all tertiary academic institutions, must reflect the distribution of those groups in the population at large, lest the firm or school be subject to a large fine. Since this policy alone would not guarantee equal representation of groups among the higher tiers of the corporate structure, if a particular group were found to be underrepresented among CEOs, then a lottery would be held to determine which firms would have to replace their CEOs with the most qualified available candidate from the underrepresented group.

While such solutions are, strictly speaking, not ruled out by hereditarian science, we can reasonably assume that they would be rejected by any society remotely resembling our own. Such measures seem prohibitively intrusive and unwieldy, conflict with most people's notions of fairness and meritocracy, and, if hereditarians are right about the relationship between IQ and professional competence, would impose enormous costs to overall economic productivity (Gottfredson, 2000). In this connection, it is notable that while Gottfredson (2005b), in insisting on the value-neutrality of hereditarian science, mentions various "compensatory" measures we might take to assist low-IQ individuals (disproportionately Black), none of her suggestions would serve to eliminate current inequalities in outcomes.

Hence, while it is true that hereditarian conclusions do not *necessarily* entail the continuation of social, political, and economic inequalities separating American Whites and Blacks, we can expect that if hereditarians are correct about the sources of current gaps, and if there are no radical revisions to the way modern societies think about notions like

fairness, desert, and distributive justice, then we can expect that Blacks will, overall, continue to claim only a minimal share of political and economic power, and will continue to be underrepresented in prestigious professions and overrepresented among the incarcerated and chronically impoverished.

6.2.3 Discrimination

As with "under/overrepresentation," the term "discrimination" has multiple senses, one neutral and one overtly normative. In the normative sense (the sense in which the term is typically used in everyday discourse), "discrimination" refers to the practice of singling out members of particular races/ethnicities, genders, religions, and other "superficial" categories for *systematically inferior* treatment (i.e., denying them benefits afforded to members of other categories of similar talents, efforts, and so forth). In the neutral sense, "discrimination" simply means treating members of different categories differently. As with "under/overrepresentation," I use "discrimination" in the neutral sense.

It should be clear that once a hereditarian perspective on group differences is assumed, it is only a few short steps to the conclusion that there is nothing in particular wrong with racial discrimination in itself. For while skin pigmentation may not *itself* be relevant to job performance, criminal inclinations, and so forth, intelligence *is* (we are assuming here) relevant to these outcomes, and skin pigmentation is, if hereditarians are correct, a probabilistic indicator of intelligence. Skin pigmentation could, therefore, be used as a source of important information about the talents and character of individuals. In the present context, then, "racial discrimination" refers to the practice of using race as an

evidentiary factor in such decisions as those concerning housing (e.g., whether to rent or sell a residence to an individual, and how much to charge), employment (hiring, promotions, and compensation), and criminal justice (whom to detain and investigate as suspects, whom to convict on what evidence, and how to mete out punishment for crimes).

Some hereditarians (Levin, 1997b; Sesardic, 2005) have indeed argued that because race is probabilistically associated with various socially-relevant characteristics, it is reasonable and appropriate for members of society to use race as a source of evidence in determining one's interactions with other members—for instance, by operating with a presumption that White job applicants will be more competent than Black applicants, or by more readily detaining and/or convicting Black individuals suspected or accused of crimes.[6]

It is worth noting that a hereditarian position on the *causes* of racial differences in such outcomes is not, strictly speaking, necessary for the argument in favor of discrimination; one could argue that if, for example, different racial groups display different dispositions to commit crimes, then *whether or not this difference can be traced to any "innate" or "genetic" factors,* it is a good idea to use race as an evidential factor in

[6] The oft-cited fact that Blacks are overrepresented among those convicted and incarcerated for crimes does not, of course, necessarily mean that Blacks *commit* more crimes—this pattern could be observed, for example, if law enforcement officers, juries, and judges are much more likely to arrest, convict, and deal harsh sentences to Black criminals than White criminals, or if innocent Blacks are much more likely to be arrested and incarcerated than innocent Whites. While there is plenty of evidence to suggest that these latter forces are likely at work, assessing just *how much* of the disparity in crime and incarceration rates can be attributed to them is difficult. For the purposes of this discussion, then, I will grant to the supporter of racial discrimination ("racial profiling," as it's called in criminal contexts) that Blacks do commit a disproportionate number of crimes. It would be remiss not to point out, however, that "crime" is hardly a natural-kind term: what activities are taken to be "criminal" is up to a society to determine, and a society can quite readily make "criminals" out of a certain segment of its populace by criminalizing particular activities (such as homosexual acts or recreational use of certain drugs—but not others).

deciding whom to arrest and whom to convict (W. Block, 2010; Epstein, 1994). However, I include racial discrimination in this discussion of the policy implications of hereditarian positions for two reasons. First, because these policy proposals have frequently been put forward by hereditarian thinkers themselves (though they sometimes seem not to notice that hereditarian arguments and pro-discrimination arguments are, strictly speaking, independent[7]). And second, because the case for racial discrimination is much more direct if the differences are expected to remain, as hereditarians insist, in *all* environmental conditions we are likely to encounter. For while it is open to someone to say "I don't care about *why* X's are, on average, more intelligent than Y's; the *mere fact* that they are is sufficient to justify preferring (all else being equal) to hire X's," this response seems decidedly perverse if there are very simple environmental manipulations we could perform that would raise the IQs of Y's to the same level as X's, thus eliminating the relevance of race to hiring decisions.[8] The case for racial discrimination is, it seems, much stronger if racial differences are "rooted in the genes" than if they are rooted in the environmental conditions that we, as a society, create.

We'll take a look at some particular kinds of discrimination that might be defended on hereditarian grounds momentarily, but first we should sketch the general logic of pro-discrimination arguments. The sorts of judgments being considered here (about applicants

[7] An example is Sesardic (2005), who defends discrimination on Bayesian grounds in a book ostensibly about the epistemic value of heritability measures for assessing group differences.

[8] Similarly, notice that if this indifference-to-causes argument is taken far enough, all that would be required for Group X to be justified in systematically discriminating against Group Y would be for X to subject Y to IQ-stifling environments for a generation or two. X's of subsequent generations could them claim innocence in discriminating against Y's, on the grounds that Y's are on average less competent, and that the reasons for the disparity are irrelevant.

for housing and employment, and about crime suspects) are almost always made under conditions of limited time and information. Those who must make these decisions, it is argued, are therefore entitled (perhaps rationally required) to use all relevant information at their disposal, *including* (but not exclusively) the race of the person(s) being judged or evaluated. The race of the target can be incorporated as a source of evidence by, for example, using Bayesian principles. If Blacks are on average less competent and productive employees than Whites, then the probability that a particular applicant will be a competent and productive employee *given that* the applicant is White is higher than the probability that a Black applicant will be a comparably suitable employee.[9] Along these lines, supporters of discrimination suggest that race can be a useful item in the evidentiary toolbox of those with a compelling interest in making an accurate judgment about the likelihood that an individual has a given characteristic, given limited time and information.[10]

Note that the consequences of adopting hereditarian-friendly policies on discrimination are similar to, but distinct from, those of adopting hereditarian thinking on differential representation considered above. The hereditarian arguments for a hands-off

[9] An obvious complexity here is that assigning conditional probabilities (e.g., probability of high IQ *given that* someone belongs to racial group R) brings us face-to-face with the *reference class problem* (Hájek, 2007): each person we encounter will belong to a dizzying number of classes (e.g., the class of American citizens, the class of people under 5'5", the class of left-handed trombone players, and so on), each of which has its own probability distribution for the trait one wants to predict. One will arrive at very different judgments about the probability that the target exhibits some quality depending on which reference class is used. So one might reasonably challenge the pro-discrimination theorist to provide a rationale for why *race* is the appropriate grounding for a Bayesian inference. This is an important issue, but for the sake of argument I will assume that those arguing in favor of discrimination will be able to meet this challenge.

[10] This line of thought further assumes that race has some predictive power *over and above* whatever other information is available to the decision-maker. This will, of course, vary by the circumstances, specifically by what information is available—the more information available to the decision-maker, the less likely that race will provide any additional predictive power beyond what is already known.

approach to differential representation discussed above do not assume that employers, admissions officers, or law-enforcement officers practice any sort of discrimination; rather, differential racial representation *should* emerge naturally because, even if employers (for example) completely disregard the races of their applicants, fewer Black than White applicants will present the attributes employers consider valuable (e.g., they will have lower levels of academic achievement, less impressive references, and so forth). The present arguments for discrimination, on the other hand, argue that *in addition* to looking at academic achievement, references, and so forth (which already will segregate racial groups), employers are rationally (and morally) justified in using race to inform their judgments about candidates for employment and promotion (with similar licenses for landlords, police officers, judges, juries, and so on).

First, let's consider employment. Again, supporters of discrimination do not argue that it is legitimate for decision-makers to consider race *exclusively*, so these arguments would not support a policy of refusing to consider hiring Blacks.[11] Nevertheless, the Bayesian principles outlined above could be used to justify a policy of hiring Black applicants only when they are significantly more qualified than White applicants (in effect,

[11] Probably, anyway. Levin (who refers to groups with different average levels of socially-desirable traits as "dominant" and "subordinate") argues: "Unquestionably, there are forces that make the rise of exceptional individuals [from subordinate groups] difficult, such as widespread beliefs about the natural subordination of B's that discourage able B's. Perhaps the defining ambition of modern liberal egalitarianism is to ease informal social pressures of this sort. *On the other hand, the cost of identifying exceptions might exceed the payoff when the overlap between dominant and subordinate groups is minute*" (1997a, n. 256, emphasis added). According to Levin, then, an organization might be justified in excluding Black applicants entirely if the overlap between Whites and Blacks on the relevant characteristics is judged to be "minute." Although Levin does not specify what counts as "minute" overlap, it seems highly doubtful that the Black-White divide on any relevant characteristic would turn out to be large enough to justify this level of disparate treatment (let us not forget—as Levin occasionally seems to—that Black and White people are in fact members of the same species).

requiring a higher standard of evidence that a Black applicant will be a competent and productive employee than that required for White applicants).[12] Against the charge that preferring White applicants to similarly-qualified Black applicants is unfair to Blacks, firms could argue that these practices are based not on unjust or irrational racial prejudice, but rather on cold actuarial calculations of expected future productivity, which firms have a legitimate interest in maximizing.

More generally, notice that since we are speaking of the obligations of individuals and corporate bodies in the public sphere, there would need to be pretty significant reasons counting against this sort of discrimination in order to prevent firms from practicing it. It is unlikely, for example, that a ban on discrimination could be derived from general duties of benevolence on the part of companies, since we generally don't expect—much less require—businesses to act benevolently against their own interests.[13] Rather, a prohibition on discrimination would likely have to derive from some general right on the part of members of groups with lower average abilities. Yet it seems doubtful that a general right of the relevant sort exists, viz., a right *not* to have certain kinds of probabilistically-informative, readily-available information about oneself used by people with compelling reasons to make accurate inferences about you given all available information. That is, it is not clear that using skin color as a proxy for other valued traits violates any right among dark-skinned people to have others make informationally-impoverished decisions.

[12] In fact, there is evidence that employers already do rely on just this sort of heuristic (Bertrand & Mullainathan, 2003; Pager, Western, & Bonikowski, 2009).

[13] Again, I am not *endorsing* the view that businesses have no duties of benevolence; rather I am observing that our current legal structures and social norms do not, as a general rule, impose such duties.

Now, should this general line of thinking be accepted, a great deal of further political deliberation would be required to establish precisely under what circumstances and to what extent employers should be allowed to adopt racial preferences in hiring. These collective deliberations would take place against a backdrop of competing value judgments regarding merit, luck, and so forth, and would presumably take the same general shape as other ongoing conversations about what limits are appropriate regarding how businesses may prioritize profitability over equal treatment (e.g., can a clothing store or an airline hire only physically attractive people, if doing so increases their appeal to customers and thereby maximizes their profits?[14]). Additionally, there is room for a variety of positions regarding the extent to which employers are obligated to invest time, money, and effort in seeking other, non-racial information about candidates' likely attributes. However, the relevant point for our purposes is that in a world where hereditarian thinking is comprehensively incorporated into policy-making (as hereditarians insist it should be), these sorts of racial preferences would no longer be off the table.

On the matter of law enforcement and criminal justice, there are three areas in which discrimination might be justified on hereditarian grounds. First, law-enforcement officers could be justified in detaining or arresting more Blacks than Whites (i.e., requiring a lower threshold of evidence of criminal wrongdoing for detaining/arresting Blacks than for detaining/arresting Whites). Second, juries could be justified in more readily convicting Blacks than Whites. Third, judges could be justified in dealing harsher sentences to

[14] For a summary of historical cases and current debate on the use of attractiveness as a criterion in hiring, see McDonald (2010).

convicted Blacks than to Whites convicted of identical crimes. I will assume the reader is familiar with the standard arguments for increased police scrutiny of certain racial groups (i.e., "racial profiling"); arguments that juries ought to be quicker to convict Blacks follow the same logic: just as law enforcement officers ought to assign a higher prior probability to a Black person's being involved in some wrongdoing—and so ought to require a lower threshold of evidence to detain or arrest such individuals—so jurors ought to assume that the probability of a defendant's guilt is higher when the defendant is Black, and thus demand a lower standard of evidence to render a vote to convict.

On the matter of punishment, Levin (1997b) argues that biological differences between Blacks and Whites entail that these groups differ in the sort of punishment (i.e., retributive vs. rehabilitative) that is most appropriate for them and for society. Since Blacks are, on average, less intelligent and less "future-oriented," Levin argues, they are more likely to be deterred from criminal behavior only by swift, harsh punishment. And moreover, given higher degrees of aggressiveness and selfishness among Blacks, many simply will not be capable of being reformed into contributing members of society. Again, the distributions of whatever traits make harsher sentences more appropriate among Blacks and Whites will overlap. But it is a consequence of Levin's strain of hereditarianism that *on average,* administering harsher, less rehabilitation-oriented sentences to Blacks will serve to allocate such sentences more efficiently (i.e., dispense more of them to people for whom they are appropriate and less of them to people for whom they are not appropriate). In other words, a judge charged with administering a sentence could, in keeping with Bayesian principles, reason that the probability that a harsh sentence is appropriate for a

particular defendant, *given that the defendant is Black,* is higher than it would be if the defendant were White. Thus, the hereditarian-cum-Bayesian might argue, given constraints on judges' time and energy, and on the information available in most cases, we are justified in adopting a defeasible presumption that convicted Blacks should be given harsher sentences than Whites convicted of identical crimes. (This presumption might be implemented or formalized in various ways, for example by a default increase of 15% to Black sentences unless the judge documents compelling reasons why the harsher sentence would not serve the usual purpose.)

6.2.4 Allocation of foreign aid

The idea that people of African ancestry are, on average, innately less disposed to high intelligence and other socially-desirable traits could also be used to argue for the lessening of foreign aid to African nations. One such argument is implied in the discussion above about directing resources to improve conditions among African-Americans in the United States: since the inhabitants of African nations are (and always will be) significantly less intelligent than those of European or Northeast Asian descent, resources spent trying to foster developed, post-industrial economies in these countries are likely to be wasted. Philosopher Stephen Kershnar (2000), however, has provided a separate argument that likewise suggests that we ought to deprioritize the rendering of aid to African nations. Kershnar offers a particularly radical take on the implications of hereditarian science for our moral deliberations; specifically, he argues that, since intelligence is an important component of agency and autonomy (i.e., the ability to engage in practical deliberation

about how to achieve one's ends, and to reflect on what ends one ought to have), individuals with lower intelligence are less fully autonomous agents than those with higher intelligence. And since Blacks exhibit, on average, less intelligence than other racial groups, Blacks have lesser "per capita intrinsic moral value" than other groups (2000, p. 205).[15]

In discussing the implications of this thesis, Kershnar argues that, all else being equal, performing acts that benefit beings with *greater* moral value is morally better than performing acts that provide equal benefit to beings with *lesser* moral value; hence, all else being equal, it is morally preferable for a powerful nation to intervene to rescue *European* victims of genocide than to rescue a comparable number of *African* victims. If we cannot intervene in both circumstances, then according to Kershnar, the morally correct choice is to rescue the Europeans; but even if we were to rescue both groups, he says, our rescuing the Europeans would be more morally praiseworthy than our rescuing the Africans.

Needless to say, many (including myself) will find Kershar's argument both absurd and repugnant.[16] But it's worth taking note of his view, as it gives a uniquely full-throated

[15] Levin's defense of the Aristotelian notion that some groups are "naturally subordinate" to others (1997a) offers a view in the same spirit as Kershnar's. While Levin does not, in this article, specifically reference racial groups, given the overall context of the discussion, and the proximity of the paper's publication to that of his book on racial differences (Levin, 1997b), it is natural to read him as intending his arguments to apply to racial groups (see Mosley, 2005; Wilson, 1998). See, in particular, Levin's claim that, "whites are on average better people than blacks" (1997b, p. 10).

[16] On the absurdity point: Kershnar's argument that more intelligent beings possess greater autonomous agency is based almost entirely upon thought experiments involving comparisons between humans and non-human animals, or between humans with normal cognitive abilities and those with serious disorders that inhibit mental functioning. Thus, the notion of "intelligence" he utilizes bears little resemblance to the use of the term in psychometrics (from which he draws his claim that racial groups differ in "intelligence"). Kershnar provides no reason for thinking that autonomy, understood in the way moral philosophy uses the term, is proportional to intelligence in the psychometric sense. On the repugnancy point, it's also worth noting that Kershnar's extrapolation of comparisons between "human beings and pigs" (2000, p. 222) to comparisons between Whites and Blacks is full of troubling implications.

defense of applying hereditarian scientific conclusions broadly and wholeheartedly in our moral and social thinking.

6.2.5 Summary of "the hereditarian society"

To summarize the features of a society that assumes hereditarianism in devising social policy: the hereditarian society is one pervaded by racial inequality, with some racial groups accumulating the majority of political and economic power, and other groups largely absent from the ranks of high-wage, high-prestige, and politically influential professions. This, of course, is not too unlike our current social conditions, but in the hereditarian society, this inequality is widely accepted as an unavoidable fact of life, and attempts to achieve greater proportionality are considered futile and are rarely pursued. Additionally, the psychological behavioral differences among racial groups are widely acknowledged as elements of "common knowledge" and openly discussed. Among the points considered in such discussions is the extent to which it is legitimate (rationally and morally) to incorporate race-based Bayesian inferences into decisions about employment, housing, criminal justice, and so forth, with some employers (for example) openly advocating for a right to discriminate against Blacks.

Importantly, not all hereditarian thinkers defend the practice of Bayesian discrimination, and even among those that do, clearly they must hold that discrimination is permissible only if the weight assigned to race in the decision-making process is in strict proportion to the actual informativeness of race given other available information (i.e., the use of race as a factor must be consistent with normatively correct applications of Bayesian

inference). So in the hereditarian society, although employers, police units, judges, and so forth might all adopt (openly) policies that apply different treatment to Blacks and Whites, some discriminatory practices will still be unjustifiable; so, for example, an employer that automatically adjusted the GPAs of Black applicants downward by a full point (e.g., treating a 3.8 GPA on a Black applicant's résumé as equivalent to a 2.8 on a White applicant's résumé) would be guilty of pernicious discrimination. In short: in the hereditarian society, decision makers might discriminate, but they would employ *just the right amount* of discrimination against Blacks.

6.3 Moral grounds for disfavoring hereditarianism

In what follows I will advance a series of moral criticisms against a society founded on hereditarian principles. I begin by offering reasons for skepticism regarding two empirical assumptions about human behavior that are implicit in the hereditarian model of social policy. Casting doubt on these two premises will establish that there are good reasons to be skeptical that the hereditarian society, as hereditarians describe it, can be realized. These two assumptions are:

1. Members of society can be trusted *not* to engage in systematic mistreatment of certain groups once the society has adopted these groups' permanent inferiority as a point of "common knowledge" (and has perhaps, in addition, been granted permission to discriminate on racial grounds).

2. Minority groups will exhibit the same levels of achievement whether or not hereditarian claims are widely publicized and are openly accepted and discussed.

If these premises are false, as I argue, then we can predict that attempts to actualize the hereditarian society would likely result in a society rife with racial injustice.

However, I think that even if we were to grant these empirical premises—i.e., if we allow that we could in fact structure society precisely as the hereditarian proposes—the resulting society would still exhibit significant moral defects. I argue, specifically, that racial inequality is, in and of itself, a morally bad-making feature of a society. While this claim might seem obvious, as we will see in Section 6.3.2) below, it is a premise that hereditarians reject (if often only implicitly). I therefore show that a society characterized by racial stratification is harmful to all members of disparaged, inferiorized groups, even those who enjoy relative success in such a society.

A preview of the key steps in the argument is shown in Figure 6.1 below. I aim to show that one could maintain that the hereditarian society is *not* morally faulty only by stringing together a long series of questionable claims (i.e., the hereditarian answers must follow the course plotted by the small blue "no" boxes). Should the hereditarian be wrong about any of these questions, we will have established that accepting hereditarianism comes with grave moral costs, which we should avoid if possible.

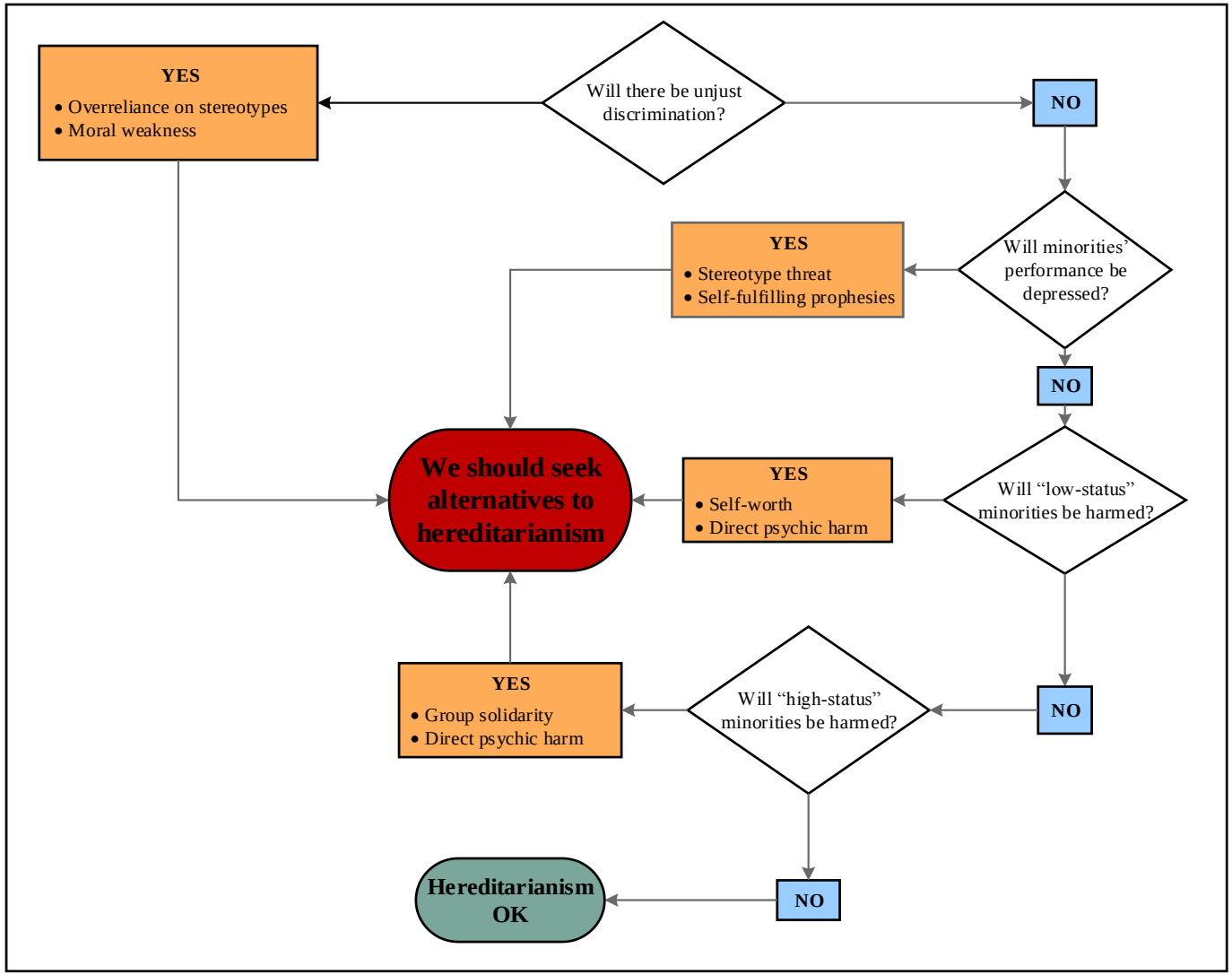

Figure 6.1: Moral questions about the hereditarian Society

6.3.1 The feasibility of the hereditarian society

6.3.1.1 On discriminating "just the right amount"

As noted earlier on, hereditarians are not necessarily univocal in supporting the sorts of discriminatory practices described in Section 6.2.3, and there is room for a wide range of opinions regarding whether and to what extent such discrimination might be justified. But we can divide these positions, broadly, into two types: one might maintain that no such discrimination is ever justified—i.e., that any beliefs about average group differences must be entirely discounted when making decisions about individuals—or one might advocate some degree of Bayesian racial discrimination, so long as the privilege is not abused. Let us begin by considering the suggestion that while hereditarian theories about the existence and causes of racial differences should be widely disseminated, decision-makers should ignore average group differences when making judgments about the character and capacities of individuals. Is it reasonable to suppose that people will refrain from such practices?

Both Herrnstein and Murray (1996) and Jensen (Miele, 2002; Rushton & Jensen, 2005) insist that no member of any racial or ethnic group need fear the widespread publicization of claims about racial differences, because these claims need not (and should not) have any bearing whatsoever on how any individual is treated publicly or privately. (That is, alleged racial differences in intelligence are relevant only to matters of broad public policy, and no one ought to make any judgment about the intelligence or any other

trait of an individual based on that person's race.) So, in this version of the hereditarian society, collective awareness among policymakers of innate racial IQ differences will lead to the withdrawal of social funding for compensatory education and the abolition of affirmative action, but employers, landlords, teachers, law enforcement officials, and the person on the street will all recognize that racial differences are statistical averages only, and will completely disregard racial information in forming and acting on judgments about individuals' likely characteristics.

This hereditarian suggestion—that there is no need to fear that widespread dissemination of hereditarian claims will affect individual interactions—has elicited such responses as "what world do they live in?" (Genovese, 1995, p. 46) and "where have these guys been?" (Judis, 1994, p. 18). I share these authors' incredulity. Still, some unpacking of this hereditarian position is worth pursuing. It is somewhat unclear whether Herrnstein, Murray, and Jensen really believe there is no *epistemic* justification for applying population-level information about racial differences to individual judgments—their emphasis on the fact that racial difference claims reflect only averages and that individuals of all IQ levels and personality types will be found among all groups seems to suggest as much. However, given these thinkers' demonstrated statistical sophistication, it is difficult to imagine that they could have overlooked the applicability of Bayesian inference principles to at least some such situations (something their fellow hereditarians have emphasized, as we saw in Section 6.2.3 above). It seems, therefore, that the prohibition against using Bayesian principles to make racially-informed judgments about individuals' characteristics is intended to be a *normative* constraint—that is, even if there is some

epistemic justification for letting population averages influence our judgments about individuals, we ought not to do so.

If this is right, we can ask two questions about this normative stricture: is it justified by more general normative principles, and is it psychologically plausible to suppose that people can or will adhere to it? Some consideration of the first question was provided above, where I suggested that, while many would see such a principle as a matter of basic fairness, the opposing view cannot be dismissed out of hand (again, we generally recognize that people and institutions have a legitimate claim to using all information at their disposal to make important decisions). But what about the second question: is it realistic to expect that individual agents and institutions will recognize the unfairness of using race to inform assessments of individual characteristics and refrain from doing so, even when it is against their own interests?

I think the answer is a resounding "no." One source of doubt regarding the feasibility of this proposal concerns the way humans apply beliefs about group characteristics to judgments about individuals. A number of studies have demonstrated that when people try to apply information about group base rates to individual judgments, they tend to overweight group-based information, *especially* when it comes to stereotypes about women and non-Whites (Nelson, Biernat, & Manis, 1990; Pager & Karafin, 2009; Quillian & Pager, 2001; Uhlmann, Brescoll, & Machery, 2010). There is considerable psychological evidence that most people already endorse—at least at an unconscious level—negative racial stereotypes (Nosek et al., 2007; Phelps et al., 2000). And moreover, thinking of group differences as reflecting innate or genetic differences tends to increase stereotyping

behavior even more (Kvaale, Gottdiener, & Haslam, 2013; Levy & Dweck, 1999), suggesting that widespread dissemination of hereditarian claims would likely exacerbate tendencies to over-rely on stereotype information. Thus, there is good reason to think that in the hereditarian society, Blacks would be subject to a good deal of unjustified discrimination at the hands of employers, landlords, and police officers, who, despite having no intentions to discriminate, over-rely on stereotype information in everyday decision making.

Rushton and Jensen downplay this worry, suggesting that the solution to the problem of pernicious stereotyping is as simple as giving people instruction in basic statistics:

> thus, any part of a general program of education must include distributional statistics so that people also learn not to stereotype or *over*generalize. This may not be as difficult a task as might be supposed. Even kindergarten children are capable of learning that although boys are typically taller than girls, many girls are taller than the average boy. (Rushton & Jensen, 2005, pp. 283–284)

However, *pace* Rushton and Jensen, the fact that people—even young children—are able to recognize the difference between distributional and universal generalizations, and can report, when asked and upon reflection, the normatively correct answer to a question like "are all girls shorter than all boys?" hardly establishes that they will be able to avoid the sort of reflexive overreliance on stereotypes recorded in the literature.

A second reason for skepticism is that we can hardly expect that everyone will sign on to a prohibition against using statistical generalizations to inform judgments about individuals, or to adhere to it even if they did. There will, in other words, surely be those

who agree with Levin and Sesardic that "reasonable" racial discrimination is entirely kosher, and there is little we could do to stop such people from enacting such discrimination in their judgments and decisions. And even among those who *did* assent to the moral impropriety of ("rational") racial discrimination, we can expect that some "weakness of will" would arise: when the choice comes between upholding this moral principle of fairness and a perceived opportunity to boost profit margins, surely the profit margins will win out some of the time. Even people convinced by (certain) hereditarians' moral arguments against discrimination would on some occasions apply their beliefs about group characteristics to their assessments of individuals (and this is all assuming, of course, that human beings can—*contra* my arguments above—be counted on to employ normatively correct statistical inferences when dealing with socially salient categories such as race).

Thus, in order for the widespread adoption of hereditarian beliefs *not* to result in (arguably legitimate) racial discrimination, we would have to rely on people to be either, to use historian Eugene Genovese's phrase, "saints or idiots." Genovese's response to Herrnstein and Murray's soothing pronouncements of individualism (referenced above) is worth quoting in full:

> Each person, [Herrnstein and Murray] solemnly aver, should be taken as an individual and treated accordingly. What world do they live in? Do they seriously believe that any such sermon would, could, or should dictate the policy of employers with bills to pay, payrolls to meet, and profits to make? May I suggest that employers would have to be either saints or idiots not to be influenced by the collective statistics in choosing between competing individuals? (Genovese, 1995, p. 46)

I conclude, then, that once widespread acceptance of the sorts of irremediable racial differences endorsed by hereditarians takes hold in the public mindset, systematic discrimination is almost sure to follow.

Again, however, not everyone thinks this would be a bad thing. As we saw above, many hereditarians (e.g., Levin, 1997b; Sesardic, 2005) argue that information about average racial differences *should* inform individual decisions (or at least, that people are free to use such information if they so choose). But this position faces the same problems as the less radical proposal of Herrnstein and Murray and Jensen. For again, even the advocate of "rational" racial discrimination must uphold that such discrimination is only legitimate if it is practiced in proper proportion to extant racial differences and the availability of other information. But the same psychological effects described above would still be operative—that is, even if people are licensed to discriminate to some degree, there is good reason to doubt that their decisions will assign only the proper weight to information about the target's race. So it would seem that the vision the hereditarian gives us—of a society in which everyone believes it has been scientifically established that Blacks are, on average, intellectually inferior, but in which statistical clear-headedness prevents these beliefs from spilling over into unjust discrimination—is a pipe dream.

6.3.1.2 Direct effects on minority achievement

In contrast to the matter of discrimination, hereditarians have typically not addressed another likely outcome of promoting widespread discussion and acceptance of permanent, genetically-based racial differences, namely the likely effects on achievement

among the putatively inferior groups. One source of concern was discussed in Chapter 5: stereotype threat. Again, findings in this literature establish that not only do negative stereotypes inhibit performance in high-stakes situations, but the effects are greater the more salient the stereotype. As hereditarians frequently observe, there are currently powerful social mores against openly airing claims about racial differences in psychological characteristics; however, they would have us do away with such restraints and give the possibility of innate racial differences a much more central role in public discourse. Adopting this recommendation can only increase the salience of negative stereotypes to members of stigmatized groups and exacerbate the effects of stereotype threat.[1]

Similarly, there is the worry that pervasive discussion of the alleged inferiority of certain groups will result in lower achievement by way of "self-fulfilling prophesies," whereby members of stigmatized groups respond to subtle behavioral cues signaling low expectations by performing to meet those expectations. For example, teachers might harbor lower expectations for minority students, and these expectations cause the teacher to offer less enthusiastic praise for minority students, or to express greater surprise when they perform well. Students will then read these cues and moderate their behavior to meet the lower expectations, e.g., by exerting lower effort (Rosenthal & Jacobson, 1968; Weinstein, Gregory, & Strambler, 2004). Decades of research on this so-called "Pygmalion Effect"

[1] To be clear, I am not here endorsing censorship—formal or informal—of scientific discussions of group differences; even setting aside the common hereditarian refrain that the pursuit of truth is an absolute good—which I find questionable—there are many good reasons to be skeptical of such censorship (Loury, 1994). Rather, I am highlighting a social and moral cost of adopting hereditarian recommendations, one that hereditarian thinkers themselves have generally neglected to discuss.

have yielded conflicting results (Jussim & Harber, 2005), so I do not wish to lean too heavily on this concern. But, as with stereotype threat, a greater emphasis on racial differences in public discourse could only worsen any such effects as may exist.

This poses a serious challenge for the hereditarian. For a key premise in the case for the hereditarian society is that racial inequality need not trouble us, because this inequality simply reflects a process of social goods being distributed according to meritocratic principles—through no one's fault, Blacks have ended up with lower IQs than Whites, and so there is nothing wrong with Whites having more than Blacks. But if merely by promulgating claims about racial differences, we cause Blacks' achievement to be *lower than it otherwise would be*, then this justification crumbles. Now a particular set of social conditions—of our own making—is causing real and demonstrable harm to those whose achievement is artificially depressed by prevailing attitudes.

Again, hereditarians show a rather surprising indifference towards this point, a point that would seem to establish a significant shortcoming of their theory. The reason, I suspect, is that hereditarians tend to put very little stock in these sorts of effects on either IQ tests or academic achievement more broadly—that is, hereditarians hold to the view that IQ tests are equally valid measures of intelligence across all groups and that IQ test performance is equally predictive of academic and professional achievement across all groups. So the hereditarian might here accuse me of begging an important empirical question about the determinants of achievement. In response I would note that *even if* stereotype threat and self-fulfilling prophesies play little role in generating the *current* achievement gap, it might still be the case that these effects *would* start to start to exert a

significant influence on Black achievement if, as the hereditarian recommends, we were to make claims about racial inferiority a much more prominent feature of public dialogue.

But nevertheless, I want to steer clear of any potential question-begging, so for the remainder of this chapter I will assume for the sake of argument that the hereditarian is correct that disseminating hereditarian scientific claims would exert no inhibiting influence on minority groups' IQ or achievement. Similar considerations apply in the case of discrimination. As argued above, I think the evidence suggests it is psychologically implausible to suppose that hereditarian science can be widely promulgated without a good deal of unjust discrimination following as a result. But hereditarians can object that they have explicitly disavowed the (morally or epistemically) improper use of statistical information about groups in judgments and decisions involving individuals, and that misuse of a theory by misguided people should not count as a strike against the theory. And moreover, for all I have said, it is by no means *certain* that there is nothing we could do (e.g., better statistical training, as Rushton and Jensen recommend) to prevent people from committing these errors. So I will allow, for the time being, that we can follow the hereditarians' suggestion for more open discussion and use of hereditarian science without paying the moral costs of unjust discrimination.

6.3.2 What is wrong with racial inequality, per se?

Taking these assumptions regarding discrimination and achievement for granted, we are now envisioning a society in which "innate" racial differences are routinely and openly discussed, but no one's academic or professional performance suffers for it. And,

while perhaps our society sanctions some degree of discrimination in places of employment, apartment complexes, and courtrooms, no one is subjected to *unjust* discrimination in these contexts.[2] And so, while this society still exhibits substantial racial stratification (reflecting the IQ distribution—with Whites and Asians at the top and Blacks at the bottom), one might nevertheless look at this society and reach the hereditarian-friendly conclusion that no one has any particular cause for complaint in the way social goods are distributed. This section sets out the challenge presented by this line of thinking, while the two that follow respond to it, showing that indeed, this society's racial stratification is a significant moral cost.

How might the hereditarian defend the view that racial inequality *per se* is not morally objectionable? The argument—generally more implicit than explicit in hereditarian writings—goes like this. Recall that I have conceded (in Chapter 5) that *individual* differences in socially-relevant behavioral traits (including intelligence) are indeed attributable to individual genetic differences. Another way of saying this is simply that under any environmental conditions some individual differences in these traits will remain. And recall that I also conceded that in the aggregate, differences in intelligence will yield differential socioeconomic outcomes. It follows from these two concessions that some sort of genetic stratification is almost certainly unavoidable—greater wealth and status will, on average, accrue to people with genotypes that (in a wide range of

[2] I.e., the hereditarian might liken the discrimination that Blacks would receive to selecting a basketball team in gym class: there is of course no guarantee that taller students will prove greater assets on the court than shorter ones, but it would be absurd for shorter students to claim they had been *wronged* by a team captain who decided to "play the percentages" by selecting taller players first.

environmental conditions) are conducive to developing high intelligence and other valued traits. But now the hereditarian can raise the question: if we accept that a stratified society is inevitable, why suppose there is anything particularly wrong with a society in which the stratification happens to track *race*?

A particularly forthright assertion of this very point comes from historian and conservative commentator John Rosenberg. In objecting to an argument for affirmative action from legal theorist Owen Fiss, Rosenberg observes,

> asserting that what's wrong with racial inequality is that it results in "the racial ordering of American society" doesn't really answer the question, for it doesn't explain what exactly is bad about a racial ordering of society...Let me put it this way: since any society that has not been leveled by centralized totalitarian authority must be "ordered" in some way or other—inherited status, wealth, talent, intelligence, where your parents when to school, etc.—liberals need to explain why exactly racial ordering is worse than other possible ways. (J. Rosenberg, 2011)

Rosenberg's challenge makes explicit a premise that I believe to be implicit in much hereditarian writing about racial inequality: *racial inequality* per se *is not morally problematic*. To feel the pull of this position, we can imagine two societies, one of which (R) exhibits racial stratification, while the other (S) shows exactly the same degree of social stratification, but all races are represented proportionally at all levels of the socioeconomic hierarchy. We can even assume that each society has an identical set of "slots" corresponding to particular socioeconomic roles—an occupation, a net worth, and so on—such that across the two societies, wealth and political power are distributed among the society's individuals in exactly the same way; the only difference lies in whether people of all races are randomly distributed across socioeconomic niches, or whether members of

some groups are disproportionally concentrated in high- or low-status roles. Now, if one wants to say that R is morally worse than S, one would have to affirm that we could accomplish a moral improvement by moving from R to S. But Rosenberg's challenge is to say why this should be so: in S there is no less poverty or wealth disparity than there is in R. If we were to move from R to S, then certainly, some Blacks would be better off, but given the rules of the thought experiment, some *Whites* would be *worse* off (and by a symmetrical amount). How can this be a moral upgrade?

In the face this challenge, unless the racial egalitarian can identify some inherent moral defect of a racially stratified society—as compared to a society that is similarly stratified, but not along color lines—then the pursuit of IQ-equalizing interventions would seem to be entirely without justification. For on the assumption that pursuing such interventions will require some sort of investment of public resources—research funds, more money for teacher salaries, and so forth—the hereditarian can question why we should make such investments. If we can identify no moral benefit to moving from a *racially*-stratified society to a *non-racially*-stratified (but still stratified) society, then it seems there is nothing to justify the cost of doing so. As I argue below, however, there are in fact significant moral drawbacks to this kind of race-based stratification.

*6.3.3 The moral costs of racial inequality (*per se*)*[3]

The first way in which I think the hereditarian society is morally deficient is in—for lack of a better term—the sheer *psychic harm* that Blacks would suffer (indeed, already do suffer) as a result of being bombarded with constant reminders of the alleged—but universally accepted—inferiority of one's people. Simply stated, it *hurts* to be told that one is a member of an inferior group.[4] So even granting the (dubious) hereditarian premise that no decrements in achievement for minority groups will result from incessant public repetitions of the hereditarian's inferiorizing mantra, this does not mean that there are no moral costs involved. And these costs are especially troublesome when it comes to those who are already struggling economically: surely it would be next to impossible for a Black person experiencing a run of bad luck not to begin to wonder if perhaps it *isn't* a run of bad luck—if instead he really *isn't* as intelligent as his White neighbors, and is therefore ultimately the source of his own misfortune. The fact that following the hereditarian's recommendations would lead to a society rife with this sort of occurrence—where the insult of publically-expressed doubt about one's intellectual adequacy is added to the injury of economic hardship—must be a serious moral blight on the hereditarian society.

---

[3] The discussion in this section, more so that any other, is framed in a way that is quite particular to the circumstances of African-Americans and the notion of "Black identity." As such I cannot be certain that the framework presented here will be applicable to cases involving hereditarian claims about other racial/ethnic groups. To take one illustrative example, scholars have noted that it is unclear whether the notion of "Hispanic identity"—a group hereditarians maintain has higher "innate IQ" than Blacks, but lower than Whites—can be given a coherent theoretical interpretation, at least under current political circumstances (Idler-Acosta, 2005; Tienda & Mitchell, 2006)

[4] The idea I have in mind is similar to Boxill's (1992, p. 82) "dignitary harm," or "the sense of wounded dignity, inferiority, and stigma experienced by the victims" of public challenges to racial equality (Boxill's example is school segregation).

The second major moral cost to the hereditarian society is the frustration and distress that Blacks experience as a result of the group's continued status near the bottom of the socioeconomic hierarchy. Whatever any individual Black person's social or economic status, it is, I contend, a source of injury to look out upon the world and its extant social conditions and to be continually confronted with reminders of the subordinate status that one's group continues to endure.

Given these quick sketches of the moral costs I see with even the "well-executed" hereditarian society, we can anticipate a number of hereditarian responses. Considering these objections will help elucidate further why I think these moral costs ought to be given serious consideration. First, regarding the psychic harm that results from persistent public repetition of claims about average genetic inferiority, one might claim that while this news might be unpleasant to hear, it is nevertheless the truth, and we must speak the truth even when it hurts. That some religious fundamentalists might find it unpleasant to hear that humans share a common evolutionary ancestor with apes is no reason to refrain from saying it. Relatedly, the hereditarian might suggest that the harm inflicted—on Blacks in particular—by *ignoring* average group differences would be greater than the harms inflicted by speaking candidly about them (cf. Gottfredson, 1988, p. 312).[5]

The hereditarian might further allege that for individual Blacks to be in any way hurt or offended by hereditarian claims is an *irrational* response. That is, one might argue that an individual Black person who feels any kind of insult in response to hereditarian

[5] Some examples of harms that might result from failing to acknowledge group differences—if they existed—were given in Section 6.1.1.

claims about average group differences commits a category mistake: hereditarian claims are not *about* any individual's intellectual capacities; rather, they are about group averages and nothing more. Whatever importance people are inclined to attach to IQ in determining a self-conception or a sense of self-worth,[6] only an individual's measured IQ (and not the mean IQ of any group to which she might belong) ought to be relevant. If someone believes that an IQ of 105 is something to be proud about and an IQ of 85 is something to be ashamed of, then it will of course be of interest to her what *her* IQ is. But once she has her own measured IQ score (or some other reliable indicator of her intelligence) in hand, incoming information about the average IQ of various racial/ethnic groups should have no more bearing on her self-image than information about the average height of the trees on her block. Therefore, the hereditarian might conclude, any psychic harm Blacks may experience as a result of statements about differing group means should *not* be counted as a genuine moral cost, because it is simply not a reasonable response to the hereditarians' scientific claims.

Finally, in response to the claim that in a racially stratified society, Blacks are harmed by constant reminders that their fellow group members are vastly overrepresented among the poorest and least valued members of society, the hereditarian might put forth a version of Rosenberg's challenge: why, the hereditarian asks, should anyone—even members of less successful groups—feel distress over the *mere fact* of stratification when everyone is (by hypothesis) being treated fairly? Is there any justifiable reason why Blacks

[6] Note that the hereditarian need not be committed to any particular stance on the attitude people ought to take towards their own IQ level.

should feel *better* about a society with identical material economic conditions but in which people of all racial groups are randomly distributed across socioeconomic strata?

This will take some explaining, but the problem is easiest to see, I think, if we consider the case of someone—call her Elise—who is what I will call, for lack of a better term, a *high-status* Black member of the hereditarian society I've described. Let us imagine that she is aware that her measured IQ is about 120, and that she has experienced precisely the academic and professional success that psychometricians would predict for someone of her cognitive abilities—she has a stimulating career and comfortable wages, lives in a nice neighborhood, and so on. Why, the hereditarian might ask, should it be of any particular concern to Elise (much less to anyone else) that some miles away in the inner city, the population of "low-status" individuals (those struggling with poverty and related social ills) contains a disproportionate number of Blacks? This is not to attribute utter callousness to the hereditarian, who can of course maintain that it would be commendable or even morally required for Elise to feel compassion for the economically less fortunate, to feel sadness or anger about the conditions they endure, and to act so as to improve those conditions. But why should she be *particularly* bothered by the lower average social status of Blacks, if in fact that group happens to include a greater percentage of low-IQ individuals?

Pressing further, the hereditarian might concede that some 60 years ago, perhaps, Elise would have had every reason to be deeply wary of the racial inequality of the time. For the conditions of *de jure* segregation and open hostility to Blacks that existed at the time ensured a permanently *vulnerable* status for all Black persons, even those, like Elise,

who were otherwise well-equipped to attain conventional success. Even if one did—against the odds—manage to accrue some measure of wealth and/or status, the prevailing racial stigma would have guaranteed the extreme fragility of this arrangement; no amount of intelligence or any other trait could protect one's status against the hostility or even the indifference of Whites. So under these circumstances, someone like Elise—even if she found herself, for a time, enjoying the benefits of her intellectual abilities—would have every reason to be deeply apprehensive of stark racial inequality, insofar as that inequality is indicative of a *shared vulnerability* among all dark-skinned people. Going from *that* sort of racially stratified society to one which was not so stratified would undoubtedly leave Elise better off.

But in the hereditarian society we are considering, things are much different. For here we have *stipulated* that people in this society acquire wealth and status in accordance with the value they bring to the marketplace (largely a function of IQ), and that—in contrast to traditional practices in American society—employers and other decision makers would not spurn intelligent and otherwise desirable candidates merely because of skin color. In other words, in the society the hereditarian would have us adopt, Elise has no reason to doubt that her intellectual capacities will be appropriately recognized and valued, and that she will enjoy all the benefits of being a high-IQ individual in a society that values high-IQ individuals. What then, the hereditarian asks, does the low average IQ of Blacks—and the social inequality it generates—have to do with Elise's well-being?

By way of responding to the hereditarian's objections, it is worth pointing out a consistent theme running through hereditarian commentary on racial differences, one

which helps to illuminate where the hereditarians' moral perspective falls short. In dismissing the potentially harmful social implications of their claims, many hereditarians evince a pronounced *individualism* regarding ethical questions. For example, when pressed by an interviewer about the social implications of his theories—specifically, about the very same (discrimination-free) racial inequality I have called attention to— Jensen responds,

> If there are racial, ethnic, social class, or any other kind of group differences in the proportions of the groups that meet these selection criteria, *so be it*, as long as every applicant, regardless of group membership or background, has been evaluated objectively on his or her own *individual* achievements. (Miele, 2002, p. 179, emphasis added)

Jensen's individualism runs deep: in the same passage he affirms his unwavering ethical commitment to "the primacy of the individual over the group." Similarly, Gottfredson laments the decline of "American individualism" that has accompanied growing concern over racial inequality, and decries the fact that "many people now distinguish between individual and group rights and accord the latter higher standing" (1988, pp. 313–314).

At its core, then, the question of whether the hereditarian society is ethically tolerable is a question about whether groups—in particular *racial* groups—are appropriate elements in our moral deliberations, or whether, instead, we should accept "the primacy of the individual." We have encountered this issue as a sticking point at each stage of the dialectic: for each of my objections, which stress the harm that hereditarian claims cause to Blacks insofar as they think of themselves as members of a group—i.e., *as Black*—the hereditarian responds: *why should that matter?* The question of the ethical status of social *groups* raises a vast array of deep philosophical issues—not least of which are some foggy

metaphysical questions regarding the ontic status of these groups—and I cannot hope to give the matter a comprehensive treatment here. But I will outline the core of my case for the moral relevance of groups, and in particular the moral relevance of the group classified as "Black" in America (and I will provide some further thoughts on this in Chapter 8 as well).

The fundamental problem with the individualism evinced by hereditarians is that it is profoundly *ahistorical*. There is a long tradition of African-American scholarship describing and defending the formation of a "shared racial identity" among American Blacks in particular (Shelby, 2005). This shared identity—whereby the experience of being racialized as Black gives rise to a perceived commonality of interests and shared destiny with others so racialized—is grounded in the peculiar and savage *history* of Blacks' encounters with Europeans and their descendents. The knowledge that one possesses a feature—"Blackness"—that has been the common element in a long history of shared suffering and shared struggle, engenders a special commitment to the well-being of others who share this stigmatizing feature and are thus connected to the same history. This "shared experience of racial oppression and a joint commitment to resist it" forms the basis of what is often called *Black solidarity* (Shelby, 2005, pp. 12–13). At the core of the critique against individualism is a conviction that group solidarity is an appropriate way for Blacks to respond to the history of racial oppression. Thus, to the extent that individualism brushes aside the historical considerations that give rise to group identity, individualism is an impoverished moral perspective.

These themes of shared identity are also prominent in the work of economist Glenn Loury, who has written extensively on the causes and moral significance of racial inequality (Loury, 1999, 2000, 2002). Crucially for our purposes, Loury emphasizes not only the legitimacy of Blacks' adopting a shared racial identity—of seeing one's own well-being as being linked to that of others who share similar experiences of racialization—but also the moral duty incumbent upon *all* members of society to acknowledge the centrality of race to the lives of African-Americans. Hence, caring about racial inequality is not a special obligation incurred by Blacks, but a moral commitment to which we are all beholden.

As Loury argues, our collective obligation to both attend to and combat racial disparities is rooted in the fact that—much as the individualist (hereditarian) might presume otherwise—we do not experience the events of the social world—events such as a walk through a struggling neighborhood, a news report documenting growing income disparity between racial groups, or a blog post touting the latest scientific proof that Blacks really are genetically inferior—merely as individuals. Rather, the moral resonance of such events is necessarily informed and shaped by one's experiences and, crucially, the sense of identity—including racial identity—to which those experiences have given rise. As Loury puts it:

> a liberal individualist of any sophistication ought to reject that brittle, two-step liberalism that enshrines some mythical "unencumbered self"—a "self" located outside the flow of history and the web of culture—as a touchstone of moral judgment in regard to questions of racial pride, kinship, and fealty. (Loury, 2002, p. 97)

Now, the chief argument *against* conferring moral significance on the status of social groups—beyond the claim that doing so is inherently unfair—is that such a practice is fundamentally *divisive*. To consider race as a factor in questions of social justice, one might argue, can only further solidify the social barriers that exist between races, and can only encourage greater mistrust and antagonism by exerting constant pressure for people to frame social issues in terms of group interests. We should, in other words, actively *discourage* people from thinking of themselves in racial terms, encouraging instead a "colorblind" mentality whereby citizens have no inclination to see themselves as part of a particular racial group.[7] If, the hereditarian-individualist might suggest, people were to abandon their notions of racial identity, then the moral issues I have raised for the hereditarian society (i.e., Blacks feeling slighted by claims about group mean IQs or feeling distressed by social disparities that happen to fall along racial lines) simply drop out of the equation. These problems, the hereditarian might argue, arise *only* because of a tendency (i.e., to include race as part of one's self-conception) that the hereditarian actively discourages in the first place.

Again, I believe this perspective fails to take account of the historical realities that are at the root of the present significance of race and racial identity to the lives of many. The tendency among historically oppressed groups to see oneself in racial terms is a deeply ingrained feature that cannot be wished away, as Loury emphasizes:

[7] Of course, this kind of colorblind individualism is at odds with some of the more extreme hereditarian voices—e.g., Kershnar (2000) and Levin (1997a, 1997b)—who argue for the moral significance of race but, on the question of where our moral commitments should lie, reach a conclusion directly opposite the one defended by Shelby, Loury, and myself.

> The implicit assumption of advocates of race-blindness is that, if we would just stop putting people into these boxes, they would oblige us by not thinking of themselves in these terms. But this assumption is patently false. (Loury, 2002, p. 142)

This critique of race-blind individualism is of course consistent with the idea that race is a "socially constructed" category that, had history been different, we might have done without—or might yet do away with. Perhaps at some point in the far future, we will reach a point where the very idea of racial identity *does* drop out of the moral equation. The rejection of colorblind individualism, then, is "not a principled rejection; it's historically contingent. It's a rejection based upon the specific facts of our society and the way in which people see themselves" (Loury, 1999). But owing to the *actual* historical and current conditions in which racial identities are formed, this rejection is necessary for the time being.

6.4 Conclusion

If what I have said in this chapter is correct, then accepting the core hereditarian claim (i.e., HED: genetic differences between races make social equality, for all practical purposes, impossible) would come at great moral cost. Again, however, the question of whether this claim is true is ultimately an empirical one, and however much we might lament its moral consequences, the *truth* of the claim is an independent matter. So the moral considerations I have outlined cannot ground an outright, in-principle rejection of hereditarianism. But these considerations *do* imply that if we have at our disposal a means of resisting the core hereditarian claim—of forestalling the acceptance of HED as a premise

in policy deliberations—we ought to pursue it. This is what I aim to provide in the next chapter: an explanatory framework that, if adopted, may obviate the moral quandaries that come with adopting hereditarianism.

CHAPTER 7: HERITABILITY, MECHANISMS, AND NATURE/NURTURE DEBATES

7.1 Introduction

This chapter explores the contrasts between behavior-genetic and mechanistic approaches to studying the ontogeny of complex human behaviors. The behavior-genetic approach is favored by those researchers—philosophers and scientists—I am calling "hereditarians," who argue that observed racial differences in socially-desirable traits (such as intelligence, altruism, and self-control) are genetic in origin and irremediable. Since I am concerned to resist such claims, this chapter argues for methodological shortcomings in the hereditarians' preferred explanatory strategies, and for the superiority of mechanistic approaches. However, in keeping with the explanatory pluralism defended in Part I, I will not be arguing for the *unqualified* superiority of mechanistic explanation. Indeed, I will argue that behavior-genetic modes of explanation are perfectly adequate for some explanatory purposes. But our pragmatic interests in the context of group IQ differences dictate a preference for mechanistic strategies.

To briefly review the arguments of the previous two chapters: hereditarians take the evidence for claims about "innate" racial differences to be sufficiently decisive that we are rationally required to take such irremediable differences as empirical premises in social policy deliberations. However, an exploration of some of the policy implications that might follow from accepting such claims shows that a society that took such hereditarian

arguments seriously would be morally defective. We therefore have powerful moral reasons to resist the claim that racial differences are irremediable, and to instead seek interventions that will eliminate observed racial gaps in socially-important traits.

Picking up this thread in the argument, the present chapter argues that the analytical methodologies used by hereditarians in pursuing nature-nurture questions are poorly suited to discovering such interventions. Moreover, I demonstrate that an alternative explanatory framework, one with the goal of developing detailed mechanistic models of psychological development, is significantly more likely to yield information that will facilitate the discovery and implementation of interventions that can eliminate observed racial differences. Mechanistic strategies are therefore a promising route to averting the moral calamity of a society characterized by permanent racial inequality.

Since I believe (see Chapter 1) that the adequacy of scientific explanations is context-dependent, and that practical and moral considerations play a role in fixing explanatory contexts, I believe that the superiority of mechanistic explanations of development in guiding intervention design makes them overall better scientific explanations in the contexts we inhabit (or at least, *ought* to inhabit). However, these assumptions are not essential for most of the arguments laid out in this chapter. I am primarily concerned to establish the narrower thesis that differences between behavior-genetic and mechanistic researchers ultimately reflect *pragmatic* rather than *empirical* differences, and that mechanistic analyses are much better suited to the design of interventions (cf. Darden, 2013; Machamer et al., 2000).

Thus, in this chapter I will be concerned with the conceptual issues that divide hereditarians and a particular class of their opponents, those who might be termed *interactionists* given their emphasis on the complex interrelations between genetic and environmental factors in development. Sometimes these theorists have been key players in the race-IQ debates (e.g. Lewontin, 1974), but in many cases these critics have been more centrally concerned with the nature of explanation in developmental science, and have argued for methodological deficiencies within the entire field of behavioral genetics (Gottlieb, 1991, 1995; Wahlsten, 1990; West & King, 1987).[1]

The chapter proceeds as follows: Section 7.2 reviews the basics of behavioral-genetic analysis and discusses two types of developmental complexity (gene-environment correlation and gene-environment interaction) that present *prima facie* challenges to the applicability of behavioral-genetic analysis to racial-difference debates. In Section 7.3 I examine in detail the debates between hereditarians and their opponents in order to draw out the differing explanatory interests of the two sides; in particular I consider the question of how we should delineate the *range of relevant environmental conditions* for development, and I argue that hereditarians have taken an overly simplistic approach to this question. Finally, in Section 7.4 I demonstrate how the mechanistic approach to explaining developmental outcomes affords better opportunities for control and manipulation than the behavioral-genetic alternative.

[1] Given their focus on the complexity of gene-environment interplay in development, the class of "interactionist" thinkers overlaps considerably with the "developmental systems theorists" discussed in Chapter 2.

7.2 Heritability and Behavioral-Genetic Analysis

Recall that a key article of evidence cited by hereditarians arguing for racial differences in genetic propensities for intelligence is the relatively high degree of heritability that measures of intelligence (e.g., IQ tests) exhibit—typically from 0.4 to 0.8 (Herrnstein & Murray, 1996; Jensen, 1969; Rushton & Jensen, 2005). The use of heritability estimates has long been subjected to strident critiques, with many researchers suggesting that most uses of heritability measures are based on fundamental confusions and hence have very little to contribute to our understanding of how phenotypes develop (for classic critiques, see Block & Dworkin, 1976; Block, 1995; Jencks et al., 1972; Lewontin, 1974). One sticking point is whether heritability analyses and ANOVA can provide *causal* information about development, either at the level of the individual organism or at the population level (e.g., Lewontin, 1974; Northcott, 2006; Sesardic, 1993; Sober, 1988). Another point of contention is whether, given a high level of heritability for a trait in two separate populations—indicating that *within* those populations, genotypic variation plays a large role in generating phenotypic variation—it is legitimate to infer that genotypic variation plays a large role in generating phenotypic variation *between* those populations.

For our purposes, these debates don't much matter. We can allow, as hereditarian philosophers like Sesardic (1993, 2005) and Levin (1997b) insist, that heritability does give us information about causes—i.e., that high heritability *within* populations can ground inferences about the causes of *between*-population differences, and that high heritability for a trait allows us to infer an individual's phenotype was caused more by its genotype

than its environment.[2] Nevertheless, I will argue, merely causal characterizations of the role of genes in producing phenotypes are too impoverished to serve the goals we bring to the table in studying individual and group variations in human behavior (Darden, 2013).

7.2.1 Heritability's limitations and the IQ gap

Two features of heritability measures, however, will be especially salient for the ensuing discussion:

1. Heritability can only be estimated for a trait in a *particular population* in a *particular environment*. That is, the heritability of a trait in one population may be different from the heritability of that trait in a genetically-identical population in a different environment.

2. Heritability measures are often insensitive to the presence of *genotype-environment correlation* (rGE) and *genotype-environment interaction* (G×E). These concepts are explained below.

These two limitations help us to identify two ways in which the "core hereditarian claim" (HED: genetic racial differences will cause unequal racial outcomes in all practically possible human environments) might be false. One such possibility (call this Hypothesis 1) follows from the first limitation of heritability mentioned above: present phenotypic differences between Black and White populations stem purely from *environmental differences* between the populations—that is, Black and White populations are genetically identical with respect to those genes that are relevant to intelligence, and if environmental conditions for Black and White populations were the same, these populations would exhibit

[2] While I am willing to allow such claims for the sake of the present argument, recall that I do offer arguments against them in Chapter 5.

identical mean phenotypic intelligence. But, Hypothesis 1 holds, Black and White children do *not* develop in identical environments; rather there are substantial differences in the environmental conditions of Blacks and Whites. Possible environmental differences were reviewed in previous chapters and include racial stigma resulting in stereotype threat, self-fulfilling prophesies about Black underperformance, and so on.

Picking up on the second limitation of heritability, an alternative hypothesis (Hypothesis 2) suggests that there *are* in fact genetic differences between Blacks and Whites with regard to genes affecting intelligence—such that even if White and Black populations developed in *identical* environmental conditions, differences in phenotypic IQ would still be observed—but that HED is still false because there are *some* environments in which mean phenotypic IQ would be equal across both populations. In other words, Hypothesis 2 holds Black and White populations have equal genetic potential for intelligence, but that genetic differences between the populations are such that the optimal environments for IQ development are different across groups. If Hypothesis 2 is correct, then a society could equalize mean IQ scores across racial populations—thereby eliminating what is, by hypothesis, the primary cause of differences in social outcomes between racial groups—by providing more targeted environmental manipulations to different citizens.[3] This possibility arises as a consequence of the phenomenon of *gene-environment interaction*.

[3] This hypothesis suggests, in other words, that there will be some correlation between racial group membership and optimal IQ environment, and the precise practical recommendations will depend on the strength of that correlation (an empirical question). But the basic idea would be that racial group membership would count as defeasible evidence that a particular child will benefit from a particular environment.

I wish to suggest that if we are concerned with eliminating the race-IQ gap—i.e., if we are as committed to racial equality as we ought to be—it doesn't really matter whether Hypothesis 1 or Hypothesis 2 is correct. *Both* of these possibilities are conducive to designing interventions that will raise Black IQ and reduce or eliminate racial inequality, provided that we understand the mechanisms by which people end up with the IQs that they do.[4]

7.2.2 Two complexities in developmental science: rGE and G×E

Objections to the use of heritability in nature-nurture debates (including the race-IQ debate) often stress two phenomena that pose conceptual and methodological challenges to inferences drawn from heritability statistics. The first is gene-environment correlation (rGE), while the second is gene-environment interaction (G×E). Both of these phenomena come in multiple forms, and at least some forms of each are thought to reveal conceptual difficulties with the ways in which methodological tools used in behavioral genetics (i.e., ANOVA or heritability) partition genetic and environmental sources of phenotypic variance, as well as methodological shortcomings that limit the utility of behavioral-genetic analyses. Specifically, *in theory* a properly-conducted heritability analysis should detect phenotypic variation due to rGE and G×E and count them as *additional* sources of variance beyond genetic and environmental sources. Crude applications of the formulae for

[4] A perhaps surprising result is that even if Hypothesis 1 is false—that is, if anti-hereditarian theorists are wrong in thinking that the current Black-White IQ gap owes entirely or even chiefly to such factors as cycles of poverty and stigma among African-Americans—we may still be obligated to invest heavily in social interventions designed to raise Black IQ, for we may well be able to reduce or eliminate racial stratification by conducting targeted interventions even in contexts that would not meet our ordinary notions of "discrimination" or "disadvantage."

calculating heritability, however, will be insensitive to these phenomena, and, what's more, will tend to count rGE and G×E as components of *genetic* variance, thus inflating the influence of genetic factors in generating phenotypic differences.

In terms of our current questions—about population differences in IQ and other socially-relevant traits—the relevant question is whether rGE and G×E play any role in generating the presently observed disparities between populations—or whether they might be expected to do so in environmental conditions we might reasonably bring about. If so, then, as I argue below, our prospects for eliminating observed differences are significantly greater. Hereditarians have typically suggested that rGE and G×E play at most minimal roles in generating the IQ gap (or at any rate, that any such role is of no practical import), but, as I argue, this stance reflects a particular explanatory perspective rather than a straightforward empirical hypothesis.

7.2.2.1 Gene-environment correlation

First, let's discuss gene-environment correlation. Gene-environment correlation (rGE) refers to circumstances in which environmental variables causally influence the development of the phenotype, but in which certain genotypes are more likely than others to encounter certain environmental conditions. The problem rGE poses for heritability is that the statistical tools used to generate heritability estimates will (at least if applied crudely) be blind to the mediating influences of the environment, and simply subsume effects of rGE under the "genetic" component of phenotypic variance.

Since we'll be considering cases where genetic influences on phenotypes—particularly IQ—are said to be more or less "direct," we should try to have a clear sense of what this means. A number of ways in which genes might influence phenotypes, and their respective levels of "directness" will be considered in detail below, but it will help to begin by describing a *limiting case* of direct genetic influence. In this vein, suppose it were the case that there is a linear relationship between intelligence (as measured by psychometric tests) and some well-defined low-level neural property—say the rate of firing of neurons in the cortex—and that the causal pathway between particular sets of genes and particular neuronal firing rates is roughly as "direct" as that between genes and eye colors in humans.[5] For simplicity's sake, we can suppose there is some combination of a handful of alleles shared by all individuals with the neural processing speed associated with IQ 100, and that every polymorphism at each of these sites is reliably associated with a specifiable increase or decrease in processing speed. This would, I take it, constitute a clear case of a maximally "direct" genetic influence on IQ. In what follows, therefore, when I speak of "direct genetic influences on IQ" or "genes that directly influence IQ," the reader should imagine the kind of limiting case just outlined.

The actual prevalence of rGE in biological development is controversial, with theorists disagreeing about whether rGE occurs frequently in nature, and if it does, how it should be incorporated into scientific explanations of phenotypic differences in a population. The matter is complicated by the fact that the general phenomenon of gene-

[5] Recall that, as argued in Chapter 2, so far as we know there are very few, if any, explanatorily relevant extra-genetic causal mediators between genes and eye colors

environment correlation can encompass a wide and diverse variety of developmental processes. Scarr (1992) divides instances of rGE into three types, each of which calls for a separate analysis; Sesardic (2005) helpfully names and explicates these:

(i) ***Passive* rGE**, in which the correlation between genotype and environment is, in terms of local causal factors in development, essentially an accident. For instance, individuals who are possessed of genes that predispose them to develop high IQ may be more likely to live in areas where there are high-quality, resource-rich schools. In that case, children who receive high-IQ genes from their parents will also, statistically speaking, develop in environments more conducive to developing high-IQ. But the environmental and genetic influences are independent causal factors that happen to be statistically correlated.

(ii) ***Reactive* rGE**, wherein individuals with particular genotypes (reliably associated with particular phenotypes) are consistently *treated* in particular ways by others in their environments. The classic (hypothetical) example is from Jencks et al. (1972): suppose that children with red hair are systematically singled out for discriminatory and abusive treatment, such that they consistently develop low IQs.[6] In this case, there is a causal chain from genotypic differences to phenotypic differences (as opposed to the accidental correlation observed in passive rGE), but it seems counterintuitive to the point of absurdity to say that the reliable

[6] A real-life example of reactive rGE is the rat pup case discussed in Chapter 3. Variation in the genes controlling certain hormones is reliably associated with variation in adult sexual behaviors, but this is only because of the way mothers *respond* (i.e., through licking) to the phenotypic traits (the hormones) that are produced by the relevant genes. Interestingly, Sesardic says that in non-human animals "for obvious reasons, the complex interactions characteristic of…scenarios like the red-haired children example do not come into play" (2005, p. 97). Presumably what Sesardic means is that non-humans do not utilize complex social categories in directing different types of treatment at different individuals. But, *pace* Sesardic, there doesn't seem to be any reason not to classify what goes on in the rat pups case as a paradigmatic example of reactive rGE.

association between genotypes and phenotypes here amounts to "a genetic difference causing a phenotypic difference," or that genes that causally produce red hair in some way constitute "genes for low IQ."

(iii) ***Active* rGE**, which occurs when individuals with particular genotypes are more likely to *seek out* or even *create* environmental conditions that influence their phenotypes. The usual example here is a situation in which some children have genes that (in a broad range of environments at least) lead to their being more likely to seek out intellectually challenging environments (reading more books, playing chess rather than pinball, etc.), thus boosting IQ. As with reactive rGE, there is a causal chain from genotype to phenotype, but it is perhaps more plausible to suppose in this case that the influence of gene-environment correlation can be counted as a "genetic" source of phenotypic variance (though causal intuitions in these cases are rather slippery, as will be discussed more fully below).

Despite substantial disagreement on the proper interpretation of rGE as a whole, there is actually some agreement on the conceptual status of types (i) and (ii) (passive rGE and reactive rGE). It is widely agreed by both hereditarians and interactionists that counting the influence of passive rGE and/or reactive rGE as part of the "genetic" component of phenotypic variance is misleading and scientifically illegitimate: while a mechanical computation of heritability might place rGE among the genetic sources of variance, all parties agree that passive and reactive rGE should be counted as distinct sources of variation in addition to purely genetic and purely environmental causes. To the extent that the possible occurrence of these types of gene-environment correlation generates

controversy, it is over the *methodological* challenge they seem to pose for hereditarianism: crude applications of formulas for estimating heritability, without attention to potential mediating environmental factors, will end up collapsing these forms of rGE together with any direct genetic influences into an amorphous chunk of "genetic causes" of IQ variation (i.e., since genes that produce red hair are reliably associated with low IQ, a heritability analysis will essentially treat genes for red hair as "genes for low IQ").

As Sesardic (2005) emphasizes, there are ways of empirically testing for passive and reactive rGE, so in principle it is possible to avoid collapsing such effects into the "genetic" component of phenotypic variance. However, conducting such tests requires that researchers know what sorts of correlations to look for (i.e., what environmental conditions might be accidentally or indirectly associated with particular phenotypic outcomes) and there is no guarantee that the associations will be sufficiently intuitively obvious that researchers will think to look for them. I return to this point in Section 7.3 below, but for now the point is simply that should we discover that passive and/or reactive rGE contribute to the IQ gap, it would be illegitimate to say that the differences are "genetic in origin," and doubly illegitimate to insist that because the differences are genetic, nothing can be done about them.

There is, however, considerable disagreement on the empirical and conceptual status of *active* rGE, and the differences in the interpretation given to active rGE helps to illuminate the importance of pragmatic explanatory contexts to nature-nurture debates. Some behavioral geneticists argue that active rGE should be counted as a source of *genetic* influence on phenotypic outcomes, and this has become the standard line among

hereditarian scientists and philosophers. The argument goes that there is a stark intuitive contrast between passive/reactive rGE and active. Passive and reactive rGE, on the one hand, involve environmental factors over which the organism has little or no control, and on which the organism's genotype has at most only a very indirect causal influence. In cases of active rGE, in contrast, the genotype *causes* the organism to structure its environment in a way that has predictable phenotypic effects, such that these phenotypic outcomes are "more or less inevitable result[s] of genotype" (Jinks & Fulker, 1970, p. 323; cf. Roberts, 1967). Thus, while they acknowledge the formal distinction between active rGE and more "direct" genetic causes, hereditarians insist that *for all practical purposes*, we should consider phenotypic differences arising from active rGE to be "genetically-caused" differences. And from here it is a short step for the hereditarian to dispense with the relevance of active rGE to their arguments: *even if* active rGE plays some role in generating observed racial differences, these differences should still be considered just as fixed and irremediable as if they were due entirely to genetic differences (e.g., Sesardic, 2005, pp. 93–95).[7]

Notice, however, that whether "we should count" active rGE as a component of genetic variation *for all practical purposes* is not an *empirical* question! One's stance on

[7] It is worth pointing out that hereditarians who are inclined to count many cases of active rGE as genetic causes of variation typically assume that genes that cause individuals to influence their environments in ways that affect the development of a trait will always or usually co-occur with genes that influence the trait *directly* and *in the same direction* (i.e., those children who already have genes that directly lead to higher IQ—via faster neural firing speeds or whatever—will *also* have genes that lead them to seek out or create IQ-enhancing environments, and those children with low-IQ-producing genes will seek out/create environments that depress IQ, or at least fail to enhance it; in a phrase: "the rich get richer"). But there is no reason, in principle, why this should always be so—it could well be the case that different genes influence neural processing speeds and tendencies to seek out enriching environments, and that these genes assort completely independently.

this question reflects instead a judgment about the *explanatory relevance* of the sorts of causal factors at work in a case of rGE. To illustrate: suppose it were the case that intellectually stimulating environments in development (more book reading and chess playing, less reality TV and video games) lead to higher IQs across all genotypes, but that, indeed, individuals with a certain genotype (G_1) are (under present environmental conditions at least) more likely to seek out such environments than some other genotype (G_2). So as a matter of fact G_1's exhibit higher IQs than G_2's. Arguing that under such circumstances, the difference ought to be classified as genetic, Jinks and Fulker ask, "to what extent could we ever get a dull person to select for himself an intellectually stimulating environment as a bright person might?" (1970, p. 323). Jinks and Fulker evidently intend the question to be rhetorical, but clearly their question is an *empirical* one. And moreover, it is a question to which some of us would very much like to know the answer: even if, under *present* environmental conditions, these different genotypes manifest their "natural" tendencies towards constructing high- or low-IQ-creating environments, are there environments in which G_2's could be induced to read more books and play more chess?

To be sure, the issue of what we might do to ensure that all genotypes end up in properly stimulating environments admits of considerable nuance—what if, for example, we could only bring this about by imposing draconian measures that force G_2 children to spend hours a day at the chess table despite finding every moment of it miserable? But the point is we cannot appeal to intuition to settle *a priori* the question of whether and how environmental discrepancies that arise as a result of active rGE can be eliminated. And

perhaps more importantly, whether we are disposed to count active rGE as part of the (unmodifiable) genetic variance for a trait is underdetermined by the empirical facts of the case, and reflects in part our interest in *controlling* or *manipulating*, rather than simply *predicting* developmental outcomes. Certainly the ability to predict outcomes is a virtue for a scientific theory and in many circumstances this will be all that is required to meet the needs of the explanatory context. In those cases subsuming rGE under genetic variance may well be acceptable. But it remains a *pragmatic* decision nonetheless.

7.2.2.2 Gene-environment interaction

The term "interaction" is prone to cause confusion here, so it will be worth taking a moment to clarify. There are, broadly speaking, two different notions of *gene-environment interaction* at play (for discussion, see Griffiths & Tabery, 2008; Jensen, 1974; Keller, 2010; Longino, 2013). One is the commonsense idea of *causal* interaction: genetic and environmental factors must interact to create phenotypes, for neither is causally sufficient on its own. The other is the more technical idea of *statistical* interaction, which arises when the impact of particular environmental manipulations on phenotype is different for different genotypes (symbolized as G×E). In these cases, the genetic and environmental contributions to population variance are non-additive, and the presence of this sort of interaction effectively nullifies the informativeness of applying an analysis of variance (ANOVA) to the phenotypic variation in a population.

An example of G×E may be seen in Figure 7.1 below (from Gray, 1992). As we can see, the effects of moving from lower elevations to higher elevations are not uniform

for different genotypes within this species of plant: increases in elevation sometimes result in a taller plant, sometimes shorter.

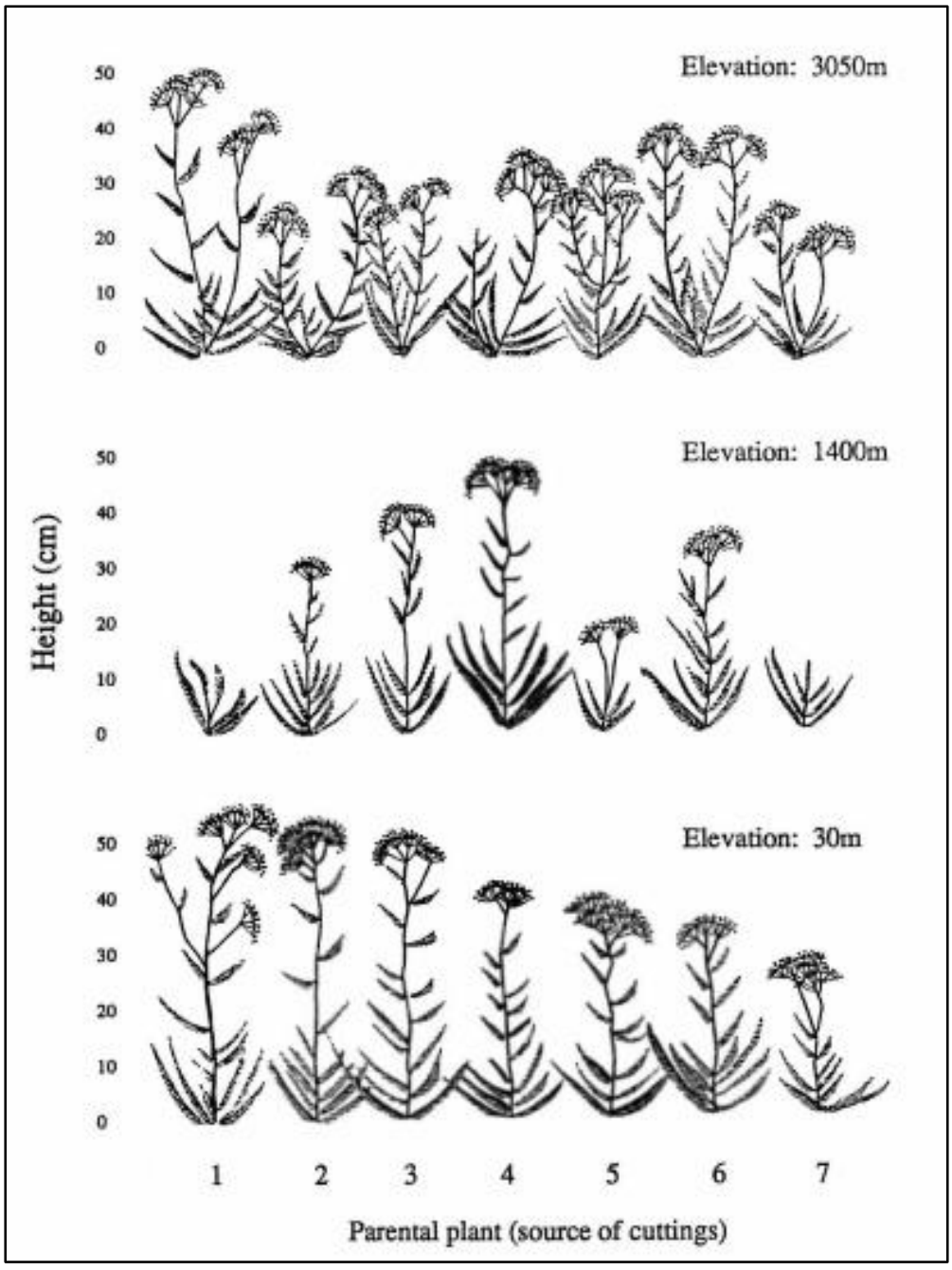

Figure 7.1: GxE in Achillea millefolium (from Gray, 1992)

More formally, the idea of statistical interaction between genes and phenotypes can be understood through the concept of a *norm of reaction*, which is a graph depicting the phenotypic values for one or many genotypes across a range of environmental conditions. The curve for any particular genotype across a range of environments can be linear or non-linear, and the curves of different genotypes may or may not cross, as demonstrated in Figure 7.2 below.

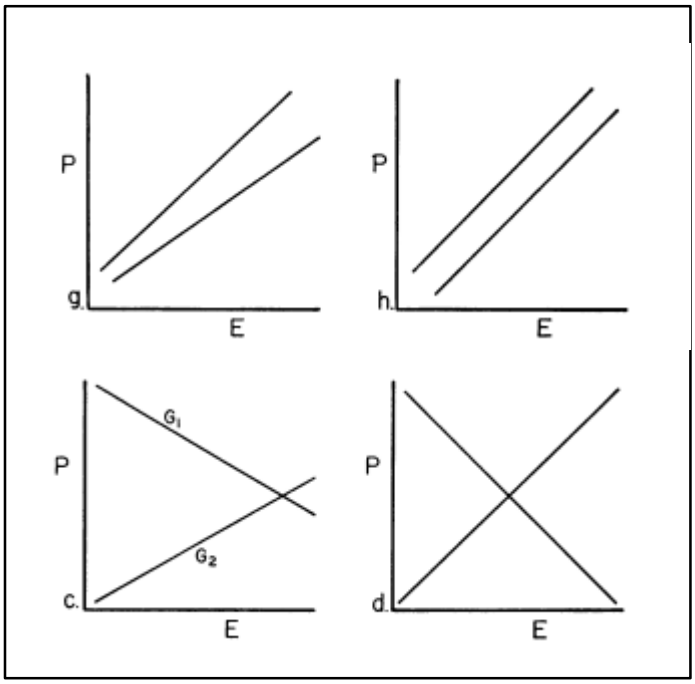

Figure 7.2: Hypothetical reaction norms for a genotype (from Lewontin, 1974)

If the curves of different genotypes do not cross, then it is said that genetic and environmental variances are "additive," and additivity is a precondition for conducting an analysis of variance to partition sources of variance; when the curves *do* cross, we have gene-environment interaction and the additivity condition ANOVA requires is not

satisfied. Intuitively speaking, if norm of reaction curves are non-linear, and if the curves of different genotypes cross, then it is not possible to say that, *in general* a particular environmental manipulation increases or decreases phenotypic values across the board: the answer to the question of whether increasing altitude (say) will increase or decrease height is: *it depends*.

For a real-life example of G×E in human behavioral development, consider the now well-known case of gene-environment interaction in the development of anti-social personality disorder, or ASPD (Caspi et al., 2002). A genetic polymorphism in humans results in differential levels of the enzyme monoamine oxidase A (MAOA), which metabolizes certain neurotransmitters. Thus, there are low-MAOA and high-MAOA individuals in the population.

Under most conditions, the incidence of ASPD among both low-MAOA and high-MAOA individuals is low, with high-MAOA variants being slightly more prone. However, among people who have been victims of maltreatment between ages 3 and 11 (including neglect and physical and sexual abuse), a dramatic asymmetry emerges: among high-MAOA variants, the incidence of ASPD rises slightly, while among low-MAOA variants it increases sharply, to as much as 85 percent among victims of serious abuse. Thus, whether a low- or high-MAOA variant is more likely to develop ASPD depends on environmental conditions. Similar effects have been observed in the effects of maltreatment on the incidence of depression, with polymorphisms for genes that control serotonin transport being differentially affected by mistreatment (Caspi et al., 2003), and

also for the effects of cannabis use on the incidence of schizophrenia, which depends on genes that regulate levels of catecholamines in the brain (Caspi et al., 2005).

It should be noted that for each of these behavioral outcomes, there was a main effect of environment but no main effect of genotype. That is, the effects of the environment, considered independently of the genome, were statistically significant, while the effects of the genotype, considered independently of the environment, were not (see Rutter, 2006). Thus, although environmental measures can, to some degree, be used to predict such outcomes, such measures are significantly impoverished when compared to the predictive power of combining genetic and environmental information. The upshot is that given two individuals with different genotypes—say, a low- and high-MAOA variant—it is impossible to say which is more likely to develop ASPD without considering environmental variables.

Returning to the question of the IQ gap, we can ask: to what extent might G×E play a role in generating the IQ discrepancy between Black and White populations? Hereditarians have correctly pointed out that at present there is little evidence that G×E plays much of a role in generating IQ differences, racial or otherwise. In other words, the norms of reaction for different genotypes do not appear to cross when it comes to IQ. But a clear lesson that emerges from the concept of the reaction norm is that while norms of reaction might *appear* to be parallel within a particular environmental range, extending the range out further—i.e., observing the development of genotypes in new environments—may well reveal the existence of G×E (Gottlieb, 1991, 1995; Griffiths & Tabery, 2008; Lewontin, 1974). That is, we must acknowledge the possibility that although the norms of

reaction for “Black” and “White” genotypes have remained separated in the environments we have observed so far, they might well meet under hitherto unobserved environmental conditions. The key questions, then, are: how broad a range of environments have we observed, and to what extent are we licensed to infer that new environments will exhibit the same patterns we have observed in existing ones? These questions I take up in the next section.

7.3 Environments and Environmental Ranges

Recall from Chapter 5 that my quick-and-dirty characterization of the hereditarian thesis is that currently-observed racial differences will manifest in *the full range of relevant environmental conditions.* I have deferred up to this point two crucial questions raised by this characterization, namely, what is a *range of environments*, and how is it determined which environments are the *relevant* ones? This section addresses these questions, and I will offer two general conclusions. The first conclusion is that a close reading of the debate between hereditarians and their critics reveals differences in the explanatory interests of the two camps. Specifically, hereditarians consider *possible* but *unobserved* environmental conditions to be irrelevant to their explanatory projects, whereas such potential environments are central to the explanatory aims of their critics.

The second conclusion is that hereditarians have misconstrued our present epistemic position, and have therefore been overconfident about what we can infer about what sorts of IQ distributions might be expected under various hitherto-unobserved social conditions.

7.3.1 Preliminaries: the "relevant" environments

It will help to begin with an example of the sort of argument the hereditarian wishes to make about the heritability and malleability of IQ, one less complicated than IQ itself. The hereditarian wishes to claim that intelligence is like height (Levin, 1996, p. 257; Plomin & Bergeman, 1991, p. 415; Rushton, 1997, pp. 47–48). Height is clearly a trait affected by both genes and environment: modern humans are much taller, on average, than humans were just a few generations ago, and people in developed countries, where food is plentiful and all necessary nutrients are readily available, are significantly taller than people in developing countries. So clearly environmental factors play a role in determining height.

However, two related facts about height are notable. The first is that as affluence increases in a region, so does the heritability of height. This should not be surprising: it is a necessary truth that as the degree of environmental variation decreases, the proportion of total variation that is due to genetic variation will *increase* (e.g., in the limiting case, where there is no environmental variation at all, all variation in the population must be due to genetic variation). Since in developed countries there is significantly less variation in the environmental conditions that affect height, more of the phenotypic variation—indeed almost *all* of it—is due to genetic variation (height is about 90% heritable in the developed world).

The second, related point is that it seems quite reasonable to speak of "genetically-determined limits" on an individual's height. It is likely true that many people in impoverished conditions could have been six inches taller if they'd had access to better nutrition, but it is doubtful whether *I* could have been six inches taller with better nutrition,

or better healthcare, or more stimulating toys, or any other environmental manipulation that plausibly (or even not-so-plausibly) makes a difference to height. Given the extremely high heritability of height in our population, it is likely that most of us are about as tall as we could have been in just about any environment—in other words, the conditions of the modern-day, developed world seem to approximate the optimal environment for human height development.

I say "just about" any environment, because clearly there are *some* environments in which I (for example) might have been significantly taller. For example: an environment in which I undergo a surgical procedure that inserts an extra few inches of bone into my shins and thighbones. It seems (at best) debatable whether such an environment ought to count as part of the "relevant" range, or at any rate whether the existence of such an environment presents any serious challenge to the claim that, once nutrition and healthcare rise above some threshold level, variation in height is due almost entirely to genetic variation, and that there are fairly strict limits on the height potential of any given genotype. So for now, let me concede that we should exclude the massively-invasive-surgery environment from the space of relevant ones.

In that case, a height-hereditarian seems on relatively firm ground if she insists that in our population, height is genetically determined, in the sense that within a broad range of environments, differences in environment do not lead to differences in height. Now again, the hereditarian wishes to claim that IQ exhibits the same pattern. It is not that there are no known environmental variables that affect IQ; to the contrary, the list overlaps significantly with factors that influence height, and includes health, nutrition, very basic

education, and the presence or absence of physical and emotional abuse. When there is substantial variation in a population with respect to these environmental conditions—as there was throughout most of human history and still is in the developing world—the heritability of IQ will be lower and there may be significant opportunities to raise IQ by improving the environment. But—just as with height—once the environment, as measured by these variables, surpasses some threshold, further improvements to the environment will no longer yield improvements in IQ—the present environment already approximates the optimum, such that, in many cases, it "maxes out" the genetic potential for intelligence possessed by each genotype.

With this characterization of the hereditarian perspective in mind, let's return to the question I set aside before, namely: how should we actually go about defining the "relevant" range of environmental conditions? What is it about extra calories and vitamins that make these components of a "relevant" set of environmental conditions, and what is it about shin-lengthening surgery that seems to require that we leave it outside the relevant range of environments? Returning to the IQ case, imagine someone who, fresh from reading *Harrison Bergeron* and completely missing the point, concludes that the way to get rid of the Black-White IQ gap is to take all White children and subject them to several years of malnutrition, social isolation, and abuse, so that their IQs are suppressed and the White mean aligns with the Black mean of 85. So here we have an environment in which the "IQ gap" vanishes, but presumably neither hereditarians nor their opponents would consider this an appropriate solution, much less a rebuttal to the hereditarians' core claim.

(HED). Given these intuitions about what should and shouldn't count as part of the relevant range, what principles might be given for making this determination?

One possibility is that we should restrict our attention to environments that are considered "normal" or "natural" for a species. Levin (1997b, p. 86), for example, suggests that "calling a phenotype *natural* for an organism [means] that the organism will display it in environments like those in which its ancestors evolved." Similarly, behavior geneticist Sandra Scarr (1992), borrowing a term from Hartmann (1958) has argued that developmental scientists should be primarily concerned with developmental trajectories in the "average expectable environment" for that type of organism, and she similarly characterizes the relevant environmental conditions in terms of the organism's evolutionary ancestors. But this does not appear particularly promising, for as we are all aware the environments in which most humans now develop are profoundly "unnatural."

And more importantly, it's hard to see why we should feel obligated to discount environmental manipulations that help us achieve some desirable outcome simply because they require us to instantiate "unnatural" environments. Take the case of phenylketonuria (described in detail in Section 7.4.1 below). Surely by Levin's and Scarr's evolutionary criteria, a diet free of phenylalanine is unnatural for human beings, so if we wanted to we could conclude that it is "natural" for people with the genotype for phenylketonuria to exhibit cognitive defects. But given that we can *prevent* this outcome by putting people with this genotype into an unnatural environment—and we have very good reasons for wanting to do so—it would be difficult to argue that this environment should lie outside the range of explanatorily relevant ones. If evolutionary considerations do not help us

identify the relevant range of environments, how else might we do so? Answering this question requires considering what we mean by the notion of a *range* of environments in the first place.

7.3.2 What is an environmental range anyway?

The question of how to define a range of environments is relatively straightforward in certain circumstances, for instance when investigating the phenotypic heights of different genotypes in a species of plant at different altitudes. Since in this case is there is a single parameter along which environments differ, one that can be easily measured and precisely quantified, a range of environments can be easily delineated: we can ask, for example, what are the reaction norms for a series of genotypes as elevation ranges from 30m to 3000m? For simplicity, let us suppose that the norms of reaction for the different genotypes are identical and all are linear, such that height for all genotypes increases with elevation across the full range. We could then characterize the different altitude environments in this range as being more or less favorable to growth.

In principle, we could do likewise for IQ in humans: we could characterize environments as being more or less favorable to IQ development (more or less "intellectually enriching," we might say), and plot the phenotypic IQs of different human genotypes against these environments.[8] The hereditarian hypothesis, then, is that environments can be ranked from more to less facilitative of IQ development, and that the

[8] In principle, this process can be done where "genotype" refers either to a single polymorphism, a group of them, or an individual's full genotype. Where one or a few genes are considered, we would assume that the rest of the individual's genotype reflects the "average" for the population.

norms of reaction of "Black genotypes" within those environments will differ from those of "White genotypes." In particular, the curves for both genotypes will rise as the environment becomes more favorable to IQ development, leveling off as conditions reach those approximating those of the modern developed world, but with the "Black" curve situated significantly below the "White" curve. Thus, the situation might look something like the (hypothetical) Figure 7.3, with the different curves representing different racial groups.

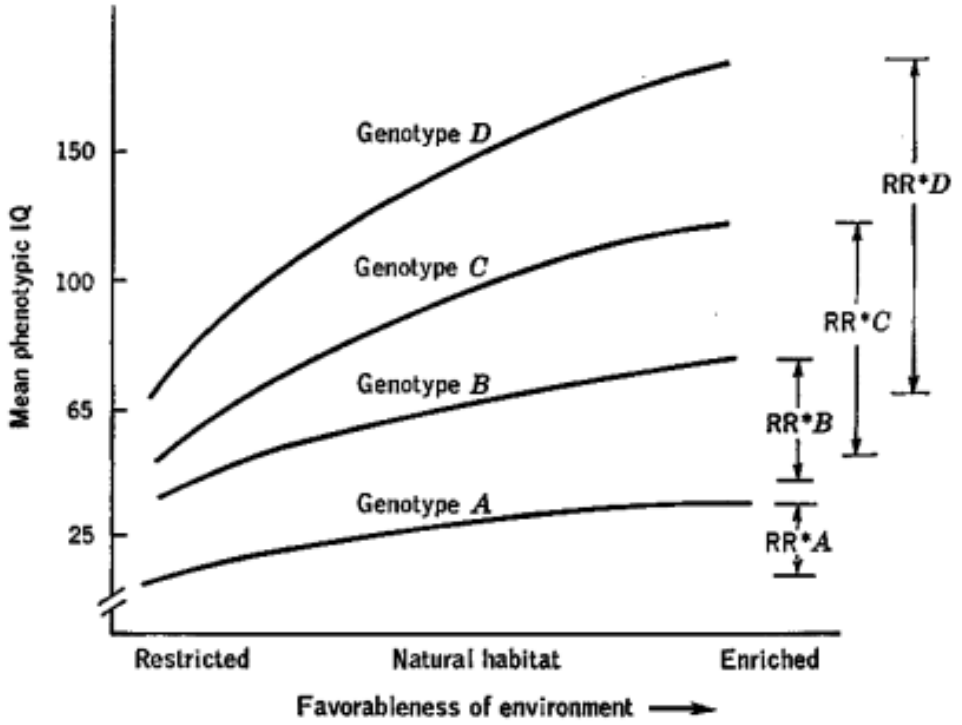

Figure 7.3: Hypothetical reaction norms for IQ (from Gottesman, 1963)

As we can see, the curves remain about the same distance apart throughout the range of environments. If this graph accurately represents the nature of group IQ

differences, then we should expect that as environments become more "enriching," we might expect to see, up to a point, overall gains in IQ among all groups, but the *gaps* will remain roughly constant.

Alternatively, we might think that the graph above does *not* reflect the nature of IQ differences, in which case we would expect that at some point, the curves for different genotypes (i.e., racial groups) will meet or cross—i.e., there might be significant G×E effects. The intuitive thought here is that improvements in, for example, economic stability and access to early-childhood education for historically-underprivileged groups might improve IQ levels among those groups to the point where they are equal with other groups (while leaving IQ levels of privileged groups unaffected). In other words, we might suppose that reducing disparities in social conditions between groups will constitute exposing all genotypes to a new environment, one in which all genotypes manifest similar IQ levels.

The usual hereditarian response to this optimistic line of thought is to insist that we have *already tried* the interventions described above, and that they have failed (the alleged failure of the Head Start program to make lasting differences in intelligence among Black children is frequently cited here—e.g., Herrnstein & Murray, 1996; Levin, 1997b). Thus, the hereditarian insists, there is no reason to expect that improving education, economic stability, nutrition, or any other variable will eliminate the IQ gap, because we have observed this gap in such a broad range of environmental conditions that we are licensed in making an inductive inference about what will happen in future environments. The hereditarian therefore assumes that an environment in which Black children are supplied

with the same nutrition and education from early in childhood as White children would essentially be *the same as* our present environment, at least with regard to its bearings on IQ in different populations. (Many hereditarians, recall, are inclined to believe that variation in IQ is mostly attributable to genetic variation once environmental conditions reach modern developed-world standards, much as is the case with height—i.e., they suggest that our current situation is like the right-most portion of Figure 7.3.)

The question of what we can infer about future environments from what we have observed so far relates to a common criticism of heritability mentioned at the beginning of this chapter, namely, that it is a *local* statistic—that is, heritability measures only the relative contributions of genetic and environmental variance to phenotypic variance *in particular population in a particular environment*. In a different environment, critics argue, the heritability of a trait might change drastically (Lewontin, 1974; D. S. Moore, 2008). Defenders of heritability (e.g., Sesardic, 2005, p. 78) respond that if heritability really provided *no* information *whatsoever* about the causes of phenotypic variance in environments other than the one measured, it is hard to see why scientists would have invested any effort whatsoever into performing such analyses. And indeed, several behavior geneticists have defended the generalizability of heritability measures to environments beyond the one originally measured (Loehlin, Lindzey, & Spuhler, 1975; Plomin & DeFries, 1976). How, according to these theorists, are we to determine whether the heritability of a trait observed in one population/environment is likely to be generalizable to other populations in other environments?

The answer, unsurprisingly, is that the more *similar* a new (hypothetical or actual) environment is to a previously-observed one, the more likely it is that the heritability of a trait in the new environment will be close to that of the old: "[heritability] values found in other populations under other circumstances will be more or less the same according to whether *the structure of the population and the environmental conditions are more or less alike*" (Falconer, 1989, p. 164, italics added).

This seems entirely plausible. The problem, however, is that we may often be in the dark about which environments are relevantly similar to which other environments. Predicting what will happen to a system under new, previously unobserved conditions requires some causal knowledge about how the system behaves. The more complex the system, the more detailed the causal knowledge required to ground inferences about the system's behavior under new conditions. For systems that are sufficiently complex (as, for example, human cognitive development), it is only through having *detailed* causal knowledge about a system that we can say with any confidence whether a change in some condition will make a difference to the outcome.

In other words, the idea of an "environmental *range*" implies a series of *adjacent* points (i.e., adjacent points along the horizontal axis of a reaction norm). And this makes sense when we are talking about environmental variables that are linear, continuous, and well-defined, such as temperature or altitude. If we have observed the behavior of a developmental system in a *range* of temperatures from -100° to 100°, we can perhaps be confident (though not certain) about how the system will behave at 103°, even without detailed mechanistic knowledge. But when it comes to the development of complex human

phenotypes, it is not so clear that we are in a position to say which environments are "next to" each other in the developmental landscape.

Applying this observation to questions about environment and IQ, my suggestion is that because we do not understand in any detail the mechanisms by which genes interact with environments to produce IQ, we are not in a position to say which interventions are likely to make a difference. This is the hereditarian's mistake: the hereditarian looks at a half-century or so of interventions designed to improve Black IQ and eliminate the racial IQ gap, and believes we are licensed in viewing proposals for further interventions (e.g., the interventions targeting working memory discussed in Chapter 5) as "more of the same." The hereditarian believes we have already observed phenotypic outcomes for Black and White genotypes *under a broad range of environmental conditions*, and are therefore licensed in thinking that we are unlikely to see anything new. But if what I have said is correct, we will not be in a position to say how broad or narrow the observed range of environments is until we understand a great deal more about how genes interact with environments to produce IQ. In particular, we will require *mechanistic* information about the development of IQ.

7.3.3 Explanatory interests and the delimitation of environments

As I noted in the sketch of my arguments given in Section 5.4, I believe that a difference in *explanatory interests* underlies much of the debate between hereditarians and their opponents. Here I wish to back up that assertion by demonstrating that hereditarians and their opponents differ in their respective explanatory interests in terms of the

importance they assign to phenotypic outcomes in environmental conditions that are *possible* but have not yet been instantiated or observed (for similar arguments, see Griffiths & Tabery, 2008; Longino, 2013; Tabery, 2009).

Critics of hereditarianism emphasize that because heritability is a local statistic (i.e., it applies only to the population observed, and leaves open the possibility that the partitioning of genetic and environmental sources of variation could go quite differently for the same genotypes observed in a different environment), the high heritability of a trait does not rule out the possibility that interventions to establish alternative environmental conditions might yield different results. Hereditarians, though generally more optimistic about the prospects of extrapolating observed heritability scores to new conditions, typically concede this point, but insist that its scientific significance is minimal: of course it is *possible* that by extrapolating development out to entirely novel conditions, we will observe different outcomes, but this possibility should not distract us from the business of explaining the world as we find it. Science is in the business of explaining the *actual world,* not in speculating about possible ones. And in the *actual world,* heritability analyses assign a large proportion of phenotypic variance to genetic factors, and so we have succeeded in explaining—i.e., identifying the relevant causes of—phenotypic variation.

Consider the following statements from hereditarians. Herrnstein, for example, states that "For me, it is no disgrace if my argument holds merely for *existing* societies, not necessarily *all possible* ones" (1976, p. 300, emphasis added). And Sesardic is particularly emphatic about the explanatory irrelevance of unobserved variation:

> Speculation about these remotely conceivable situations may then begin to dominate the picture so much that we witness the curious triumph over the *possible* over the *actual*. We should resist this and try to interpret the heredity-environment discussion so that it deals with our real world, not with uncontrollable "might-have-beens" and counterfactuals gone wild. (Sesardic, 2005, pp. 84–85, emphasis in original)

Even psychologist Earl Hunt, a comparatively moderate voice in the IQ debates, states quite matter-of-factly that "epidemiological [i.e., heritability] studies are likely to be of much more help than laboratory studies, because *we are concerned not with what might happen but with what actually does happen in our society*" (1997, p. 544, emphasis added).

But it is difficult to see what grounds these theorists' confidence that describing the patterns of phenotypic variation we observe under *present* conditions should take precedence over developing and testing hypotheses about the outcomes of novel interventions. Indeed, it is notable that hereditarian thinkers with a preference for restricting the space of explanatory relevance to the "real world" generally *assert* rather than *argue for* this stricture. And notice, too, that these thinkers do not even attempt to defend their circumscription of the space of relevant possibilities on empirical grounds—rather, they quite explicitly frame the injunction in *pragmatic* terms, i.e., in terms of *what we are concerned with*.

Now contrast the hereditarian perspective with the approach evident in the research of developmentalists, which consists in experimental studies that allow for observation of developing systems under a wide range of controlled but specific interventions, many of which push the system outside of the range of environments that might be considered "natural" (Gottlieb, 1991, 1995; Griffiths & Tabery, 2008; Meaney, 2010). In so doing, the

developmentalists are—in theory at least—able to characterize the behavior of the developmental system in more robust (i.e., *mechanistic*) terms, just as Celia Moore has accomplished with the rat pups discussed in Chapter 2. The information gleaned from such experimental work arguably constitutes a significant advancement in our scientific knowledge than would have been possible merely by partitioning the genetic and environmental sources of variation in a population under preexisting conditions. So why should behavior-geneticists be so dismissive of endeavors to expand our knowledge in this way?

In the behavior-geneticist's defense, it is true that pursuing an experimental approach to behavioral research—one that seeks to develop mechanistic models of development that can support robust predictions about the outcomes of novel interventions—is a much more daunting proposition than carrying out behavioral-genetic analyses. For one thing, the experimental model—at least as developmentalists have pursued it in their animal laboratories—is (for ethical reasons) of limited use when studying human beings. For another, the developmentalist/experimental approach forces us to acknowledge and account for considerably greater complexity in our models of the causes of developmental outcomes—i.e., we must be comfortable admitting to a great deal more ignorance about development than we would if we were to accept the output of heritability analyses as fully explanatory. And finally, it is true that for *any* explanation that one might try to offer in *any* domain of science, it is always possible that observed regularities will cease to obtain under hitherto unobserved circumstances, and we should not hold the very

practice of scientific explanation hostage to the necessarily limited nature of the conditions that we can observe.

On these grounds, it is perhaps not unreasonable for the behavior-geneticist to question whether the developmentalist's approach is properly scientific, in that it makes the ideal of *perfect* scientific knowledge the enemy of *good* workaday explanation (Scarr, 1995). So perhaps we cannot begrudge behavior-geneticists a preference for the comparatively clean explanatory strategies of partitioning observed phenotypic variation into a handful of components.[9] But at the same time, we should recognize that behavior-geneticists have ultimately not put forth any particular reason for thinking that it is *illegitimate* to be concerned with outcomes in unobserved possible environments, as many developmental biologists clearly are, and as I suggest we ought to be when it comes to the race-IQ gap. In short, when the hereditarian says that "we" are not concerned with developmental possibilities outside a narrowly-defined range centering on our present circumstances, we can respond: *what do you mean WE, kemosabe?*

7.4 Advantages of a mechanistic approach

The discussion in the previous section raises the question of *why* one might prefer one explanatory approach to behavioral development (e.g., a developmentalist or mechanistic perspective) over another (e.g., hereditability). I have alluded to what I take to be the primary difference here, but I now defend it explicitly: while heritability analyses

[9] In the next chapter I will, however, consider the possibility that those behavior geneticists (i.e., hereditarians) who have explicitly defended the irrelevance of novel environments in the case of the race-IQ gap are deserving of criticism—but this criticism would be of a *moral* rather than *scientific* sort.

may be sufficient for *explaining* or *predicting* outcomes given existing environmental conditions, mechanistic models provide all this *plus* significantly greater opportunities to *control* or *manipulate* the system in question. A mechanistic understanding of the development of a system affords greater opportunity to intervene in the operations of that system to produce a desired result. And this is true, it turns out, for intellectual development. Understanding in detail how genetic factors causally interact with non-genetic factors to produce brains that display intelligent behavior gives us greater potential to design interventions that change outcomes.

To illustrate this point, I discuss an example of how a mechanistic understanding of a "genetic disorder" (phenylketonuria, or PKU) has improved our ability to exert influence on the outcomes of a developmental process. These insights, I argue, would not have come about from adopting only a behavior-genetic perspective that identifies genetic and environmental sources of variation in the population.

7.4.1 Phenylketonuria

Phenylketonuria is a metabolic disorder characterized by inability to synthesize the metabolic enzyme phenylalanine hydroxylase (PAH). Inheritance of PKU follows an autosomal recessive pattern, and is linked to any of a number of mutations in the gene for PAH. The absence of the PAH enzyme results in an inability to metabolize the amino acid phenylalanine (Phe) into tyrosine (Tyr). Phenylalanine is a common nutrient found in animal protein and starchy foods, and in the absence of the PAH enzyme, any Phe that a person consumes accumulates in the bloodstream, while Tyr levels are deficient. Since

other amino acids—which are crucial for the synthesis of neurotransmitters—compete with Phe for uptake in the developing brain, elevated Phe levels reduce the availability of these other amino acids, and brain development is disrupted, with lifelong cognitive deficits as a result.

Children with the PKU mutation are thus born physically and cognitively normal, but are slow to develop, and without the appropriate interventions, will suffer lifelong cognitive impairments. Fortunately, a simple environmental manipulation has proved incredibly effective in preventing the symptoms of PKU from manifesting. Neonatal screenings (which are standardly performed in the developed world) are able to detect elevated levels of Phe and high ratios of Phe to Tyr, indicating the absence of normally-functioning PAH (due to the presence of the PKU mutation). Infants with PKU are immediately placed on Phe-restricted diets, and if this diet is maintained through the lifespan, brain development proceeds normally and children can grow up without any significant cognitive impairments.[10]

Now let's consider how hereditarian vs. mechanistic analyses might approach the case of PKU. The historical story of the discovery of PKU and its treatment is interesting (see Centerwall & Centerwall, 2000), but for the purposes of my analysis I'll imagine the story had gone somewhat differently, in order to highlight the methodological points that are relevant for assessing the hereditarian case for irremediable racial differences in IQ. Suppose that prior to the discovery of the biochemical basis for the cognitive deficits

[10] There is some evidence that even on a Phe-restricted diet, some outcomes for PKU patients are "suboptimal" (Enns et al., 2010; MacLeod & Ney, 2010; van Calcar et al., 2009). Research into the use of amino acid dietary supplements to correct these suboptimal outcomes is ongoing, and preliminary results are promising.

associated with PKU, researchers had conducted a heritability analysis. They would have discovered that the heritability of the cognitive deficits associated with PKU was essentially 1 (i.e., all of the variation in the symptoms would be explained by variation in the genotype—specifically, in the presence or absence of the mutated PAH gene).

Now let's imagine that researchers had inferred the presence of the genetic polymorphism responsible for the cognitive deficits associated with PKU, but knew little or nothing about the mechanisms by which the mutation results in microcephaly and other developmental abnormalities.[11] We can imagine that a few unsophisticated hereditarians about PKU might have declared the case closed: PKU is a genetic cognitive disorder, with mental disability manifesting for the PKU genotype in all relevant environments (where "all relevant environments" means simply "the environments we have observed so far").

Critics of such crude hereditarianism would surely have objected, and pointed out that the high heritability of the cognitive deficits may not hold under other environmental conditions—the reaction norm for the PKU genotype might be such that under certain environmental conditions, cognitive functioning improves or even equalizes with that of non-PKU genotypes. Suggestions might have been made for improving the cognitive-developmental outcomes of people with the PKU mutation. It might have been thought, for example, that children with the PKU mutation simply need better overall nutrition (e.g., more calories or more vitamins), or more intensive education, or more encouraging and nurturing caregiving, and so on. *Absent more specific information*, these are all plausible

[11] This, I would suggest, is not unlike our present situation when it comes to understanding the development of IQ: some genetic markers have been identified—though only tentatively (see Chabris et al., 2012)—that have small associations with IQ, but little is known about the mechanisms by which these genetic factors influence general intelligence.

ways of improving an individual's cognitive development. These interventions, however, would not have been effective in improving cognitive development in PKU patients, because they would not have made a difference to the operation of the mechanism that produces cognitive abnormalities (i.e., the crowding out of neurotransmitter-building amino acids by excess Phe).

At this point, a more sophisticated class of hereditarians might have said: "Look, we've tried a number of plausible interventions and exposed PKU children to a *wide range* of environments, but nothing works. Of course it is *possible* that in some hitherto-unobserved environmental condition, radically different from any environment any human has ever experienced, PKU children will exhibit normal cognitive development. But in science we must be concerned with *actual* conditions, or at least those that are normal for the species. It is time to accept that PKU mutations cause cognitive deficits in the full range of normal human environments." (For hereditarian arguments to this effect regarding racial differences in IQ, see Levin, 1997; Scarr, 1992; and Sesardic, 2005.)

These hereditarians would have been right about one thing: antecedently, there is no particular reason to count Phe-restricted diets as part of the "normal," "natural," or "relevant" environmental conditions of human development. Certainly Phe-restricted diets were practically nonexistent in the environments in which modern humans evolved, and only a tiny fraction of humans consume such diets today. By these standards, Phe-restricted diets ought to be considered "extreme" and explanatorily irrelevant environmental conditions.

I sincerely doubt, however, that anyone would seriously suggest that the developmental trajectory of humans with the PKU mutation *outside* of the "normal" human environmental range (e.g., on a Phe-restricted diet) is of no scientific interest. Clearly, the details of how the PKU mutation eventuates in cognitive deficits is not only theoretically relevant, but practically and morally important as well. Therefore, behavior-genetic analysis that simply associates the presence of the mutation with cognitive deficits is inadequate.

We are fortunate that this is not how the discovery of PKU and the development of its treatment actually went. Rather, because a sketch of the mechanism was available (again, I'm idealizing the historical record to some degree), an intervention could be designed. Consider the mechanism sketch in Figure 7.4:

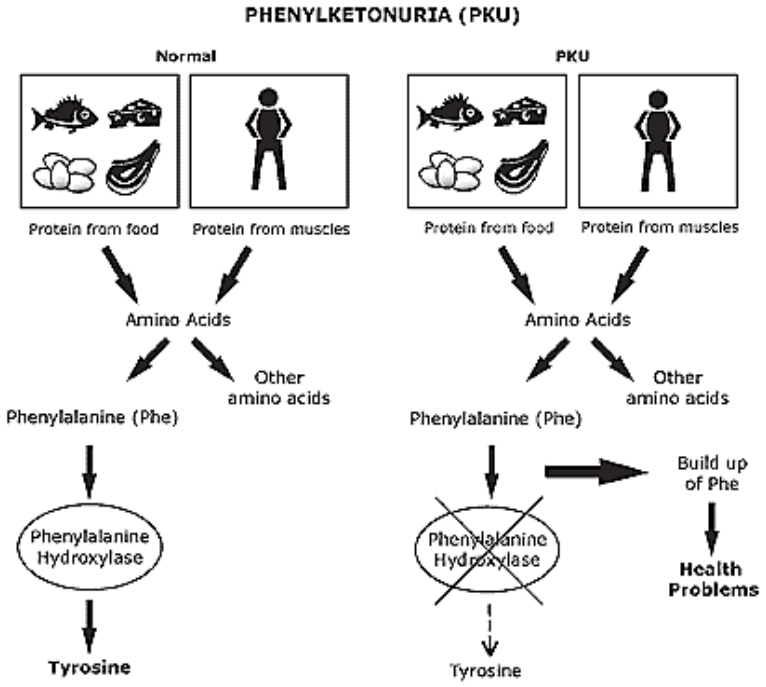

Figure 7.4: Mechanism sketch of normal Phe metabolism vs. PKU

This sketch allows us to target particular parts of the mechanism, such as the build-up of Phe, for interventions. We can predict that by restricting the intake of Phe, neurodevelopmental abnormalities will be averted. And knowing that Tyr deficiency contributes to lingering symptoms even among patients on a Phe-restricted diet allows us to predict that dietary supplements will further improve outcomes.

This sort of information, in contrast, is absent under a behavioral-genetic analysis. If all that is known is that a genetic factor is correlated with (or even *causes*) a particular phenotypic outcome in certain environments, then if we wish to manipulate that outcome our only option (other than trying to manipulate the genotype directly) is to randomly attempt intuitively-plausible interventions and hope they work. But with even a fairly abstract sketch of the relevant mechanisms, we can seek to target our interventions to specific entities or activities that are likely to make a difference to the outcome.

7.5 Conclusion: Mechanisms, interventions, and group differences in IQ

Now let us apply these lessons to arguments for the ineliminability of racial differences in IQ and other traits. Hereditarians, again, treat the persistence of such differences under observed conditions—in spite of various interventions directed at eliminating them—as evidence that these differences will manifest in all "practically possible" environments (Levin, 1997b, p. 89). But if I am correct, this conclusion is at best premature. For just as it would have been a mistake for our hypothetical PKU-hereditarians to conclude that *no* intervention could make a difference to the manifestation of the PKU

phenotype, prior to specification of the relevant mechanisms, it is a mistake to assume that we know whether the interventions that have been directed at closing the race-IQ gap have been aimed at the right entities or the right activities.

What is needed, then, much like with PKU, is a detailed understanding of how different genotypes interact with different environments to influence traits like intelligence and personality. The mechanisms by which an individual develops a particular degree of intelligence, say, will operate at multiple size levels, and will include such concrete entities as DNA sequences and the proteins they synthesize, ingested nutrients, hormones produced in response to environmental stimuli (e.g., stress-inducing dangers or affection from a caregiver), and more abstract entities like representations of linguistic entities in the environment and the working memory system.

Needless to say, modeling the mosaic, multi-level mechanisms that underlie the development of complex human behaviors is a daunting task, much more so than modeling the mechanisms of PKU. Yet it is only with these kinds of mechanistic models that we would be in a position to say which sets of environmental conditions are causally similar to which others, and hence whether a novel intervention is likely to yield different results from other, previously-attempted interventions. I conclude, then, that insofar as we are concerned with eliminating racial IQ gaps—and as I have argued, I think we should be very concerned with this—we have all-things-considered reasons to prefer mechanistic strategies over behavior-genetic ones in the context of studying intellectual development.

CHAPTER 8: IS RACIAL SCIENCE RACIST?

8.1 Introduction

It is frequently alleged, in both academic and popular contexts, that certain kinds of scientific investigations of race constitute a form of racism (often called "scientific racism"). Specifically, the charge of racism is frequently levelled at the sorts of scientific investigations discussed in the previous three chapters, those that investigate the possibility of "innate" or "genetic" average racial differences in important behavioral traits (e.g., IQ, aggression, parenting behaviors, etc.). For the purposes of this chapter, I'll refer to such research programs collectively as "racial science."[1]

Perhaps surprisingly, allegations of racism directed at racial science are typically made without explicit appeal to any particular definition of "racism." Rather, it is apparently assumed that racism—whatever it is—can be unambiguously recognized, and that the description of some aspect of racial science on offer makes it plain that it qualifies as racist. My purpose in this chapter is to subject such claims—claims that investigation of potential racial differences is (either inherently or *de facto*) racist—to careful and systematic analysis by examining the institution of racial science through the lens of established philosophical theories of racism. My project here is largely exploratory: I aim to shed further light on the moral status of racial science by asking what

[1] Note that this focus on psychological investigations is a somewhat narrow usage of the broad term "racial science," which in other contexts could mean the scientific investigation of whether the facts of human biodiversity license divisions of the species into "races," or scientific investigation of physiological differences (including those of medical import) between human populations. Many of the same issues discussed in this chapter may also apply to investigation of physiological differences, but these are beyond the scope of the present discussion.

sort of case can be made that racial science is racist under several different theories of what racism is. However, my analysis tentatively concludes that although charges of racism are often more difficult to sustain than many have supposed, many aspects of racial science are indeed deeply morally problematic. I begin with a description of the motivations for pursuing this question, followed by an evaluation, in Section 8.2, of the suggestion that it is racist—or in some other way morally censurable—to so much as pose and investigate the sorts of questions that are at the core of racial science. Section 8.3 then considers whether racial science may be counted as racist under each of three recent philosophical analyses of racism: racism as *pernicious belief*, racism as an *institution*, and racism as race-based *ill-will* or *disregard*.

8.1.1 Motivations

Below I will discuss several reasons why a careful consideration of whether racial science should be counted as racist is an important project. I would like to begin, however, by motivating the topic of this chapter in a slightly different way. Specifically, I would like to present some rather striking statements that have recently been offered by some of the more daring practitioners of racial science. While this may seem like a cheap emotional appeal, my intent is not to shock the reader into agreeing that these scientists are very bad, very racist men. Rather, these statements are intended to prompt reflection on how we might distinguish between empirical scientific claims that we perhaps find distasteful or mildly irresponsible, and claims about race that so flagrantly flout moral norms as to be guilty of racism.

First, consider the remarks of Donald Templer (a former professor of psychology) at the 2004 conference of the White nationalist[2] organization *American Renaissance*. A reporter at the conference provides the following account of Templer's address:

> Many psychologists recommend psychological therapy for black prisoners, but Prof. Templer disagrees: *'They need 60 hours a week of work therapy. That would give them less time for manufacturing alcohol and weapons, trafficking drugs, and giving each other AIDS.'* (Jobling, 2004, emphasis added)

Next, consider the words of a thinker we have encountered several times before, the late Canadian psychologist J. Philippe Rushton. In 1999, Rushton mailed 40,000 copies—unsolicited—of an abridged version of his book *Race, Evolution, and Behavior* to North American academics. The book contains the following passage:

> In Africa, the female-headed family is part of an overall social pattern. It consists of early sexual activity, loose emotional ties between spouses and sexual union and the procreation of children with many partners. It includes fostering children away from home, even for several years, so mothers remain sexually attractive. Males likewise compete more for females and fathers are less involved in child rearing…In Black Africa and the Black Caribbean, as in the American underclass ghetto, groups of pre-teens and teenagers are left quite free of adult supervision. (Rushton, 2000, p. 16)[3]

[2] For those who are innocent of such matters, note that there is a distinction between White *supremacist* views and White *nationalist* views (Swain & Nieli, 2003). White supremacists believe that northern Europeans and their descendants constitute a "master race" fit to dominate all other groups. White nationalists, in contrast, believe that it is legitimate for members of each race to band together to promote their own interests against those of other races; hence Whites should maintain solidarity in order to direct a greater share of social and economic influence out of the hands of non-Whites and into the hands of Whites (while recognizing that other racial groups have the right to do likewise).

[3] It is worth pointing out the irony of Rushton's dismissive attitude towards Black parenting styles (see also Levin, 1997b, p. 371) given that hereditarians persistently argue that differences in parenting behavior have virtually no effects on children's behavioral outcomes (Jensen, 1996; Scarr, 1992).

Lest there be any doubt about what these racial scientists are claiming, note that it is clear from the context of both men's work that these statements are intended to express the view that the behaviors ascribed to Blacks are not the result of local cultural forces, but rather reflect Blacks' "genetic programming."[4]

It is safe to assume, I think, that most decent-minded people will find both Templer's and Rushton's words deeply offensive. And while there is perhaps something distasteful about *any* claim that posits "innate" and immutable racial differences there is (it seems to me at least) a stark qualitative difference between the force of the two statements quoted above and some of the claims we considered in the previous three chapters—for example, the "core hereditarian claim" (HED) that Blacks will exhibit lower IQs than Whites across all relevant environmental conditions. The latter seems at least plausibly defensible as a legitimate scientific claim, whereas Templer's and Rushton's statements—while appearing to express a series of empirical claims—immediately strike us as outside the bounds of acceptable ways for a scientist to express an empirical proposition.

Now, I do not wish to lean too heavily on intuitions about these two statements—what I have to say in this chapter will, I believe, be just as plausible (or not) even if the reader's intuitions about these statements are completely at odds with my own. But on the assumption that the reactions described above on not entirely idiosyncratic, I think the words of Templer and Rushton should serve both to illustrate and to motivate the questions I take up in this chapter. To wit, something is very clearly *wrong* with what these scientists are saying about race. And yet much of what surrounds these claims (in Rushton's case, at least—Templer's informal remarks are a

[4] I am grateful to Mike Zenz for pointing out the need to clarify this point.

somewhat different matter) *looks* like ordinary scientific practice: Rushton's book, while interspersed with claims like the above, also summarizes and synthesizes a wide array of empirical studies found in peer-reviewed scientific literature.[5] So while we might have our doubts about Rushton's scientific inferences, it seems we would need something more than just these doubts to elevate a critique of Rushton (or Templer) beyond a mere empirical disagreement and into the realm of serious moral criticism.

Thus, part of the challenge taken up in this chapter will be to determine whether it is possible to articulate what is *wrong* with the sorts of statements presented above, while preserving the belief (rejected by some, but assumed here) that the state of scientific evidence regarding "innate" or "genetic" racial differences is something about which reasonable people may disagree. With these reflections in mind, let us now consider some more formal motivations for addressing the question at hand, viz., *is racial science racist?*

Firstly, in contemporary society, racism is rightly regarded as one of the most serious moral ills we face, and a charge of racism is one of the gravest moral critiques that can be leveled at a person. It is thus important to know whether this very serious charge is warranted by the institution of racial science and those who practice it, if for no reason other than the importance of dispensing serious moral criticism where it is warranted, and withholding it where it is not.

Secondly, and relatedly, several authors (Blum, 2002; Miles, 1989) have expressed concern over the possibility of "conceptual inflation" in applications of the term "racism." These authors

[5] This is certainly not to say that Rushton's methodologies should be assumed trustworthy; indeed his methods have been widely criticized (including among those we might deem "moderate" voices in the IQ debates—Cronshaw, Hamilton, Onyura, & Winston, 2006; Hunt & Sternberg, 2006; Wicherts, Dolan, Carlson, & van der Maas, 2010). These scholars have been particularly critical of Rushton's persistent reliance on extremely low estimates of average IQ in sub-Saharan Africa—estimates that were generated using unrandomized convenience samples, and seemingly subject to a number of other methodological confounds as well.

point out that if the definition of "racism" is allowed to expand too freely, and accusations of racism made too indiscriminately, the word will quickly lose its force as a serious condemnation of a grievous moral ill. Thus, we must be cautious in our application of the term, and we must ensure that accusations of racism are constrained by a coherent theory of what racism is and is not.

Thirdly, it should come as no surprise that practitioners of racial science forcefully deny that they or their work are in any way racist. Plausibly, much of the divergence of opinion on this question owes to different conceptions of what "racism" is held by proponents and detractors of racial science. The possibility of such crosstalk highlights the importance of specifying in precise terms different notions of racism and assessing their plausibility on independent grounds. By making our commitments regarding the nature of racism explicit, we can be sure that we are embarking on a clear and productive endeavor when we ask whether racial science is racist. Moreover, we can minimize the extent to which critics of racial science can use accusations of racism as a cheap polemical cudgel, and we can also ensure that practitioners of racial science are not given an "easy out" by way of adopting (explicitly or implicitly) an inordinately weak or narrow conception of racism.

Finally, there are many questions of great practical import in the vicinity of racial science. As we have seen throughout the last several chapters, the questions addressed by racial science are relevant to public policy in the domains of education, employment, and criminal justice, among others. But in addition, there are also more immediate questions regarding the extent to which racial science should be pursued, funded, encouraged, tolerated, and popularized. To the extent that racial science is genuinely racist, this would provide strong *pro tanto* reasons for marginalizing it through various means.

8.1.2 Two caveats

With these motivations on the table, let me briefly mention two caveats about my aims in this chapter. The first is that my chief concern in this chapter will be to examine what case can be made for *public* accusations of racism. Plausibly, a much higher evidential standard is required to justify a *public* pronouncement that a person or institution is racist than to speculate privately that this might be so. Because such public accusations are already quite common—even in relatively sober academic and journalistic publications—clearly many believe that such evidential standards are met. It is therefore of particular importance that we seek to resolve—or at least shed some light on—whether such pronouncements are justified.

The second caveat is that while in previous chapters I offered some discussion of the plausibility of certain claims about racial differences given current evidence and methodologies, in this chapter I aim to stay as neutral as possible on any *factual* claims about racial differences as they are investigated by racial science; my goal is to assess the justification for accusations of racism *without* begging any questions about how the empirical matters will turn out. As we will see, questions about the potentially racist nature of racial science cannot be entirely separated from the empirical content of the relevant theories and the state of the evidence. But I will endeavor to establish, so far as I can, what sort of case can be made for the racist nature of racial science even if we remain agnostic about the substantive empirical claims currently debated in the field.

8.1.3 Defining racism

I should also clarify from the outset my approach to the question of what *racism* is, and how we might decide whether the term is legitimately applied to racial science. Clearly in order to

determine whether racial science is racist, we must have a theoretical account of what racism is. So what is a "theoretical account of racism" or a "theory of racism"? As I use these and related terms here, I take these accounts to be good old-fashioned conceptual analyses of the word "racism." That is, I will be treating these theoretical accounts of racism as attempts to systematize and clarify folk beliefs about what racism is in accordance with ordinary usage of the term. Thus, these theories should specify what sorts of entities (people, attitudes, institutions, etc.) can properly have "racism" predicated of them, provide reasonably clear criteria (if not necessary and sufficient conditions) for something's being racist, illuminate any inherent normative implications of the word, and so forth, all while preserving some critical mass of intuitions about proper applications of the term "racism."

Importantly, however, the discussion below is not meant to identify the One True Theory of racism, nor even to compare the relative merits of the various theories on the market (though I will mention advantages and disadvantages of each theory as relevant). Like other commenters on racism (Blum, 2002, 2004; Faucher & Machery, 2009), I endorse a *pluralist* position on what racism is: there are multiple understandings of racism that are applicable in different contexts. Similarly, I endorse pluralism about what sorts of things can be properly called racist. Thus, people, beliefs, attitudes, jokes, symbols, and institutions can all be racist, though they will be so in somewhat different ways (cf. Blum, 2002, Chapter 1).

One particular distinction is important here, one that is frequently blurred in public discourse surrounding racism: I will assume that there is a distinction between, on the one hand, racist *persons* and, on the other hand, racist *attitudes*, *motivations*, or *actions*. That is, an *attitude* held by a person or an *action* performed by a person can be racist without the *person* being racist.

For example, telling a racially insensitive joke might qualify as a racist action, yet the person who tells it may well not be a racist person if, for example, she tells the joke because she thinks it will impress her friends and not because she is motivated by genuine antipathy towards the targets of the joke. For present purposes, we may define a racist person roughly as someone for whom racist motivations and attitudes form a central component of his moral character (Blum, 2002). We should bear this distinction in mind as we explore the possibility that certain aspects of racial science may be racist—arriving at an affirmative answer does not necessarily entail that any racial scientists are racist *people*.

While I am generally neutral on most questions regarding conceptual constraints on a theory of racism, I do adopt one fundamental commitment about what "racism" means: *racism is necessarily morally wrong*. There are definitions of "racism" that hold otherwise (Mills, 2003; Shelby, 2002), but it seems clear that everyday usage takes racism to be a serious moral ill, and—most germane to our purposes here—accusations of racism leveled at racial science are quite clearly intended to be moral condemnations.[6]

Finally, note that in evaluating the applicability of philosophical theories of racism to racial science, I do not intend to take the success or failure of these theories at counting racial science as racist as a litmus test for their adequacy. In other words, I am treating the racist nature of racial science as an *open question* rather than as a *datum*. My aim in applying each of these theories to racial science, rather, is to "cast a broad net" over the question of whether racial science is racist,

[6] I should note here that I will be sidestepping—as much as possible—questions arising in normative ethics regarding the "best" moral theory (however this might be defined). In particular, for present purposes I am agnostic about whether all morally wrong actions share some particular feature (such as failing to maximize utility or demonstrating lack of respect for persons). Rather, I will be proceeding from the assumption that we can clearly recognize certain things as morally wrong even if we might disagree about the source of their wrongness.

by asking whether and to what extent *any* of the theories of racism that have had some traction in the literature provide good grounds for attributing racism to racial science.

8.2 Is it racist to ask the questions?

Before taking up individual theories of racism and their applicability to racial science, we should first consider one approach that would classify racial science as racist right out of the gate. As we have seen, there is much to worry about, morally speaking, when it comes to the possibility of racial differences in important psychological traits, and so one response to these moral quandaries would be to insist that research exploring the possibility of racial differences should not be performed at all—indeed, one might suggest that to even take up the question of whether there are such racial differences is an inherently racist endeavor. Louise Antony, for example, refers to Herrnstein's investigations of race and intelligence as his "racist undertakings" (1993, p. 144). And, while stopping short of suggesting that pursuing racial science is racist, neuroscientist Steven Rose (2009) argues that research on intelligence and race (or gender) should not be performed because, as he titles his article, "science and society do not benefit."

At first blush, this perspective seems inconsistent with the way science is generally practiced, and with the way we usually think about how scientific research can contribute to the formulation of policy. That is, so long as we concede it is *possible* that racial and ethnic groups differ in intelligence (genotypic or phenotypic), and that the possibility of such differences is (*prima facie*) relevant to the construction of social policy, what reasons could we have for declining to investigate this empirical question, using the best methods we have for investigating empirical questions? And, moreover, if racism is morally bad, counting such investigations as racist would

require showing that they are morally defective in some way. Yet we might ask: how could executing an empirical investigation of a non-trivial empirical question be *ipso facto* morally condemnable? On what grounds can thinkers like Antony and Rose condemn such practices?

Well, the Antony/Rose strategy is to mount an accusation of racism against racial science by emphasizing the biological non-reality of race. The argument goes like this: science is concerned with, among other things, identifying relationships among naturally-occurring properties (e.g., between mass and gravitational attraction, between bacteria and disease, and so forth). But if race is not an "objective" biological category—i.e., if studying the things we call *races* fails to "carve nature at its joints"—then the usual reasons for directing scientific attention to correlations among properties do not apply. Therefore, absent some further explanation, it would be a mystery why questions about correlations between race and other properties (intelligence, altruism, and so forth) continue to be the subject of so much scientific attention and effort. Faced with this situation, one might hypothesize that *racism* constitutes the "further explanation" for the continued interest in racial science: it is because scientists (and perhaps the public) harbor racist attitudes of some sort that the question of whether "races" differ in intelligence seems like an interesting or even reasonable question to ask. As Antony puts it:

> [I]f such classifications as race fail to reflect deep regularities in human biology, and reflect instead only historically and culturally specific interests, then there is no reason, *apart from racist ones,* to investigate the relation between race and some presumably biological feature of human beings…even from [an extreme empiricist] perspective it must be an arbitrary choice to investigate one set of such correlations rather than another. Why intelligence and *race*? Why not intelligence and number of hair follicles? (Antony, 1993, p. 144, emphasis in original)

While I think there is something right about these questions that Antony, Rose, and other scholars have raised about racial science, in the end I think there are significant limitations to these critiques. Throughout the rest of this section, therefore, I will first note some points of agreement and highlight some important considerations to which these criticisms draw our attention, after which I will explain why, even so, I think that possible relationships between our racial/ethnic categories and important psychological characteristics are in fact legitimate subjects of scientific study. It follows from this that the modern institution of racial science is not inherently racist. Nevertheless, I close this section by expanding upon the insights of Antony, Rose, and others, in order to show why the superficiality (and perhaps artificiality) of racial categories does still undermine the moral legitimacy of racial science to some extent.

8.2.1 The contingency of race and of racial science

First, recall my arguments from Section 5.5: I endorse the claim that race is *not* an objective biological category. Rather, racial categories are *arbitrary* divisions imposed on continuous human biological variation. While there are clearly systematic differences in genetic structure among human populations (where "populations" can range in size from a few villages to an entire continent), our common-sense "racial" divisions do not have any particular explanatory significance among the myriad ways we could partition genetically distinguishable populations. Thus, while there are legitimate scientific reasons to study differences among genetically distinct groups, we do not get any particular inferential payoff from using *racial* groups. So Antony is correct that the choice to study relationships between behavior and *race*, as opposed to other categories, cannot, strictly speaking, be grounded in the biological informativeness of race. But,

does this establish that the use of race as a theoretical parameter in scientific inquiry is without justification, and should be discarded? Antony and Rose seem to think so. But neither gives us a clear picture of what alternative they have in mind, of what the scientific study of human behavior (and in particular, human behavioral *differences*) might look like in the absence of race. So it will be instructive, I think, to begin with the biological non-reality of race and then spell out what a "race-free" behavioral science might look like.

The easiest way to do this, I think, is to turn back the clock and imagine an alternative history in which the notion of race simply never emerges. The story goes like this: because race is not a real biological property—out there in the world waiting to be discovered, like blood types or viruses—our use of the particular set of racial categories we've settled on (and indeed, our use of the concept of race at all) seem to be *contingent*, in a way that our reliance on other, genuine biological properties is not. That is, idealizing somewhat, we can suppose that by doing halfway-decent science for long enough, our species was bound to discover certain elements of the biological furniture of the world, like blood types, viruses, and so on. But (by hypothesis) not so with race: if race is not the sort of category that can be discovered through biological inquiry, then we can imagine that in our alternative history, all the facts about human genetic and somatic variation are the same, and the development of the biological and behavioral sciences through time is more or less the same, but no race concept emerges—it simply never occurs to anyone to impose upon human variation the peculiar kind of structure that characterizes our racial way of thinking (i.e., dividing people up into a small, discrete number of geographic groups).[7]

[7] Of course, it is possible that our tendency to "think racially" might be inevitable (or nearly so) even if race is not a category discoverable through scientific inquiry, if there is some universally shared innate mechanism that predisposes us to think in racial terms. Some theorists have argued for this possibility (Gil-White, 2001; Hirschfeld,

In this world, then, the somatic characteristics that now form the basis of our racial categorizations (such as skin color, hair texture, and so on) simply never become imbued with the deep significance that we (as racial thinkers) assign to them. We need not suppose that in this alternative history, people would be completely blind to various perceptually salient somatic differences among people of different continental ancestries, but we can imagine that these characteristics came to have no greater significance than, say, traits like height or eye color—things we notice but do not use to divide people up in any serious way.

Now, had this been our history, what might our behavioral sciences look like? Well, most obviously, "race" would not be a theoretical parameter in the study of human behavior. But in other ways, I suggest, things might look very much the same. We can suppose, for example, that in this alternative history the field of behavior genetics emerges, similarly populated by scientists who are interested in the relationship between genetic variation and behavioral variation.[8] Eventually these behavior geneticists would get around to studying the genetic correlates of the very same sorts of traits to which we are accustomed (intelligence, alcoholism, "big five" personality factors, and so on). And, at some point these behavior geneticists might well see fit to investigate whether and how genetic differences among people with different geographical origins are associated with behavioral differences. Perhaps such studies would find associations between the genes prevalent in particular geographic populations and particular behaviors, and perhaps not.

2012; Machery & Faucher, 2004), but I believe Adam Hochman provides persuasive arguments against it (Hochman, 2013; see also Kurzban, Tooby, & Cosmides, 2001).

[8] It may be that by introducing this possibility, I am straying from what Antony and Rose actually have in mind. That is, perhaps the alternative they envision is a world in which the nature and causes of human differences are simply not seen as particularly urgent matters for science to investigate (this is perhaps especially plausible regarding Antony, given her Chomskyan emphasis on the theoretical importance of *similarities* among individuals and of human *universals*). Whether or not this is what Antony and/or Rose have in mind, it strikes me as a rather far-fetched (if admirable) suggestion—the impulse to understand how and why we differ from one another appears quite deeply ingrained, at least for many.

Either way, whatever language they might use to describe their hypotheses and their results, these behavior geneticists would lack the vocabulary of race. And not only would it never occur to them to think of themselves as studying "racial differences in behavior," it also would never have occurred to them to divide people up using the categories *we* think of as races. From their perspectives, and the perspective of their audience, they would simply be investigating differences among individuals.

Now that we have a picture of what behavioral science might look like devoid of the concept of race, two questions arise about the scenario I have sketched. First: is this state of affairs an improvement, morally speaking, over our familiar system that treats race as an important theoretical parameter? Second: Can we get there from here? That is, given that the history sketched above is *not* our history, and we *do* have a concept of race, could we, as Rose and Antony suggest, do away with race as a theoretical parameter?

It is tempting to say "yes" to the first question. For we might think that whatever ills might accompany the investigation of individual differences in this alternative society, at least these investigations would not be saddled with the heavy moral baggage that comes with our particular history with race. However, the moral status of this alternative society is exceedingly difficult to assess given the sheer unruliness of the counterfactuals we would need to consider. How far would we have to turn back the clock to reach a time before people thought in racial terms?[9] How would

[9] There is disagreement about just when the modern notion of *race* arose in the course of human history. One popular view is that the idea of race was absent in the ancient world but emerged in the eighteenth and nineteenth centuries as Whites sought to reconcile enlightenment ideals with slavery and colonialism (Blum, 2002; Smedley, 1999; Zack, 1996). On the other hand, those concerned to argue for the biological reality and fixity of race (Miele & Sarich, 2005) also claim that the modern idea of race has existed for as long as different continental populations have been in contact. There is no need to settle this question here; rather, the key point for our purposes is that imagining a world in which neither scientists nor laypersons developed the notion of race—as we now understand it—requires imagining a world where most of the events of *at least* the last several hundred years—and perhaps several thousand—were entirely different.

our society be structured? How would we conceive of matters like social inequality? So I am pessimistic about the prospects of a satisfactory answer to this question, and it is by no means certain that studying human variation without the bogeyman of race would have raised any fewer moral quandaries in the end. However, even if we grant that yes, it would have been better had we never devised the artifice of race and built it into our social science, we must ask whether it is possible, now that we've got it, to actually dispense with it.

8.2.2 The contingent necessity of racial science

In responding to this question, the first thing to note is that even if race it not a legitimate biological category, it is clearly an important *social* category, in that our assignations of people—ourselves and others—to racial categories clearly do a good deal of explanatory work when it comes to accounting for both interpersonal relations and the structure of society as a whole. This being the case, it is difficult to imagine how we could ignore questions about how race—however arbitrary the notion might be at the end of the day—relates to other important aspects of social behavior, including intelligence, violence, poverty, and so on.

The essential problem for the Antony/Rose suggestion of abandoning race as a theoretical parameter in social science is that *in the here and now*, we cannot help but see the world in racial terms. And what's more, in the here and now, racial science is an established institution. The pervasive social inequality among the groups we call races is highly salient to us (most of us, anyway), and racial science has stepped forward to offer explanations for this inequality. It would perhaps be better if neither of these things were true. But given that they *are* true, it seems there is

little choice but to continue to seek a more complete understanding of why those classified as Black do worse than those classified as White, and how we might change the situation.

Given these considerations, it is unclear what alternative Antony and Rose might offer us. Suppose that tomorrow the world's scientists unanimously agree to abandon race as a variable. How are we to know if we are making progress towards ensuring that members of all groups are allowed to develop to their full cognitive potential? Clearly, with circumstances being what they are, the decision to use race as a theoretical parameter in scientific inquiry cannot be—as Antony borders on suggesting—inherently racist. Indeed, it would be difficult to accuse those scholars who acknowledge the existence of an IQ gap, but who study its contours for the express purpose of figuring out how to eliminate it (Flynn, 1999; Fryer & Levitt, 2004; Nisbett, 2009) of engaging in "racist undertakings."

I conclude, then, that studying the relationship of race with behavioral outcomes (including intelligence) is something that should be done, at least in the near-term. And, *a fortiori*, pursuing such questions is not necessarily a racist or even morally objectionable endeavor; So, as concerns the question posed in the title of this chapter, racial science is not automatically *racist* science.

8.2.3 Racial preoccupation and the moral evaluation of racial science

However, it does not follow from the above discussion that nothing is out of bounds in pursuing scientific research into the possibility of differences among human groups. To return to Antony's original point: we must keep in mind that race/ethnicity is but one (or at most two) out of a very large number of variables that we might choose to include in our scientific investigations of human behavior. And indeed, many scholars have drawn attention to the seemingly

disproportionate degree of attention directed towards questions of race over other potential correlates of behavioral differences we might study. In addition to Antony's fanciful (but incisive) example of "number of hair follicles," Rose (2009) points out that little scientific attention is directed towards measured IQ differences among geographic regions in Wales, and Loury (2002, pp. 85–86) makes similar observations regarding regional IQ variation in the United States and the tendency of IQ to decline with age (we will return to Loury's observations in Section 8.3.3 below).

These examples should give us pause in considering the common defenses that racial scientists offer for dedicating time and energy to studying racial differences. These scientists typically justify their investigations by pointing to the intrinsic value of knowing the truth of the matter, or the indispensability of this knowledge to the development of effective policy, or both. The question of whether truth is intrinsically valuable is a complicated one that I cannot address here,[10] but as we have seen I am in agreement with the latter, at least in principle. So there are legitimate scientific and social justifications for studying race and behavior. And, I would argue, the principle of charity demands that we (for the most part) take racial scientists at their word that their reasons for studying race and intelligence are among the legitimate ones. Crucially, however, we can also recognize that every scientist-hour dedicated to investigating the relationship between race and behavior is time spent *not* studying potential behavioral correlates of a multitude of *other* social and biological categories, some of which, it would seem, are of equal theoretical and practical importance (and so ought to have equal claim on our attention). So, while racial scientists might be forgiven for choosing *race* over *number of hair follicles*, the neglect of intra-racial geographic and age variation seems more difficult to explain away.

[10] For a defense of the intrinsic value of Truth, see Lynch (2004); for a more skeptical view, see Kitcher (2001)

The worry here is that by dedicating a disproportionate amount of energy and attention to the relationship of *race* to behavior, racial science tips its hand, revealing that its motivations are not so pure and disinterested after all. In other words, it is all well and good for a community of racial scientists to insist that they do what they do because it is essential that we find out the truth about whether and how the races differ. But this response may appear disingenuous when we begin to take into account all the other potential research parameters that might shed light on some policy question or other—and this is especially true given the clear risk of harm attached to racial science because of the discipline's sordid history.

Of course, spelling out exactly what is a "proportionate" amount of attention to direct towards questions of race as opposed to other categories will not be an exact science. But nevertheless we can advance the following principle: to the extent that racial science exhibits a disproportionate preoccupation with race to the exclusion of alternative ways of studying differences in human behavior, it cannot be defended on the grounds that it is nothing more than an attempt to gain some worthwhile knowledge about the world. If race-IQ connections are doggedly pursued while geographic and other potential covariates go largely unexplored, then we can demand an answer to the question: *why race*? If there is no satisfactory answer to this question, then there are *prima facie* grounds for moral disapproval of this institution that goes out of its way to gather scientific data that (irrespective of its truth value) stigmatizes and belittles groups that have been subjected to quite enough of that sort of thing already.[11] Moreover—and this bears

[11] This is not necessarily to impugn any particular researcher: as I have argued, the relationship of race to behavior is—in present circumstances at least—something worth studying and writing about. Hence, if a particular scientist decides that race is a more interesting subject to explore than age or geography, this certainly does not automatically imply ill-will or any other fault. But given the eminent salience of moral concerns about racial science, it does not seem unreasonable to expect those who would pursue such research to carefully weigh their intrinsic interest against the harms that may well result.

emphasizing, since the point is evidently lost on many racial scientists—we can mount this moral criticism *even though* we might concede that gathering data about race and behavior is, in the abstract, a legitimate scientific endeavor.[12]

To sum up the discussion so far, then: proposals like Antony's that seek to establish the inherently racist nature of racial science do not stand up to scrutiny. Relatedly, the suggestion, à la Rose, that we ought to pull the plug on all racial science forthwith fails to grapple with the constraints imposed by our present circumstances—the attempt to abandon racial science at this point presents us with the altogether unpromising business of trying to stuff various genies back into their bottles and cats back into their bags. It seems, then, that there is no alternative but to continue the scientific study of how racial categories (ontologically artificial as these categories may be) relate to important aspects of social behavior. And if some researchers pursuing such questions genuinely find that the evidence draws them to some "unpleasant truths" about racial differences, that must be acceptable. But for all that, we can and should point to the moral deficiencies inherent in any scientific community whose pursuit and publicization of "unpleasant truths" manages *only* to heap the unpleasantness upon one or a few already-stigmatized racial groups.

[12] In fairness, I should point out here that perhaps the most popular whipping boy for critics of racial science—Herrnstein and Murray's *The Bell Curve*—is not nearly as preoccupied with race as its detractors often insinuate. The vast majority of Herrnstein and Murray's argument regarding the relationship between genes, IQ, and success is built upon analyses of all-White samples, and the discussion of racial differences occupies a comparatively small portion of a large book. Murray and other supporters of the book have, reasonably I think, complained that popular treatment of the book has misrepresented it as a work primarily about racial differences. However, it also seems fair to point out, as does Randall Kennedy (1994), that the publicity campaigns that accompanied the book gave pride of place to Herrnstein and Murray's claims about racial differences (though of course we do not know what level of influence the authors had over such decisions), and the controversality of these claims in particular plausibly explains much of the book's commercial success.

To conclude this section: I have established that while it is too hasty to condemn all racial science as *ipso facto* racist, and while it is naïve to think we can abandon racial science cold-turkey, there may still be grounds for moral condemnation of much racial science, even *without* engaging in detailed analysis of its claims and methodologies. My claim is that it is sufficient for moral condemnation that racial science privileges—*without scientific or moral justification*—the investigation of *racial* differences over other potential objects of study, and in so doing disproportionately and unjustifiably burdens members of racially stigmatized groups with additional layers of insult. Note, however, that establishing that there are grounds for moral censure of racial science does not yet answer the central question of this chapter, which is whether racial science is *racist*. It is to that question I now turn.

8.3 Theories of racism and their applicability to racial science

8.3.1 Racism as inferiorization or pernicious belief

Lawrence Blum (2002) suggests that a core type of racism is *inferiorization:* claiming or implying (in word or in deed) that some racial groups are inferior to others.[13] Thus, for Blum, "Blacks are intellectually inferior" is a paradigmatically racist proposition (2002, p. 21). The notion of racism as inferiorization comports well with much ordinary discourse about racism: expression of a belief that one racial group is inferior to another in any serious way is generally taken to be grounds for an accusation of racism. So, applying Blum's idea of racism as inferiorization would seem to count much of racial science as racist right off the bat: indeed, the

[13] Blum's other type of racism is "racial antipathy," which will be discussed below.

distinguishing feature of the sort of scientific claims I'm here considering is that they make claims about the superiority and inferiority of various racial groups along various important psychological dimensions. But although Blum's proposal works reasonably well as an account of what many ordinary folks mean by "racism," it is not particularly helpful for settling questions about whether racial science is racist.

There are two related problems with using Blum's inferiorization account as a basis for charging racial science with racism. The first is that this move appears to be question-begging in the present context. The very reason why racial science is controversial is that it so frequently makes claims about the relative average endowments of members of different racial groups with regard to important psychological characteristics; in effect, assignation of superior and inferior positions in a racial hierarchy.[14] And yet it is the position of the racial scientists that their research enterprises are *not* racist. Thus, insofar as conducting such research and disseminating the results requires making claims that certain racial groups are, on average, inferior, holding that inferiorization is *ipso facto* racism leaves the racial scientist with no hope of avoiding the charge (except by halting the research or concealing its results). In other words, practitioners of racial science simply deny that engaging in inferiorization amounts to racism.

The second problem for Blum's suggestion is that it is not clear whether beliefs or statements can be racist if they are *true*, or at least *justified*. For racism (again, at least as the term

[14] One might deny that the claims of racial science genuinely inferiorize groups. Levin (1997b), for example, argues at considerable length that because there is no objective, universal standard by which to measure human characteristics, saying that Blacks are less intelligent than Whites (or, as Levin also holds, that Blacks are less motivated, less autonomous, and less altruistic than Whites) does not amount to saying that Blacks are "inferior." I see no particular reason to disagree with Levin about this: "superior" and "inferior" are, plausibly, inherently evaluative notions, which are not necessarily entailed by claims about greater or lesser degrees of intelligence—strictly speaking such claims are evaluatively neutral. However, conceding this point does not render Blum's account any less plausible. It seems sufficient for Blum's theory that most people *treat* claims about differing levels of intelligence (and other traits) as making evaluative—*inferiorizing*—judgments.

is normally used) is necessarily a moral fault, and, plausibly, it cannot be a moral fault to believe what one has good evidence for believing. As Levin puts it: "to say that blacks are genetically less intelligent and more impulsive than whites is not racist, since racism is by definition *bad*, and facts, however unwelcome, are morally neutral" (1996, p. 306). So this might seem to settle the matter: insofar as racial science makes *factual* assertions about racial inferiority, these claims are not subject to the moral evaluation—much less moral criticism—that would be required to sustain charges of racism.

There are two nuances here, however, that might re-open to possibility of directing moral criticism at the inferiorizing statements put forth by racial science. First, there is the matter of the empirical justification for these claims. Clearly, beliefs or statements that are woefully unjustified will not be immune from moral critique in the same way that well-supported ones are. That is, we might amend Blum's proposal and suggest that a belief might be racist if *both* of the following conditions are satisfied: (a) it inferiorizes a particular group, and (b) no reasonable person could believe it. Thus, a belief that Jews ritually drink the blood of infants cannot be defended on the grounds that it is a reasonable belief about an empirical matter, and so it is, plausibly, a racist belief. And indeed, some have taken the claims of racial science to be so dramatically deficient in empirical support that belief in them is hopelessly irrational (e.g., Dummett, 2004).

This raises the rather tricky issue of how to distinguish between rational and irrational belief. For just as there is disagreement (both *among* racial scientists and *between* these scientists and their critics) about whether the available evidence supports particular empirical claims about group differences, there is also *meta*-disagreement about the range of the "reasonable" positions to take on what conclusions the evidence supports and it is often not obvious how to settle such

disagreements (see Warfield & Feldman, 2010). So it seems quite possible that very few, if any, claims in racial science will turn out to be like the claim of Jewish neonate hematophagy, *so obviously* without empirical support that *no rational person* could *possibly* believe them (research in racial science is, after all, generally published in peer-reviewed scientific journals and academic presses). It is therefore by no means obvious that legitimate charges of racism can be premised on the indisputable irrationality of endorsing the conclusions of racial science.[15] Note, however, that the arguments just presented only establish that it is not racist to *believe* claims of racial science (if that is the belief one takes to be best supported by the evidence); it remains possible that making *public pronouncements* (whether in general or in particular contexts or particular ways) that endorse such propositions is racist or otherwise morally censurable.

Second, we might of course mount a moral critique of certain beliefs or claims if it is the case that prejudicial attitudes play a *causal role* in one's coming to adopt such beliefs (for example, a person is drawn to accept claims about lower Black IQ in order to justify her irrational dislike for Blacks). Many have also made this claim about racial science, its practitioners, as well as non-scientists who publicly and privately endorse the claims of racial science (Richardson, 1984; Sarkar, 1998). That is, some have claimed that those who conduct experiments and analyses that purportedly show Blacks to be inferior are motivated by a dislike for Blacks, and are thereby led to design experiments that will show Blacks to be inferior, and they subject their results to less

[15] Of course, a full consideration of this issue would require examining in detail the empirical bases for a range of racial science claims. An interesting test case is psychologist Kevin MacDonald's claims that Jews have an innate disposition towards "ethnocentricity" that drives them to practice in-group eugenics and to organize to promote Jewish interests against those of other groups—often through surreptitious means. The case that MacDonald's claims are racist (on grounds that they are so absurd that belief in them is morally censurable) is, plausibly, much stronger than the equivalent case regarding claims that the Black-White IQ gap is partly genetic in origin (which has traditionally enjoyed mainstream support among psychometricians—see Gottfredson, 1997; Snyderman & Rothman, 1987, 1988).

critical scrutiny than they would if they were found to support conclusions which the researchers were not already motivated to believe. In that case, the adoption of the inferiorizing beliefs would be morally censurable, since they could be causally traced to a censurable attitude (i.e., a prejudicial dislike for Black people).

However, establishing that a researcher or a consumer of racial science has such prejudicial attitudes, and that *these attitudes* are the reason why they have adopted this belief—rather than a genuine assessment of the available evidence—requires a very high evidential standard, one that, I submit, observers not intimately familiar with the persons in question will rarely be in a position to meet. Thus, this is not a promising avenue for answering our central question of whether there are grounds for *public* attributions of racism to racial science. This is not to deny that to the extent that it is in fact *true* of an individual that her endorsement of racial differences is (in part) motivated by immoral attitudes, this is a moral defect and she is worthy of moral criticism; my point is that such assumptions about the workings of an individual's psychology will rarely be legitimate grounds for public denunciations of racial science (or racial scientists) as racist.

8.3.2 Racism as an institution

The term "institutional racism" (made popular by Carmichael & Hamilton, 1967) generally seeks to draw attention to the ways in which broader social structures can systematically disadvantage particular groups, even in the absence of overt and deliberate acts of interpersonal racism (Headley, 2000).[16] Let us say that an institution is racist if it systematically disadvantages

[16] Although, again, my aim is not to adjudicate among competing theories of racism, I will note that I am sympathetic with what appears to be the majority opinion among philosophers, namely that while the notion of institutional racism has been useful in refocusing discussion away from (thankfully, increasingly rare) instances of outright

members of particular racialized groups, i.e., leaves those groups worse off than they would be if the institution were absent or structured differently. Racial science is plausibly a racist institution in this sense, since widespread publicization of claims that Blacks have been "scientifically demonstrated" to be, on average, less intelligent, more dangerous, and less dedicated parents than Whites will almost certainly make Blacks worse off in a number of ways, as I argued at length in Section 6.2.3. Thus, it seems there may be good grounds for classifying racial science as a racist institution.

However, as we saw earlier, the normative status of these outcomes (inequality, licensed discrimination, and so forth) is up for debate, and some would maintain that nothing is amiss with any of these results—they are simply signs of a well-functioning and economically efficient meritocracy. And if this is so, then the case for racial science as a racist institution is significantly undermined, since it is now no longer clear that there is anything *wrong* with racial science after all. In other words, it would seem that undergirding the theory that a racist institution is one that makes certain groups worse off is an assumption that those groups do not *deserve* to be worse off, or that their being worse off is in some way an *unjust* state of affairs. But if certain hereditarians are right in both their empirical and normative stances, this condition does not hold in the "hereditarian society."

Of course, I reject the premise that these outcomes are morally acceptable, but the key point for present purposes is that characterizing the effects of racial science as morally bad (which is necessary in order to classify it as a racist institution) now requires taking on a number of

racial bigotry, the notion that racism is fundamentally about relations of power does not serve well as a *general* analysis of racism (Blum, 2002; Garcia, 1997; Taylor, 2004).

substantive empirical and/or normative assumptions. That is, in order to characterize racial science as not just something that makes people of color worse off, but is also *morally condemnable* on those grounds, one must take on significant empirical assumptions and/or contentious theoretical commitments about notions like *desert* or *merit,* about the value of economic efficiency as compared to other social goods, and so forth. Thus, it will be very difficult to land an accusation of institutional racism against racial science without begging the question about some key normative issues. In order to sustain the charge, one would need to engage in a protracted defense of rather specific theses concerning much broader ethical questions. So, as with *racism as pernicious belief*, this line does not appear particularly promising.

8.3.3 Racism as ill-will or disregard

A final view is one that divorces racism from the reasonableness of empirical propositions about racial differences or the racially disparate consequences of social institutions. This account—defended, in various forms, by Jorge Garcia (1997, 1999), Lawrence Blum (2002), Paul Taylor (2004), and Joshua Glasgow (2009)—characterizes racism as a kind of a *stance* or an *attitude* towards people of different races. The account is most associated with Garcia, so I will focus on his formulation here. According to Garcia, the "heart" of racism is in the possession of *ill-will* or moral *disregard* towards the well-being of particular groups of people. Garcia proposes that we conceive of racism as:

> fundamentally a vicious kind of racially based disregard for the welfare of certain people. In its central and most vicious form, it is a hatred, ill-will, directed against a person or persons on account of their assigned race. In a derivative form, one is a racist when one either does not care at all or does not care enough (i.e., as much as morality requires) or does not care in the right ways about people assigned to a

> certain racial group, where this regard is based on racial classification (Garcia, 1999, pp. 399–400, emphasis added).17

Because, again, I am mainly concerned with grounds for *public* assertions about the racist nature of racial science, I do not think we will get much mileage out of the idea of *ill-will*, as it would be difficult to establish that someone conducting racial science actually feels *antipathy* towards or *wishes harm* upon a particular group, even while he might fiercely maintain that group's genetic inferiority.[18] However, it seems to me that at least some racial science is plausibly classified as racist under the notion of *indifference* or *disregard*, as I will now argue.

I am inclined to argue that the sort of indifference or disregard of which Garcia speaks is evident in much of racial science. This indifference takes the form of a *resignation* to the prospect of racial inequality discussed at length in Chapter 6. Recall that when racial scientists argue for a "genetic" basis for group differences, they are evincing a particular explanatory framework, one that gives us little reason to continue inquiry into the mechanistic details of development that may lead to inequality-rectifying interventions. And this privileging of genetic causes, in turn, renders permanent racial inequality a seemingly unavoidable outcome—an outcome we must, the racial scientist insists, accept with equanimity.

While I am wary of imputing any particular attitude or motivation to any particular scientist, some examples are necessary to elucidate the sort of statements I have in mind. Consider the following statement from Rushton and Jensen at the end of a long article arguing for

---

[17] Note that this conception of racism has the advantage of potentially subsuming another kind of case we have encountered, namely beliefs about racial inferiority that are motivated by ill-will. In such cases, the ill-will/disregard theory would hold, the person is guilty of holding a racist attitude, but it is not the *belief* that is racist, but rather the ill-will which is a necessary condition for the formation of the belief.

[18] The matter may be quite different, of course, when it comes to laypersons who are *consumers* of racial science, and who use the findings as a tool to heap denigration upon stigmatized groups—some of whom this author has had the unfortunate experience of encountering in the nether regions of the internet.

genetically-based IQ differences between racial groups: "Ultimately, the public *must accept* the pragmatic reality that some [racial] groups will be overrepresented and other groups underrepresented in various socially valued outcomes" (2005, p. 283, emphasis added). Similarly, recall Jensen's statement (discussed in Chapter 6) that if focusing only on individual merit—without regard for the status of groups—leads to disparate group outcomes, "so be it."

I would like to suggest, in contrast to the disposition these thinkers exhibit, that proper concern and regard for disadvantaged racial groups should precipitate an *extreme reluctance* to accept conditions of racial stratification. Indeed, the mere suggestion that the current and long-standing regime of racial inequality will continue indefinitely should, it seems to me, occasion profound sorrow and empathy for the plight of those who would be consigned to the social ladder's lower rungs. And with this regret should come a fierce and dogged commitment to finding viable alternatives to the racially stratified society. (I have laid out one set of strategies for pursuing such alternatives in the previous chapter—strategies which, as we have seen, hereditarians reject as detached from science's proper concern with "describing the real world.") But to react to the possibility of permanent racial inequality by shrugging one's shoulders and pronouncing "so be it," or to counsel levelheaded detachment as we accept this "unpleasant fact," is, I submit, to exhibit insufficient regard for the welfare of marginalized racial groups.

This is not to say that racial scientists are entirely uncaring or devoid of concern for *individuals* in socioeconomically disadvantaged positions. It is just that their insensitivity to the moral salience of the status of *groups*—which I demonstrated in Chapter 6—evinces an attitude that falls short of the moral regard one *ought* to have for members of historically-stigmatized groups. Again, Glenn Loury's thoughts are helpful here, as he elegantly captures the moral

inadequacy of responding with indifference to circumstances in which there is wide group disparity, but where *individual* injustices are absent:

> What does…abstract individualism…suggest that we do now? Throw up our hands? Declare that no questions of justice are raised? Scratch our heads and say that we don't quite know what to do? Too bad, we lament, but…*There is, I believe, a gaping hole in any normative framework that can provide us with no better answers than those.* (Loury, 2002, p. 124, emphasis added)

To put the matter another way: I argued at length in Chapter 6 that there are significant moral costs to disseminating claims about "innate" racial inferiority—even if those claims are correct as measured against one particular explanatory framework. One need not buy any particularly controversial assumptions in order to acknowledge this much. Although hereditarians are typically reluctant to address the issue, it seems impossible that they could deny that public claims about Black intellectual inferiority are hurtful to Black persons. These are, if you will, eggs the hereditarian believes must be broken in pursuit of the omelet of Truth. And perhaps they are correct that at a certain point we must accept a certain degree of insult to Blacks' self-image as part of the cost of doing scientific business. And yet, if my arguments in Chapter 7 are correct—if there is a viable alternative explanatory framework that affords the possibility of intervening to *eliminate* racial differences—then there is a course of action we may yet take that will not require us to simply live with these costs. To reject this course of action—to reject the further pursuit of inquiry in favor of declaring the matter settled—is not merely to advance an empirical hypothesis, but to make a value judgment about the scale and significance of the moral costs involved—costs borne entirely by non-Whites.

I suspect this point is lost on many hereditarians, who speak as if the only relevant question is a straightforwardly empirical one: can environmental manipulations eliminate racial differences

or not? Having answered that question to their satisfaction, they are content to accept whatever follows. Interestingly, however, Levin acknowledges the possibility of gene-environment interactions (G×E) playing a role in racial IQ gaps, such that although heritability analyses might allow for the difference to be attributed to genetic differences under *current* environmental conditions, this would not be the case under other conditions. In short, Levin acknowledges the possibility of environmental conditions that would eliminate racial inequality. And yet, he *still* explicitly rejects the idea that society (or at least, White people) need put forth any effort to instantiate such environments:

> Whether or not there are environments in which the attainments of the races would be equal, if *in fact*, in the environment(s) in which blacks and whites have *actually* functioned, the extant race difference in attainment was caused by genetic factors rather than white misdeeds, this difference is not an injury, hence not an injury for which whites are to blame, hence not a condition whites are obliged to remedy. (Levin, 1997b, p. 261, emphasis in original)

It is difficult to read this in a way that does not suggest a certain degree of disregard for Black individuals, in the form of an utter complacency and indifference to the moral costs—again, costs shouldered entirely by non-Whites—of a racially stratified society, even when we *know* there is something we could do about it.

Before I close this discussion, allow me to appeal to one more of Loury's observations, one that, in addition to driving home the point about disregard, will reconnect our conversation to the concerns about hereditarian rhetoric that I raised at the beginning of this chapter (recall the quotes from Rushton and Templer). I have said that by discounting the psychic harm done to Black persons as a result of promulgating racial science—by and urging that we all do likewise—racial scientists exhibit *insufficient moral regard* for Black persons. But what justifies this claim—in

particular, how do we know what degree of moral concern is *sufficient*? Why suppose that accepting subordinate social status for people of color implies a unique disregard for their wellbeing? Well, for starters we might ask how we would approach the matter if it were not Black people but some other population that was the target of these scientific investigations. Loury offers the following trenchant comparison:

> The American population is aging, and it is known that intelligence declines as a person ages, after some point in the life cycle. It is a demographic certainty that there will be relatively more older people in the American population in the years to come, and it is a legal fact that laws against age discrimination have abrogated mandatory retirement. These things taken together imply, as a mathematical necessity, that the American workforce is going to be made "dumber" by those baby boomers who insist on staying in the workforce beyond one's prime years. Where can one read about the dire consequences of this development for the productivity of the American economy? Nowhere. Why not? The reason, I suggest, is that those older, soon-to-be-less-intelligent workers are our mothers and fathers. We are not about to set them to one side and engage in an elaborate discourse about their fitness. (Loury, 2002, pp. 85–86)[19]

To push Loury's point a bit further: imagine that a group of academics calling themselves "age realists" began publishing a series of books and articles boldly proclaiming that old people, as a matter of immutable biological principle, are just not as smart as younger people. Imagine further that these writings were bursting with the same rhetorical devices common in hereditarian works about race: it is time that we stop pretending that differences in cognitive ability between

[19] Of course, there are reasons to question Loury's syllogism: even if IQ does decline with age, the sort of intelligence measured by IQ tests may not be all that important for job success, or perhaps older workers might compensate in other ways—for example by bringing a wealth of experience to work-related activities. But notice—as I suspect Loury does—that these are precisely the sorts of moves that hereditarians reject when it comes to the relationship between race, IQ, and economic productivity and success. Whenever it is suggested that the Black-White IQ gap might not matter so much, because a high IQ is not necessary to succeed in the workforce or in life (e.g., Chomsky, 1972), the idea is immediately dismissed, and the status of IQ as the best predictor of employee productivity is reaffirmed (Gottfredson, 2000; Herrnstein & Murray, 1996; Herrnstein, 1976).

the young and old do not exist; we must swallow the "unpleasant truths" that science has delivered to us about the productivity of older people in the workforce; the elimination of mandatory retirement constitutes a wasteful sacrifice of economic prosperity and meritocratic principles upon the altar of political correctness; it's really too bad if old people's feelings are hurt by persistent messages that science has proved that they are getting dumber, but only greater harm can result from keeping silent about the cognitive deficiencies of seniors in the workforce; and so on.[20]

Loury's intuition—and mine—is that we would not stand for this. The reason, I think, is that to speak in such a way about "our 'dumb' moms and dads" (Loury, 2002, p. 86) would seem *disrespectful*. Even if there are intractable biological forces at work here, we might say, we owe our elders better. But note that there is no discernable reason why the older members of our society should deserve greater respect than the people of color.[21] Therefore, if the way that racial science frames the narrative around racial differences falls short, in any way, of what we would consider acceptable regarding the cognitive changes besetting our mothers and fathers, then we ought to conclude that racial science exhibits lesser regard for people of color than morality requires—and this, again, is at the core of Garcia's account of racism.

To conclude our discussion of *ill-will/disregard* and its relation to racial science: even if I am wrong about the particular cases considered above, it seems to me that the following overall approach is promising: *to the extent* that insufficient moral regard for some group(s) contributes

[20] By the way, Loury's estimation that there is "nowhere" one can read about the economic effects of cognitive decline in an aging workforce appears to be largely correct. An admittedly cursory search on Google Scholar yields at most a handful of related papers (seemingly only one—Skirbekk, 2004—that mentions productivity losses due to declining cognitive ability among the more senior elements of the workforce).

[21] We might of course afford older persons a certain *kind* of unique respect, in recognition of the wisdom of their lived experience. But here we are considering a different, more basic kind of respect—roughly, basic respect for Kantian persons.

to a theorist's (or anyone's) readiness to accept the moral costs of permanent racial inequality, then the adoption of a narrow hereditarian framework exhibits some degree of racism. I conclude, then, that while the first two popular conceptions of racism—racism as *pernicious belief* and as *oppressive institution*—lack the resources to count racial science as racist, this third option—racism as *disregard*—provides an understanding of racism according to which a fair amount of racial science is indeed racist, at least to some degree.

8.4 Conclusion: What to do?

At this point, it would be quite reasonable for a practitioner of racial science—or anyone, for that matter—to pose the question of how, exactly, we are to proceed if what I have said about the morally problematic character of racial science is correct. After all, I have conceded that scientific investigation into the causes of observed racial differences is permissible, and indeed necessary, for the time being at least.

I confess to being short on answers here, and this will need to be a topic for further research. The discussion above suggests that the morally deleterious features of racial science can be mitigated by undertaking these investigations while holding the proper *moral regard* and *respect* for those who stand to endure the most harm—something I think we can safely say present manifestations of racial science have failed adequately to do. But further work will be required to spell out exactly how we might put this general stance of respect and regard into practice. I hope to have at least established, however, that adopting this stance—and using it as a guiding star in navigating the treacherous waters of racial science—is a moral imperative.

REFERENCES

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20–29.

Anderson, E. (2004). Uses of value judgments in science: A general argument, with lessons from a case study of feminist research on divorce. *Hypatia*, *19*(1), 1–24.

Andreasen, R. O. (2004). The Cladistic Race Concept: A Defense. *Philosophy of Science*, *19*(3), 425–442.

Andreasen, R. O. (2006). A New Perspective on the Race Debate. In E. Sober (Ed.), *Conceptual Issues in Evolutionary Biology* (3rd ed., Vol. 2). Cambridge, MA: MIT.

Antony, L. M. (1993). Quine as feminist: The radical import of naturalized epistemology. *A Mind of One's Own*, 185–225.

Antony, L. M., & Hornstein, N. (2003). *Chomsky and his Critics*. John Wiley & Sons.

Appiah, K. A. (1996). Race, Culture, Identity: Misunderstood Connections. In K. A. Appiah & A. Gutmann (Eds.), *Color Conscious: The Political Morality of Race* (pp. 30–105).

Ariew, A. (1999). Innateness is canalization: In defense of a developmental account of innateness. In V. Hardcastle (Ed.), *Where biology meets psychology: Philosophical essays* (pp. 117–138). MIT Press Cambridge, MA.

Ariew, A. (2006). Innateness. In M. Matten & C. Stephens (Eds.), *Handbook of the Philosophy of Biology*.

Baillargeon, R. (2002). The acquisition of physical knowledge in infancy: A summary in eight lessons. *Blackwell Handbook of Childhood Cognitive Development*, *1*, 46–83.

Bamshad, M. J., Wooding, S., Watkins, W. S., Ostler, C. T., Batzer, M. a, & Jorde, L. B. (2003). Human population genetic structure and inference of group membership. *American Journal of Human Genetics*, *72*(3), 578–89. doi:10.1086/368061

Barros, D. B. (2013). Negative causation in causal and mechanistic explanation. *Synthese*, *190*(3), 449–469.

Beebee, H. (2004). Causing and nothingness. *Causation and Counterfactuals*, 291–308.

Bertrand, M., & Mullainathan, S. (2003). *Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination*. National Bureau of Economic Research.

Birch, J. (2009). Irretrievably confused? Innateness in explanatory context. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *40*(4), 296–301.

Block, N. (1995). How heritability misleads about race. *Cognition*, *56*(2), 99–128.

Block, N., & Dworkin, G. (1976). *The IQ controversy: Critical readings.* Pantheon.

Block, W. (2010). *The Case for Discrimination*. Ludwig von Mises Institute.

Blum, L. (2002). *"I'm Not a Racist, But. . ." The Moral Quandary of Race* (p. 259). Ithaca, NY: Cornell University Press.

Blum, L. (2004). What do Accounts of "Racism" do? In M. Levine & T. Pataki (Eds.), *Racism in mind* (pp. 56–77). Ithaca, NY: Cornell University Press.

Bolnik, D. (2008). Individual Ancestry Inference and the Reification of Race as a Biological Phenomenon. In B. Koenig, S. Lee, & S. Richardson (Eds.), *Revisiting Race in a Genomic Age*. Piscataway, NJ: Rutgers University Press.

Boxill, B. R. (1978). The morality of preferential hiring. *Philosophy & Public Affairs*, 246–268.

Boxill, B. R. (1992). *Blacks and social justice*. Rowman & Littlefield.

Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies*, *61*(1), 127–148.

Buschkuehl, M., & Jaeggi, S. M. (2010). Improving intelligence: A literature review. *Swiss Medical Weekly*, *140*(19-20), 266–272.

Bushneil, I. W. R., Sai, F., & Mullin, J. T. (1989). Neonatal recognition of the mother's face. *British Journal of Developmental Psychology*, *7*(1), 3–15.

Carey, S., & Spelke, E. (1996). Science and core knowledge. *Philosophy of Science*, 515–533.

Carmichael, S., & Hamilton, C. V. (1967). *Black power: The politics of liberation in America*. Random House LLC.

Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., … Poulton, R. (2002). Role of genotype in the cycle of violence in maltreated children. *Science*, *297*(5582), 851–854.

Caspi, A., Moffitt, T. E., Cannon, M., McClay, J., Murray, R., Harrington, H., … Braithwaite, A. (2005). Moderation of the effect of adolescent-onset cannabis use on adult psychosis by a functional polymorphism in the catechol-O-methyltransferase gene: longitudinal evidence of a gene X environment interaction. *Biological Psychiatry*, *57*(10), 1117–1127.

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., … Braithwaite, A. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science Signaling*, *301*(5631), 386.

Cavalli-Sforza, L. L., Menozzi, P., & Piazza, A. (1994). *The History and Geography of Human Genes*. Princeton, NJ: Princeton University Press.

Ceci, S. J., & Liker, J. K. (1986). A day at the races: A study of IQ, expertise, and cognitive complexity. *Journal of Experimental Psychology: General*, *115*(3), 255.

Centerwall, S. A., & Centerwall, W. R. (2000). The discovery of phenylketonuria: the story of a young couple, two retarded children, and a scientist. *Pediatrics*, *105*(1), 89–103.

Chabris, C. F., Hebert, B. M., Benjamin, D. J., Beauchamp, J., Cesarini, D., van der Loos, M., … Atwood, C. S. (2012). Most reported genetic associations with general intelligence are probably false positives. *Psychological Science*, *23*(11), 1314–1323.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press.

Chomsky, N. (1972). The Fallacy of Richard Herrnstein's IQ. *Social Policy*, *3*(1), 19–25.

Chomsky, N. (1988). *Language and Problems of Knowledge*. Cambridge, MA: MIT.

Chomsky, N. (1990). On the Nature, Use and Acquisition of Language. In W. Lycan (Ed.), *Mind and Cognition: A Reader*. Cambridge, MA: Blackwell.

Churchland, P. M. (1996). *The engine of reason, the seat of the soul: a philosophical journey into the brain*. The MIT Press.

Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by g. *Intelligence*, *32*(3), 277–296.

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*(12), 547–552.

Cosmides, L., & Tooby, J. (1997). Evolutionary psychology: A primer. *Url: Http://www. Psych. Ucsb. Edu/research/cep/primer. Htlm*.

Cowie, F. (1998). *What's Within? Nativism Reconsidered*. Oxford University Press.

Crain, S., & Thornton, R. (1988). Acquisition of wanna contraction. In *Annual Meeting of the Linguistic Society of America, New Orleans, LA*.

Craver, C. F. (2007). *Explaining the brain*. Oxford University Press Oxford.

Cronshaw, S. F., Hamilton, L. K., Onyura, B. R., & Winston, A. S. (2006). Case for Non-Biased Intelligence Testing Against Black Africans Has Not Been Made: A Comment on Rushton, Skuy, and Bons (2004). *International Journal of Selection and Assessment, 14*(3), 278–287.

Cummins, R. (1975). Functional Analysis. *Journal of Philosophy*, *72*, 741–765. doi:10.2307/2024640

Darden, L. (2013). Mechanisms versus causes in biology and medicine. In *Mechanism and Causality in Biology and Economics* (pp. 19–34). Springer.

Davidson, D. (1967). Causal relations. *The Journal of Philosophy*.

De Regt, H. W. (2009). The epistemic value of understanding. *Philosophy of Science*, *76*(5), 585–597.

De Regt, H. W., & Dieks, D. (2005). A contextual approach to scientific understanding. *Synthese*, *144*(1), 137–170.

Devitt, M. (2008). *Ignorance of Language*. Oxford University Press.

Dickens, W. T., & Flynn, J. R. (2006). Black Americans reduce the racial IQ gap evidence from standardization samples. *Psychological Science*, *17*(10), 913–920.

Dowe, P. (1992). Wesley Salmon's process theory of causality and the conserved quantity theory. *Philosophy of Science*, 195–216.

Dowe, P. (2000). *Physical Causality*. Cambridge, UK: Cambridge University Press.

Dretske, F. (1972). Contrastive statements. *The Philosophical Review*, 411–437.

Dummett, M. (2004). The Nature of Racism. In M. P. Levine & T. Pataki (Eds.), *Racism in Mind*. Ithaca, NY: Cornell University Press.

Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(24), 13763–8. doi:10.1073/pnas.231499798

Edwards, A. W. F. (2003). Human genetic diversity: Lewontin's fallacy. *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology*, *25*(8), 798–801. doi:10.1002/bies.10315

Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S. L., Wiebe, V., Kitano, T., … Pääbo, S. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. *Nature*, *418*(6900), 869–872.

Enns, G. M., Koch, R., Brumm, V., Blakely, E., Suter, R., & Jurecki, E. (2010). Suboptimal outcomes in patients with PKU treated early with diet alone: Revisiting the Evidence. *Molecular Genetics and Metabolism*, *101*(2), 99–109.

Entine, J. (2008). *Taboo: Why Black athletes dominate sports and why we're afraid to talk about it*. PublicAffairs.

Epstein, R. A. (1994). Standing Firm, on Forbidden Grounds. *San Diego L. Rev.*, *31*, 1.

Evans, G. W., & Schamberg, M. A. (2009). Childhood poverty, chronic stress, and adult working memory. *Proceedings of the National Academy of Sciences*, *106*(16), 6545–6549.

Falconer, D. S. (1989). *Introduction to Quantitative Genetics* (3rd ed.). New York: John Wiley & Sons.

Faucher, L., & Machery, E. (2009). Racism: Against Jorge Garcia's Moral and Psychological Monism. *Philosophy of the Social Sciences*, *39*(1), 41–62.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314.

Fisher, P. A., Stoolmiller, M., Gunnar, M. R., & Burraston, B. O. (2007). Effects of a therapeutic intervention for foster preschoolers on diurnal cortisol activity. *Psychoneuroendocrinology*, *32*(8), 892–905.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, *95*(1), 29.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, *101*(2), 171.

Flynn, J. R. (1999). Searching for justice: the discovery of IQ gains over time. *American Psychologist*, *54*(1), 5.

Flynn, J. R. (2007). *What is intelligence?: Beyond the Flynn effect*. Cambridge University Press.

Fodor, J. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.

Fodor, J. (2001). Doing without what's within: Fiona Cowie's critique of nativism. *Mind*, *110*(437), 99–148.

Foraker, S., Regier, T., Khetarpal, N., Perfors, A., & Tenenbaum, J. (2009). Indirect evidence and the poverty of the stimulus: The case of anaphoric one. *Cognitive Science*, *33*(2), 287–300.

Fryer, R. G., & Levitt, S. D. (2004). Understanding the black-white test score gap in the first two years of school. *Review of Economics and Statistics*, *86*(2), 447–464.

Gannett, L. (1999). What's in a cause?: The pragmatic dimensions of genetic explanations. *Biology and Philosophy*, *14*(3), 349–373.

Garcia, J. L. A. (1997). Current conceptions of racism: A critical examination of some recent social philosophy. *Journal of Social Philosophy*, *28*(2), 5–42.

Garcia, J. L. A. (1999). The heart of racism. In L. Harris (Ed.), *Racism (Key Concepts in Critical Theory)* (Vol. 27, pp. 399–434). Amherst, NY: Humanity Books.

Garfinkel, A. (1981). *Forms of Explanation: Rethinking the Questions in Social Theory* (p. 186). Yale University Press.

Genovese, E. D. (1995). Living with inequality. *National Review*, *46*, 44–46.

Gil-White, F. J. (2001). Are ethnic groups biological "species" to the human brain? *Current Anthropology*, *42*(4), 515–553.

Glasgow, J. (2003). The New Biology of Race. *Journal of Philosophy*, *100*(9), 456–474.

Glasgow, J. (2009). Racism as Disrespect. *Ethics*, *120*(1), 64–93.

Glazer, N. (1994). The lying game. *The New Republic*, *16*, 94–95.

Glennan, S. S. (1996). Mechanisms and the nature of causation. *Erkenntnis*, *44*(1), 49–71.

Glymour, C. (1998). What went wrong? Reflections on science by observation and The Bell Curve. *Philosophy of Science*, 1–32.

Godfrey-Smith, P. (2007). Innateness and Genetic Information. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The Innate Mind, Volume 3: Foundations and the Future*. Oxford: Oxford University Press.

Gopnik, A. (2000). Explanation as orgasm and the drive for causal knowledge: The function, evolution, and phenomenology of the theory formation system. In F. C. Keil & R. A. Wilson (Eds.), *Explanation and Cognition*. The MIT Press.

Goth, A. (2001). Innate predator-recognition in Australian brush-turkey (Alectura lathami, Megapodiidae) hatchlings. *Behaviour*, *138*(1), 117.

Gottfredson, L. S. (1988). Reconsidering fairness: A matter of social and ethical priorities. *Journal of Vocational Behavior*. doi:10.1016/0001-8791(88)90041-3

Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, *24*(1), 13–23.

Gottfredson, L. S. (2000). Skills gaps, not tests, make racial proportionality impossible. *Psychology, Public Policy, and Law*, *6*(1), 129.

Gottfredson, L. S. (2005a). Suppressing intelligence research: Hurting those we intend to help. *Destructive Trends in Mental Health: The Well-Intentioned Path to Harm*, 155–186.

Gottfredson, L. S. (2005b). What if the hereditarian hypothesis is true? *Psychology, Public Policy, and Law*, *11*(2), 311–319. doi:10.1037/1076-8971.11.2.311

Gottlieb, G. (1991). Experiential canalization of behavioral development: Theory. *Developmental Psychology*, *27*(1), 4.

Gottlieb, G. (1995). Some conceptual deficiencies in "developmental"behavior genetics. *Human Development*, *38*(3), 131–141.

Gould, S. J. (1983). *The Mismeasure of Man*. New York: Norton.

Graves, J. L. (2002). The misuse of life-history theory: JP Rushton and the pseudoscience of racial hierarchy. *Race and Intelligence: Separating Science from Myth*, 57–94.

Gray, R. (1992). Death of the gene: developmental systems strike back. In *Trees of life* (pp. 165–209). Springer.

Greenspan, R. (2001). The flexible genome. *Nature Reviews Genetics*, *2*(5), 383–387.

Griffiths, P. (2002). What is innateness? *The Monist*, *85*(1), 70–85.

Griffiths, P., & Machery, E. (2008). Innateness, canalization, and "biologicizing the mind." *Philosophical Psychology*, *21*(3), 397–414.

Griffiths, P., Machery, E., & Linquist, S. (2009). The vernacular concept of innateness. *Mind & Language*, *24*(5), 605–630.

Griffiths, P., & Tabery, J. (2008). Behavioral genetics and development: Historical and conceptual causes of controversy. *New Ideas in Psychology*, *26*(3), 332–352.

Gross, S., & Rey, G. (2012). Innateness. In E. Margolis, R. Samuels, & S. Stich (Eds.), *Oxford Handbook of Philosophy of Cognitive Science*. Oxford: Oxford University Press.

Hacking, I. (1999). *The Social Construction of What?*. Harvard University Press.

Hackman, D. A., Farah, M. J., & Meaney, M. J. (2010). Socioeconomic status and the brain: mechanistic insights from human and animal research. *Nature Reviews Neuroscience*, *11*(9), 651–659.

Hájek, A. (2007). The reference class problem is your problem too. In *Synthese* (Vol. 156, pp. 563–585). doi:10.1007/s11229-006-9138-5

Hall, N. (2004). Two Concepts of Causation. In J. Collins, N. Hall, & L. A. Paul (Eds.), *Causation and Counterfactuals*. Cambridge, MA: MIT.

Harman, G. (1990). The intrinsic quality of experience. *Philosophical Perspectives*, *4*, 31–52.

Hartmann, H. (1958). *Ego Psychology and Problem of Adaption*. International Universities Press.

Haslanger, S. (2004). Oppressions: Racial and other. In M. Levine & T. Pataki (Eds.), *Racism in Mind*. Ithaca, NY: Cornell University Press.

Headley, C. (2000). Philosophical approaches to racism: A critique of the individualistic perspective. *Journal of Social Philosophy*, *31*(2), 223–257.

Hempel, C. G. (1960). Science and Human Values. In R. Spillar (Ed.), *Social Control in a Free Society*. Philadelphia: University of Pennsylvania Press.

Herrnstein, R. (1976). Whatever Happened to Vaudeville? A Reply to Professor Chomsky. In N. Block & G. Dworkin (Eds.), *The IQ controversy: Critical readings* (p. 299). New York: Random House.

Herrnstein, R. (1990). Still an American Dilemma. *Public Interest*, *98*, 3–17.

Herrnstein, R., & Murray, C. (1996). *The Bell Curve: Intelligence and Class Structure in American Life*. Free Press.

Hiddleston, E. (2005). Review of "Making Things Happen." *Philosophical Review*, *114*(4), 545–547.

Hirschfeld, L. A. (2012). Seven Myths of Race and the Young Child. *Du Bois Review: Social Science Research on Race*, *9*(01), 17–39.

Hochman, A. (2013). Do We Need a Device to Acquire Ethnic Concepts? *Philosophy of Science*, *80*(5), 994–1005.

Hsu, N. S., Novick, J. M., & Jaeggi, S. M. (2014). The development and malleability of executive control abilities. *Frontiers in Behavioral Neuroscience*, *8*, 221.

Hunt, E. (1997). Nature vs. nurture: The feeling of vuja de. In R. J. Sternberg & E. Grigorenko (Eds.), *Intelligence, Heredity, and Environment*. New York: Cambridge University Press.

Hunt, E., & Sternberg, R. J. (2006). Sorry, wrong numbers: An analysis of a study of a correlation between skin color and IQ. *Intelligence*, *34*(2), 131–137.

Idler-Acosta, J. (2005). *The very idea of hispanic identity*. University of Maryland, College Park.

Jacob, F., & Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, *3*(3), 318–356. doi:10.1016/S0022-2836(61)80072-7

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short-and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences*, *108*(25), 10081–10086.

Jencks, C., Smith, M., Acland, M., Bane, M., Cohen, D., Gintis, H., … Michelson, S. (1972). *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books.

Jensen, A. (1969). How much can we boost IQ and scholastic achievement. *Harvard Educational Review*, *39*(1), 1–123.

Jensen, A. (1974). *Educability and group differences*. Nature Publishing Group.

Jensen, A. (1996). The puzzle of nonshared environment. In R. J. Sternberg & E. Grigorenko (Eds.), *Intelligence, Heredity, and Environment*. New York: Cambridge University Press.

Jerne, N. K. (1985). The generative grammar of the immune system. *Bioscience Reports*, *5*(6), 439–451.

Jinks, J. L., & Fulker, D. W. (1970). Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of the human behavior. *Psychological Bulletin*, *73*(5), 311.

Jobling, I. (2004). Racial Heresies for the 21st Century. *American Renaissance*, *15*(4).

Judis, J. (1994). Taboo you. *The New Republic*, *211*, 18.

Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, *9*(2), 131–155.

Kalinowski, S. T. (2011). The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. *Heredity*, *106*(4), 625–32. doi:10.1038/hdy.2010.95

Kaplan, J. M., & Winther, R. G. (2012). Prisoners of Abstraction? The Theory and Measure of Genetic Variation, and the Very Concept of "Race." *Biological Theory*, 1–12. doi:10.1007/s13752-012-0048-0

Karbach, J., & Unger, K. (2014). Executive control training from middle childhood to adolescence. *Frontiers in Psychology*, *5*.

Keller, E. F. (2010). *The mirage of a space between nature and nurture*. Duke University Press.

Kennedy, R. (1994). The phony war. *The New Republic*, *211*, 1.

Kershnar, S. (2000). Intrinsic Moral Value and Racial Differences. *Public Affairs Quarterly*, *14*(3), 205–224.

Kitcher, P. (2001). *Science, Truth, and Democracy*. New York: Oxford University Press.

Klein, K., & Boals, A. (2001). The relationship of life event stress and working memory capacity. *Applied Cognitive Psychology*, *15*(5), 565–579.

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences*, *98*(26), 15387–15392.

Kvaale, E. P., Gottdiener, W. H., & Haslam, N. (2013). Biogenetic explanations and stigma: A meta-analytic review of associations among laypeople. *Social Science & Medicine*, *96*, 95–103.

Lai, C. S. L., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., & Monaco, A. P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*, *413*(6855), 519–523.

Landau, B., & Gleitman, L. R. (2009). *Language and experience: Evidence from the blind child* (Vol. 8). Harvard University Press.

Laurence, S., & Margolis, E. (2001). The poverty of the stimulus argument. *The British Journal for the Philosophy of Science*, *52*(2), 217–276.

Levin, M. (1996). Why Race Matters: A Preview. *Journal of Libertarian Studies*, *12*, 287–312.

Levin, M. (1997a). Natural subordination, Aristotle on. *Philosophy*, *72*(280), 241–257.

Levin, M. (1997b). *Why Race Matters: Race Differences and What They Mean* (p. 415). Praeger.

Levy, S. R., & Dweck, C. S. (1999). The impact of children's static versus dynamic conceptions of people on stereotype formation. *Child Development*, *70*(5), 1163–1180.

Lewis, D. (1973). Causation. *Journal of Philosophy1*, *70*(17), 556–567.

Lewontin, R. C. (1972). The Apportionment of Human Diversity. *Evolutionary Biology*, *6*, 381–398.

Lewontin, R. C. (1974). The analysis of variance and the analysis of causes. *American Journal of Human Genetics*, *26*(3), 400.

Lidz, J., Waxman, S., & Freedman, J. (2003). What infants know about syntax but couldn't have learned: experimental evidence for syntactic structure at 18 months. *Cognition*, *89*(3), 295–303.

Loehlin, J. C., Lindzey, G., & Spuhler, J. N. (1975). *Race differences in intelligence.* WH Freeman/Times Books/Henry Holt & Co.

Longino, H. E. (2001). What do we measure when we measure aggression? *Studies In History and Philosophy of Science Part A*, *32*(4), 685–704.

Longino, H. E. (2013). *Studying Human Behavior: How Scientists Investigate Aggression and Sexuality*. University of Chicago Press.

Loury, G. C. (1994). Self-Censorship in Public Discourse A Theory of "Political Correctness" and Related Phenomena. *Rationality and Society*, *6*(4), 428–461.

Loury, G. C. (1999). Why Should We Care About Racial Inequality per se. *Issues in Ethics*, *10*(2).

Loury, G. C. (2000). Who cares about racial inequality. *J. Soc. & Soc. Welfare*, *27*, 133.

Loury, G. C. (2002). *The Anatomy of Racial Inequality*. Cambridge, MA: Harvard University Press.

Lykken, D. (1998). The genetics of genius. In A. Steptoe (Ed.), *Genius and the Mind: Studies of Creativity and Temperament in the Historial Record*. Oxford University Press.

Lynch, M. P. (2004). *True to life: Why truth matters*. mit Press Cambridge, MA.

Lynn, R. (2013). Rushton's r–K life history theory of race differences in penis length and circumference examined in 113 populations. *Personality and Individual Differences2*, *55*(3), 261–266.

Lynn, R., & Vanhanen, T. (2002). *IQ and the wealth of nations*. Greenwood Publishing Group.

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, 1–25.

Machery, E., & Faucher, L. (2004). Why do we think racially? A critical journey in culture and evolution. *Categorization in Cognitive Science*.

MacLeod, E. L., & Ney, D. M. (2010). Nutritional management of phenylketonuria. *Annales Nestlé (English Ed.)*, *68*(2), 58–69.

Mameli, M. (2008). On innateness: The clutter hypothesis and the cluster hypothesis. *Journal of Philosophy*, *105*(12), 719–737.

Mameli, M., & Bateson, P. (2006). Innateness and the Sciences. *Biology & Philosophy*, *21*(2), 155–188. doi:10.1007/s10539-005-5144-0

Marks, J. (2010). Ten Facts about Human Genetic Variation. In M. Muehlenbein (Ed.), *Human Evolutionary Biology2* (pp. 265–276). Cambridge, UK: Cambridge University Press.

Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information, Henry Holt and Co. *Inc., New York, NY*.

Mays, V. M., Cochran, S. D., & Barnes, N. W. (2007). Race, race-based discrimination, and health outcomes among African Americans. *Annu. Rev. Psychol.*, *58*, 201–225.

McDonald, J. (2010). "Lookism": The Next Form of Illegal Discrimination. *Bloomberg Law Reports--Labor & Employment*, *4*(46).

Meaney, M. J. (2001a). Maternal care, gene expression, and the transmission of individual differences in stress reactivity across generations. *Annual Review of Neuroscience*, *24*(1), 1161–1192.

Meaney, M. J. (2001b). Nature, nurture, and the disunity of knowledge. *Annals of the New York Academy of Sciences*, *935*(1), 50–61.

Meaney, M. J. (2010). Epigenetics and the biological definition of gene X environment interactions. *Child Development*. doi:10.1111/j.1467-8624.2009.01381.x

Miele, F. (2002). *Intelligence, race, and genetics: Conversations with Arthur R. Jensen.* Westview Press.

Miele, F., & Sarich, V. (2005). *Race: The reality of human differences*. Westview Press.

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

Miller, P. B. (1998). The anatomy of scientific racism: racialist responses to Black athletic achievement. *Journal of Sport History*, *25*, 119–151.

Millikan, R. G. (1989). In defense of proper functions. *Philosophy of Science*, 288–302.

Mills, C. (1998). "But What Are You Really?" The Metaphysics of Race. In *Blackness Visible: Essays on Philosophy and Race* (pp. 41–66).

Mills, C. (2003). ``Heart''Attack: A Critique of Jorge Garcia's Volitional Conception of Racism. *The Journal of Ethics*, *7*(1), 29–62.

Mitchell, S. D. (2009). *Unsimple Truths: Science, Complexity, and Policy* (p. 160). University of Chicago Press.

Mitton, J. B. (1977). Genetic Differentiation of Races of Man as Judged by Single-Locus and Multilocus Analyses. *The American Naturalist*, *111*(978), 203–212.

Montague, A. (Ed.). (1964). *The Concept of Race*. New York: Free Press.

Moore, C. L. (1984). Maternal contributions to the development of masculine sexual behavior in laboratory rats. *Developmental Psychobiology*, *17*(4), 347–356.

Moore, C. L. (1992). The Role of Maternal Stimulation in the Development of Sexual Behavior and Its Neural Basis. *Annals of the New York Academy of Sciences*, *662*(1), 160–177.

Moore, D. S. (2008). Espousing interactions and fielding reactions: Addressing laypeople's beliefs about genetic determinism. *Philosophical Psychology*, *21*(3), 331–348.

Morange, M. (1994). *A History of Molecular Biology (M. Cobb, Trans.)*. Cambridge, MA: Harvard University Press.

Mosley, A. (2005). Affirmative Action. In A. Cohen & H. Wellman (Eds.), *Contemporary Debates in Applied Ethics*. Wiley-Blackwell.

Nelson, T. E., Biernat, M. R., & Manis, M. (1990). Everyday base rates (sex stereotypes): Potent and resilient. *Journal of Personality and Social Psychology*, *59*(4), 664.

Neville, H. J., Stevens, C., Pakulak, E., Bell, T. A., Fanning, J., Klein, S., & Isbell, E. (2013). Family-based training program improves brain function, cognition, and behavior in lower socioeconomic status preschoolers. *Proceedings of the National Academy of Sciences*, *110*(29), 12138–12143.

Nisbett, R. E. (2009). *Intelligence and how to get it: Why schools and cultures count*. WW Norton & Company.

Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., & Turkheimer, E. (2012). Intelligence: new findings and theoretical developments. *American Psychologist*, *67*(2), 130.

Northcott, R. (2006). Causal efficacy and the analysis of variance. *Biology and Philosophy*, *21*(2), 253–276.

Northcott, R. (2008). Causation and contrast classes. *Philosophical Studies*, *139*(1), 111–123.

Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., … Greenwald, A. G. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology*, *18*(1), 36–88.

Novembre, J., Johnson, T., Bryc, K., Kutalik, Z., Boyko, A. R., Auton, A., … Bustamante, C. D. (2008). Genes mirror geography within Europe. *Nature*, *456*(7218), 98–101. doi:10.1038/nature07331

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*(5719), 255–258.

Oyama, S. (2000). *The ontogeny of information: Developmental systems and evolution*. Duke university press.

Oyama, S., Griffiths, P., & Gray, R. D. (2003). *Cycles of contingency: Developmental systems and evolution*. Mit Press.

Pager, D., & Karafin, D. (2009). Bayesian bigot? Statistical discrimination, stereotypes, and employer decision making. *The Annals of the American Academy of Political and Social Science*, *621*(1), 70–93.

Pager, D., Western, B., & Bonikowski, B. (2009). Discrimination in a Low-Wage Labor Market A Field Experiment. *American Sociological Review*, *74*(5), 777–799.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference* (p. 384). Cambridge University Press.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, *12*(5), 729–738.

Plomin, R., & Bergeman, C. S. (1991). The nature of nurture: Genetic influence on "environmental" measures. *Behavioral and Brain Sciences*, *14*(03), 373–427.

Plomin, R., & DeFries, J. C. (1976). Heritability of IQ. *Science (New York, NY)*, *194*(4260), 10–12.

Quillian, L., & Pager, D. (2001). Black Neighbors, Higher Crime? The Role of Racial Stereotypes in Evaluations of Neighborhood Crime. *American Journal of Sociology*, *107*(3), 717–767.

Regier, T., & Gahl, S. (2004). Learning the unlearnable: The role of missing evidence. *Cognition*, *93*(2), 147–155.

Rey, G. (1997). *Contemporary Philosophy of Mind: A Contentiously Classical Approach* (p. 384). Wiley.

Richardson, R. C. (1984). Biology and ideology: The interpenetration of science and values. *Philosophy of Science*, 396–420.

Richardson, R. C. (2000). Review of M. Levin, Why Race Matters. *Ethics*, *110*(4), 847.

Risch, N., Burchard, E., Ziv, E., & Tang, H. (2002). Categorization of humans in biomedical research: genes, race and disease. *Genome Biol*.

Roberts, R. C. (1967). Some concepts and methods in quantitative genetics. In J. Hirsch (Ed.), *Behavior-genetic analysis.* (pp. 214–257). New York: McGraw Hill.

Rodgers, J. L. (1998). A critique of the Flynn Effect: massive IQ gains, methodological artifacts, or both? *Intelligence*, *26*(4), 337–356. doi:10.1016/S0160-2896(99)00004-5

Rose, S. (2009). Darwin 200: Should scientists study race and IQ? NO: Science and society do not benefit. *Nature*, *457*(7231), 786–788.

Rosenberg, J. (2011). What's so bad about racial inequality...or racial stratification? *Discriminations*. Retrieved from http://www.discriminations.us/2011/01/what's-so-bad-about-inequality-or-racial-stratification/

Rosenberg, N., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K. K., Zhivotovsky, L. a, & Feldman, M. W. (2002). Genetic structure of human populations. *Science*, *298*(5602), 2381–5. doi:10.1126/science.1078311

Rosenthal, R., & Jacobson, L. (1968). Pygmalion in the classroom. *The Urban Review*, *3*(1), 16–20.

Rushton, J. P. (1997). *Race, Evolution, and Behavior: A Life History Perspective* (2nd ed.). New Brunswick, NJ: Transaction Publishers.

Rushton, J. P. (2000). *Race, Evolution and Behavior: A Life History Perspective (Abridged Edition)*. Charles Darwin Research Institute.

Rushton, J. P., & Jensen, A. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law*, *11*(2), 235.

Rutter, M. (2006). Implications of resilience concepts for scientific understanding. *Annals of the New York Academy of Sciences*, *1094*(1), 1–12.

Sabban, E. L., & Kvetňanský, R. (2001). Stress-triggered activation of gene expression in catecholaminergic systems: dynamics of transcriptional events. *Trends in Neurosciences*, *24*(2), 91–98.

Sackett, P. R., Hardison, C. M., & Cullen, M. J. (2004). On interpreting stereotype threat as accounting for African American-White differences on cognitive tests. *American Psychologist*, *59*(1), 7.

Salmon, W. C. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.

Salmon, W. C. (1989). *4 Decades of Scientific Explanation*. University of Pittsburgh Press.

Samuels, R. (2002). Nativism in cognitive science. *Mind & Language*, *17*(3), 233–265.

Samuels, R. (2007). Is Innateness a Confused Concept. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The Innate Mind, Volume 3: Foundations and the Future*. Oxford: Oxford University Press.

Sandoval, J. (1979). The WISC-R and internal evidence of test bias with minority groups. *Journal of Consulting and Clinical Psychology*, *47*(5), 919.

Sarkar, S. (1998). *Genetics and reductionism*. Cambridge University Press.

Scarr, S. (1992). Developmental theories for the 1990s: Development and individual differences. *Child Development*, *63*(1), 1–19.

Scarr, S. (1995). Commentary. *Human Development*, *38*, 154–157.

Scarr, S., & Weinberg, R. A. (1983). The Minnesota adoption studies: Genetic differences and malleability. *Child Development*, 260–267.

Scholl, B. J., & Leslie, A. M. (1999). Modularity, development and "theory of mind." *Mind & Language*, *14*(1), 131–153.

Schoofs, D., Preuß, D., & Wolf, O. T. (2008). Psychosocial stress induces working memory impairments in an n-back paradigm. *Psychoneuroendocrinology*, *33*(5), 643–653.

Sesardic, N. (1993). Heritability and causality. *Philosophy of Science*, *60*(3), 396.

Sesardic, N. (2005). *Making sense of heritability*. Cambridge University Press.

Sesardic, N. (2010). Race: a social destruction of a biological concept. *Biology & Philosophy*, *25*(2), 143–162. doi:10.1007/s10539-009-9193-7

Shelby, T. (2002). Is racism in the "heart"? *Journal of Social Philosophy*, *33*(3), 411–420.

Shelby, T. (2005). *We who are dark: The philosophical foundations of black solidarity*. Harvard University Press.

Sheldon, J. P., Jayaratne, T. E., & Petty, E. M. (2007). White Americans' genetic explanations for a perceived race difference in athleticism: The relation to prejudice toward and stereotyping of Blacks. *Athletic Insight: The Online Journal of Sport Psychology*, *9*(3), 31–56.

Skirbekk, V. (2004). Age and individual productivity: A literature survey. *Vienna Yearbook of Population Research*, 133–153.

Smedley, A. (1999). *Race in North America: Origin and evolution of a worldview*. Basic Civitas Books.

Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences*, *11*(1), 1–74.

Snyderman, M., & Rothman, S. (1987). Survey of expert opinion on intelligence and aptitude testing. *American Psychologist*, *42*(2), 137.

Snyderman, M., & Rothman, S. (1988). *The IQ controversy, the media and public policy*. Transaction Publishers.

Sober, E. (1988). Apportioning causal responsibility. *The Journal of Philosophy*, *85*(6), 303–318.

Sober, E. (1998). Innate knowledge. In E. Craig (Ed.), *Routledge Encyclopedia of Philosophy* (pp. 794–797). Routledge.

Spelke, E. (1994). Initial knowledge: Six suggestions. *Cognition*, *50*(1), 431–445.

Spelke, E., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89–96.

Steele, C. M. (2010). *Whistling Vivaldi: And Other Clues to How Stereotypes Affect Us*. WW Norton & Company.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. CUP Archive.

Strevens, M. (2007). Review of Woodward, Making Things Happen*. *Philosophy and Phenomenological Research*, *74*(1), 233–249.

Sue, D. W., Capodilupo, C. M., & Holder, A. (2008). Racial microaggressions in the life experience of Black Americans. *Professional Psychology: Research and Practice*, *39*(3), 329.

Swain, C. M., & Nieli, R. (2003). *Contemporary voices of white nationalism in America*. Cambridge University Press.

Tabery, J. (2009). Making Sense of the Nature–Nurture Debate. *Biology and Philosophy*, *24*(5), 711–723.

Tang, H., Quertermous, T., Rodriguez, B., Kardia, S. L. R., Zhu, X., Brown, A., … Risch, N. J. (2005). Genetic structure, self-identified race/ethnicity, and confounding in case-control association studies. *American Journal of Human Genetics*, *76*(2), 268–75. doi:10.1086/427888

Taylor, P. C. (2004). *Race: A Philosophical Introduction*. Polity.

Te Nijenhuis, J., & van der Flier, H. (2013). Is the Flynn effect on g?: A meta-analysis. *Intelligence*, *41*(6), 802–807.

Templeton, A. R. (1999). Human Races: A Genetic and Evolutionary Perspective. *American Anthropologist*, *100*(3), 632–650.

Templeton, A. R. (2002). The Genetic and Evolutionary Significance of Human Races. In J. Fish (Ed.), *Race and Intelligence* (pp. 31–56). Malwah, NJ: Lawrence Erlbaum.

Tienda, M., & Mitchell, F. (2006). *Multiple origins, uncertain destinies: Hispanics and the American future*. National Academies Press.

Troxel, W. M., Matthews, K. A., Bromberger, J. T., & Sutton-Tyrrell, K. (2003). Chronic stress burden, discrimination, and subclinical carotid artery disease in African American and Caucasian women. *Health Psychology*, *22*(3), 300.

Tung, J., Barreiro, L. B., Johnson, Z. P., Hansen, K. D., Michopoulos, V., Toufexis, D., … Gilad, Y. (2012). Social environment is associated with gene regulatory variation in the rhesus macaque immune system. *Proceedings of the National Academy of Sciences*, *109*(17), 6490–6495.

Turner, R. J., & Avison, W. R. (2003). Status variations in stress exposure: Implications for the interpretation of research on race, socioeconomic status, and gender. *Journal of Health and Social Behavior*, 488–505.

Uhlmann, E. L., Brescoll, V. L., & Machery, E. (2010). The motives underlying stereotype-based discrimination against members of stigmatized groups. *Social Justice Research*, *23*(1), 1–16.

Van Calcar, S. C., MacLeod, E. L., Gleason, S. T., Etzel, M. R., Clayton, M. K., Wolff, J. A., & Ney, D. M. (2009). Improved nutritional management of phenylketonuria by using a diet containing glycomacropeptide compared with amino acids. *The American Journal of Clinical Nutrition*, *89*(4), 1068–1077.

Van Fraassen, B. C. (1980). *The scientific image*. Oxford University Press.

Waddington, C. H. (1957). *The strategy of the genes, 1957. George Allen, Londres*. London: George Allen.

Waddington, C. H. (1975). *The evolution of an evolutionist*. Edinburgh University Press Edinburgh.

Wahlsten, D. (1990). Insensitivity of the analysis of variance to heredity-environment interaction. *Behavioral and Brain Sciences*, *13*(01), 109–120.

Warfield, T., & Feldman, F. (2010). Disagreement. Oxford: Oxford University Press.

Waters, C. K. (2007). Causes that make a difference. *The Journal of Philosophy*, 551–579.

Watkins, K. E., Dronkers, N. F., & Vargha-Khadem, F. (2002). Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. *Brain*, *125*(3), 452–464.

Weinstein, R. S., Gregory, A., & Strambler, M. J. (2004). Intractable Self-Fulfilling Prophecies Fifty Years After Brown v. Board of Education. *American Psychologist*, *59*(6), 511.

Wendler, D. (1996). Innateness as an explanatory concept. *Biology and Philosophy*, *11*(1), 89–116.

West, M. J., & King, A. P. (1987). Settling nature and nurture into an ontogenetic niche. *Developmental Psychobiology*, *20*, 549–562. doi:10.1002/dev.420200508

Wicherts, J. M., Dolan, C. V, Carlson, J. S., & van der Maas, H. L. J. (2010). Raven's test performance of sub-Saharan Africans: Average performance, psychometric properties, and the Flynn Effect. *Learning and Individual Differences*, *20*(3), 135–151.

Wicherts, J. M., Dolan, C. V, Hessen, D. J., Oosterveld, P., Van Baal, G. C. M., Boomsma, D. I., & Span, M. M. (2004). Are intelligence tests measurement invariant over time? Investigating the nature of the Flynn effect. *Intelligence*, *32*(5), 509–537.

Wilson, C. (1998). Natural Domination: A Reply to Michael Levin. *Philosophy*, *73*(4), 573–592.

Witherspoon, D. J., Wooding, S., Rogers, a R., Marchani, E. E., Watkins, W. S., Batzer, M. a, & Jorde, L. B. (2007). Genetic similarities within and between human populations. *Genetics*, *176*(1), 351–9. doi:10.1534/genetics.106.067355

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford University Press.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358*(6389), 749–750.

Wynn, K. (1998). Psychological foundations of number: numerical competence in human infants. *Trends in Cognitive Sciences*, *2*(8), 296–303.

Xing, J., Watkins, W. S., Shlien, A., Walker, E., Huff, C. D., David, J., … Jorde, L. B. (2010). Toward a More Uniform Sample of Human Genetic Diversity: A Survey of Worldwide Populations by High-density Genotyping. *Genomics*, *96*(4), 199–210.

Zack, N. (1996). *Bachelors of science: Seventeenth-century identity, then and now*. Temple University Press.

Zack, N. (2002). *Philosophy of Science and Race*. New York: Routledge.