IQ and Immigration Policy

A dissertation presented

by

Jason Richwine

to

The Department of Public Policy

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Public Policy

Harvard University
Cambridge, Massachusetts

May 2009

© 2009 – Jason Richwine
All rights reserved.

Dissertation Advisor: Professor George J. Borjas Author: Jason Richwine

IQ and Immigration Policy

Abstract

The statistical construct known as IQ can reliably estimate general mental ability, or intelligence. The average IQ of immigrants in the United States is substantially lower than that of the white native population, and the difference is likely to persist over several generations. The consequences are a lack of socioeconomic assimilation among low-IQ immigrant groups, more underclass behavior, less social trust, and an increase in the proportion of unskilled workers in the American labor market. Selecting high-IQ immigrants would ameliorate these problems in the U.S., while at the same time benefiting smart potential immigrants who lack educational access in their home countries.

CONTENTS

	Acknowledgments	v
Part One	**PRELIMINARIES**	
	Introduction	2
One	The Science of IQ	6
Part Two	**THE IMMIGRANT IQ DEFICIT**	
Two	Immigrant IQ	26
Three	Hispanic IQ	60
Four	Causes of the Deficit	67
Part Three	**CONSEQUENCES AND SOLUTIONS**	
Five	The Socioeconomic Consequences	79
Six	The Labor Market Consequences	105
Seven	IQ Selection as Policy	123
Appendix A	List of National IQ Scores	135
Appendix B	Details of IQ Calculations	142
Appendix C	List of Countries by 1970 Education Level	145
	References	147

ACKNOWLEDGMENTS

I am indebted to the American Enterprise Institute for its generous support, without which this dissertation could not have been completed. In particular, I must thank Henry Olsen, head of AEI's National Research Initiative, for bringing me to AEI and supporting my research.

The substance of my work was positively influenced by many people, but no one was more influential than Charles Murray, whose detailed editing and relentless constructive criticism have made the final draft vastly superior to the first. I could not have asked for a better primary advisor.

I want to thank my Harvard committee members, starting with George Borjas, who helped me navigate the minefield of early graduate school and has now seen me through to the finish. Richard Zeckhauser, never someone to shy away from controversial ideas, immediately embraced my work and offered new ideas as well as incisive critiques. Christopher Jencks generously signed on as the committee's late addition and offered his own valuable input.

I also need to thank my colleagues Bruno Macaes, Scott Rosen, Deepa Dhume, and Abigail Haddad for important discussions along the way, as well as Batchimeg Sambalaibat for her patient Stata tutoring. Eager AEI interns Maria Murphy, Stephen Meli, Jordan Murray, and Emma Jackson provided useful research assistance. And thanks to Dan Black and Steve McClaski at the Center for Human Research Resources for enduring all of my pesky emails.

As usual, the people I have thanked are not liable for any errors that may have escaped attention. That responsibility is my own.

Part One:

PRELIMINARIES

INTRODUCTION

In the first couple of decades after World War II, immigrants were a small portion of the American population, coming mainly from Europe due to formal and informal restrictions on non-white immigration in place since the 1920s. Immigrants at the time had slightly less education but earned slightly more income than natives. The situation began to change after 1965, when the U.S. abolished national origin quotas, set aside specific visas for Western hemisphere immigrants, and gave preference to applicants who had relatives residing in the U.S. (Lynch and Simon 2003, 16). The new policy, combined with periodic increases in visa allowances and a growing illegal immigrant presence, helped to change the type of immigrants who came to the U.S. Immigrants have become increasingly less skilled, in terms of education and income, relative to the native population (Borjas 1999, 21-22).

This situation is not necessarily problematic. European immigrants in the late nineteenth and early twentieth centuries were similarly unskilled, but fears that they would damage American society proved to be baseless. The optimistic argument says that if today's immigrants gradually get better educations and move up the socioeconomic ladder, then they could assimilate culturally and economically just as Europeans did. However, this optimism is unwarranted if the average immigrant lacks the raw cognitive ability, or intelligence, to pursue higher education and take on skilled labor. Just as low intelligence will limit an individual's career choices, low average intelligence in a group will inhibit its overall success. This dissertation assesses the average intelligence of current immigrants living in the U.S. and explores its implications.

Although a precise definition of intelligence is impossible, it has been broadly described as "…the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly, and learn from experience" (Gottfredson 1994). To approximate intelligence, I

use the statistical construct known as IQ, which helps to explain the variance in human performance on a wide range of cognitive tasks. The next chapter provides a much more detailed discussion of the science behind IQ; for now, it is sufficient to state that IQ is a reliable and valid operational measure of general intelligence.

The major finding presented here is that the average IQ of immigrants is substantially lower than that of the native population, and the difference does not disappear by the second or third generation. The result is a lack of socioeconomic assimilation, and an increase in undesirable outcomes such as underclass behavior and loss of social trust. The upside is that calling attention to this problem may help focus policy on attracting a different kind of immigrant—the poor with great potential. A summary of the chapters follows.

Chapter 1 reviews the science of IQ. I show that the existence of general intelligence is widely-accepted, that it can be reliably measured using IQ tests, and that it is determined partly by genes. I also review the history of research on immigrant IQ, showing that, contrary to conventional wisdom, there was no consensus among early twentieth century intelligence researchers that European immigrants had low average IQs.

Chapter 2 moves on to the empirical heart of the dissertation, the demonstration that the IQ of current immigrants is considerably lower than that of the native population. Four different datasets are analyzed, and average immigrant IQ is estimated to be in the low 90s, on a scale where white natives are at 100. When broken down by national origin, the estimates differ greatly. Mexican immigrants average in the mid-80s, other Hispanics are in the low 90s, Europeans are in the upper 90s, and Asians are in the low 100s. IQ scores go up slightly in the second generation, but the scores of Mexicans and other Hispanics remain well below those of whites, and the differences persist over several generations.

Chapter 3 looks specifically at Hispanic American IQ estimates from a variety of secondary sources. The results are consistent with the second and third generation Hispanic immigrant IQs detailed in the previous chapter. The chapter also uses the historical experience of Hispanic Americans to argue that today's immigrant IQ deficit is not a short-lived (or even illusory) phenomenon as it was for European immigrants in the early twentieth century.

Chapter 4 discusses the possible causes of the deficit. First, the U.S. may be attracting immigrants from the low-side of the IQ distribution in their home countries. Second, material deprivation—such as inadequate nutrition, healthcare, and early schooling—could depress immigrant IQ scores. Third, cultural differences that deemphasize education may be a factor. Finally, genetic differences among ethnic groups could contribute to the difference. The chapter assesses the plausibility of these explanations, concluding that the material environment and genes probably make the greatest contributions to IQ differences.

Chapter 5 is the first of two chapters that analyze the effects of immigrant IQ on American society. This chapter first reviews the numerous socioeconomic correlates of IQ, arguing that many of the correlations reflect a causal relationship between intelligence and the outcome in question. The chapter moves on to describe the typical skills of people with IQs in the low 90s. The rest of the chapter focuses on two areas of social policy in which IQ's importance is rarely mentioned. First, low IQ is a likely underlying cause of the Hispanic underclass, since a natural impetus to disengage from the cultural mainstream is the inability to succeed at the same level. Second, there is evidence that relatively high IQ is a necessary precondition for developing societies with high amounts of "social capital." Ethnic diversity undermines social capital, but high-IQ minorities may mitigate the diversity problem.

Chapter 6 uses a model of the labor market to show how immigrant IQ affects the economic surplus accruing to natives and the wage impact on low-skill natives. All workers, no

matter what their IQ, benefit natives as a whole to some degree by lowering the prevailing wage in the sectors in which they compete. The lower wage translates to lower prices for consumers. However, higher IQ immigrants take the skilled jobs that maximize the economic surplus and minimize the adverse impact on wages for low-skill natives.

Chapter 7 concludes by exploring the policy implications of these findings. I argue that selecting immigrants on the basis of IQ has some obvious and subtle benefits. IQ selection would obviously reverse the cognitive decline of immigrants, but it would also benefit a large number of intelligent yet underprivileged people who would be ineligible under selection systems that emphasize educational attainment. Giving high IQ citizens of poor countries the chance to get an education that matches their cognitive skill would be a win-win situation.

Chapter 1: THE SCIENCE OF IQ

Before beginning the main analysis, it is important to establish exactly what IQ is and how it is measured. A number of myths and misconceptions surround the science of cognitive ability (Sternberg 1996), and the national media frequently misstate our current knowledge about it (Snyderman and Rothman 1988). It is still not unusual to hear a commentator claim that IQ is not real, or is not useful, or is merely a proxy for education or privilege. As the first part of this chapter demonstrates, the actual psychological literature says otherwise. The second part of the chapter examines how others have viewed immigration through the lens of IQ in the past, and then summarizes the small amount of modern research on the topic.

THE AMERICAN PSYCHOLOGICAL ASSOCIATION STATEMENT ON IQ

Strictly speaking, few aspects of IQ research are without controversy, but a general consensus about its fundamentals has emerged among most psychologists. After the media furor surrounding publication of Richard Herrnstein and Charles Murray's *The Bell Curve* (1994), the American Psychological Association (APA) published a statement (Neisser et al. 1996) on the current science regarding intelligence, which is an authoritative summary of a vast literature. The APA report cannot entirely end debate on any issue, but I use it to show that the treatment of IQ in this study is firmly grounded in the psychological mainstream.

The APA did not address *The Bell Curve*'s central claim about IQ determining social class structure, but it did affirm that its handling of IQ as a science was sound. Among the specific conclusions drawn by the APA were—IQ tests reliably measure a real human trait, good tests of IQ are not culturally biased against minority groups, and IQ is a product of both genetic inheritance and early childhood environment. A similar report signed by 52 experts, entitled "Mainstream Science on Intelligence," also stated those same facts (Gottfredson 1994). Every

bold subheading in this section is a direct quote from the APA report. The discussion that follows each quote is my own summary of the literature.

"…the *g*-based factor hierarchy is the most widely accepted current view of the structure of abilities…" The existence of general intelligence was inferred by early psychometricians who noticed high positive score correlations among tests that covered very different topics. For example, people who are good at rotating three-dimensional objects in their mind also tend to be good at understanding verbal analogies, applying rigorous logic to solve math problems, detecting patterns in a matrix of shapes, repeating backward long sequences of digits that are read aloud, and so on. In fact, performance on any two tasks that tax the brain tend to be correlated, no matter how substantively different the tasks appear to be. These correlations are due to the existence of general intelligence. The average person who scores well on both math and verbal tests is not blessed with separate talents for each subject. He scores well on both because he is generally smart.

Psychometricians can quantify just how much performance is due to a general mental factor by performing a factor analysis of scores on a wide variety of cognitive tests. This process attempts to find the underlying factors within a matrix of correlations between tests. If the tests were unrelated to each other, then factor analysis would fail to simplify the data—10 unrelated tests would mean that each test can explain only 10% of the score variance. However, psychometricians have found that a single underlying factor, which they call *g*, almost always accounts for a large proportion of the variance, usually more than half (Carroll 1993, 57). The people who do well on cognitive test batteries are the ones who have high *g*.

One cannot claim that *g* is precisely the same thing as intelligence, because intelligence itself has proved impossible to define satisfactorily (Jensen 1998, 46-49). However, *g* corresponds so well to our everyday conception of what it means to be generally smart that the

two terms are often used interchangeably. It must be noted, however, that IQ and *g* are *not* the same thing. An IQ test is used to approximate the *g* factor, and the best IQ tests are those that are highly "*g*-loaded," meaning correlated with *g*. For example, the Armed Forces Qualification Test (AFQT), a cognitive assessment used by the military, correlates at about 0.83 with *g*, meaning *g* explains nearly 70% of the variance in AFQT scores, with 30% explained by several much smaller factors, including random error. A person's IQ is simply his score on an IQ test. This score is a very good—but nevertheless not perfectly exact—approximation of his general intellectual ability, or *g*. Throughout this study, I will maintain the distinction by referring precisely to either IQ or *g*.

Since the APA report was written, neurologists have begun to demonstrate a physiological basis for *g* inside the brain, providing even more convincing evidence that *g* is essentially mental ability. We know that brain size and IQ (not necessarily *g* itself) are correlated (Andreasen et al. 1993), but Haier et al. (2004) showed that a specific set of small regions of the brain account for much of that correlation. Now even more recent studies by neurologists have better isolated the *g* factor as a real property of the brain. For example, Colom et al. (2006) administered complete IQ test batteries and brain MRIs to a group of 48 adults. They found that the correlation between amount of "gray matter"—bundles of interconnected neurons in the brain—and subtest performance went up linearly with the *g*-loading of the subtest. In other words, the more a subtest taps *g*, the more a person's amount of gray matter affects his performance.

A common objection to the idea of a single, unitary *g* is that some people seem quite lopsided in their abilities—everyone knows the literature buff who trembles at the sight of a math textbook, or the science nerd who can't seem to put two sentences together. But these differences are often exaggerated, because people tend to compare themselves only to their

immediate peers. In many cases, their peer group is far from representative of the nation as a whole. At an elite college, for example, a physics major may be in the 99th percentile of mathematical ability in the general population and "only" the 90th in verbal ability. That difference is real and tangible when this person compares himself to his friends; in fact, it might have determined his choice of major. However, in everyday life and in most lines of work, the difference is negligible.

This is not to say that abilities more narrow than *g* are non-existent. They do exist, but most psychometricians see them as lower-order factors still dependent in large part on *g*. Carroll's (1993) authoritative survey establishes a hierarchical, "three-stratum" model of intelligence. At the top of the hierarchy is *g*, followed by a handful of broad second-order abilities, followed by many narrow third-order abilities. The three-stratum model emerges from the fact that certain first-order abilities tend to cluster together into broader second-order categories. For example, tests of visualization and spatial perception correlate more highly together than either one correlates with vocabulary tests. Carroll classifies these visualization and spatial perception skills as part of a second-order "broad visual perception" category. Other second-order factors include "crystallized intelligence" (learned knowledge), "fluid intelligence" (abstract reasoning ability), and memory power.

Crucially, all of the second-order factors are dominated by *g*, the single third-order intelligence factor. Individuals with higher *g*'s will tend to have higher abilities in all of the second- and first-order categories. Individuals with the same *g* will still differ to some degree in lower-order factors, but much of the variance in these narrower abilities is eliminated by controlling for *g*. If certain mental abilities were independent and distinct, multiple *g*'s could emerge at the top of the hierarchy—but, as Carroll shows, this does not happen. As the quote

from the APA report that began this section put it: "…the *g*-based factor hierarchy is the most widely accepted current view of the structure of abilities…"

The APA statement does warn that not all psychometricians subscribe to the view of a dominant *g*. In fact, a small group favors multidimensional models, such as Howard Gardner's (1983) theory of multiple intelligences (MI) and Robert Sternberg's (1985) triarchic theory. These are interesting attacks on the mainstream view, but they remain the viewpoints of a small minority. Gardner and other MI theorists usually acknowledge the data showing high subtest correlations that produce a general intelligence factor, but they argue such correlations could be due to a common upbringing that enriches different types of intelligence independently (Gardner 2006), suggesting a valid empirical test of MI has yet to be devised

Most psychometricians are unconvinced by this theory, because Gardner has not demonstrated that separate "intelligences" can be observed independent of *g*. The predominant view is that MI theory is really just a variant of the hierarchical structure described by Carroll, the model that I embrace for this study. The debate over MI cannot be resolved here, but even if MI theorists could somehow succeed in splitting *g* into independent factors, traditional IQ scores would remain important measures of ability.

"Intelligence test scores are fairly stable during development." IQ tests have a high reliability coefficient, which is the correlation between the test scores of the same individual. As the quote indicates, tests remain generally reliable throughout a person's life, starting around the beginning of elementary school. The APA report cites a correlation of 0.86 between a person's IQ—actually his average score on several IQ tests, to reduce measurement error—taken around the ages of 5 to 7 with his average score at ages 17 to 18. If the younger age range is bumped up to 11-13, then the correlation with the late teenage years becomes 0.96 (Bayley 1949, table 4). The correlation remains quite high throughout middle age (Larsen et al.

2008). This is not to say that no one other than infants or the elderly ever sees his IQ score change substantially—distracting testing conditions, illnesses, and simple random measurement error can all affect scores.

"…a sizable part of the variation in intelligence test scores is associated with genetic differences among individuals." Like many human traits, an individual's IQ is determined by an interaction of his genes and his childhood environment—no major expert today believes that IQ is a product of just one or the other. Since attempts to disentangle each factor's effects are quite difficult, researchers have generally relied upon studies of twins to estimate the genetic component of IQ scores. Identical twins ("monozygotes") share the same genetic code; therefore, monozygotes raised in separate homes are subjects in a natural experiment that holds genes constant while varying the environment.

Results from twin studies emphasize that there are three different factors that explain the variance in IQ scores—genes, the shared environment, and the nonshared environment. The shared environment encompasses a person's experiences that do not differ from his siblings in the same household—parental income and occupation, school attended, number of books in the home, etc. The nonshared environment is the set of personal experiences that are not directly related to the household situation—peer groups, for example, or environmental events affecting brain development *in utero* or during infancy. According to the APA summary of the twins data, the proportions of IQ variance explained by genes, shared environment, and nonshared environment among children are 0.45, 0.35, and 0.20, respectively. Heritability then increases with age, with genetic variance rising to 0.75, shared environment falling to near zero, and nonshared environment at around 0.25.

Psychologists typically rely on identical twins to determine genetic contributions to IQ, given the genetic equivalence of monozygotes, but the studies are not perfect. For example,

although the genetic proportion of IQ variance is large, it does not necessarily limit the impact of the environment on IQ. Theoretically, people with certain genotypes could choose (or be given) more favorable environments that tend to enrich intelligence, which would lead some environmental benefits to be attributed to genes (Jencks 1980; Dickens and Flynn 2001).

Additionally, studies that use regular biological siblings rather than twins have the advantage of much larger sample sizes, but they inevitably require questionable assumptions built into elaborate models of genetic transmission. Studies that have attempted modeling—e.g., Feldman et al. (2000) and Daniels et al. (1997)—have generally found lower genetic heritability estimates in the 0.35 to 0.45 range, although the estimates vary considerably depending on the model specification. Even if the APA has underestimated the environmental contribution to IQ by excessive reliance on twin studies, no one claims an insignificant role for genes.

"**The differential between the mean intelligence test scores of Blacks and Whites…does not result from any obvious biases in test construction and administration…"** This quote from the APA actually makes two points. First, groups differ in average IQ, and, second, the differences are not due to any obvious test bias. By far the most frequently studied group difference is the APA-affirmed 1.0 standard deviation IQ differential between whites and blacks. Since IQ has a normal distribution—i.e., a bell curve—in populations, this difference places the average black at roughly the 16th percentile of the white IQ distribution.

Several other group differences have been examined, albeit to a lesser extent. The APA notes that Hispanics have reliably tested somewhere between whites and blacks, and East Asians probably have slightly higher IQs than whites. Also, although unmentioned by the APA, Jews have a substantially higher average IQ compared to non-Jewish whites (Murray 2007a; Entine 2007, 303-311).

All of these observed group differences in IQ lead to the question about whether the tests are biased, in the sense that they measure IQ less accurately for some groups compared to others. The answer is "no." The APA report focused on evidence showing no test bias against specifically blacks, but the authors of "Mainstream Science on Intelligence" go a step further by stating: "Intelligence tests are not culturally biased against American blacks or other native-born, English-speaking peoples in the U.S. Rather, IQ scores predict equally accurately for all such Americans, regardless of race and social class."

Briefly, the evidence concerning test bias comes in two forms, external and internal. The external validity of tests refers to how well they predict outcomes for each group in question. For example, if a score of 1300 on the SAT corresponded to a college GPA of 3.0 for whites, and the same 1300 led to an average GPA of 3.5 for blacks, then the SAT might be biased against blacks, since it has underpredicted their college achievement. However, no such result has been uncovered for the SAT or for any other widely-used standardized test. When the predictive value of tests differ at all by race, they tend to *over*predict black achievement. Tests also show the same internal validity for all of the groups in question. This means that test items show the same relative difficulty within groups, and that the factor structure of subtests is roughly the same for each group as well. Jensen (1980) is still the definitive account of test bias (Reeve and Charles 2008).

Since the publication of the APA report, another potential bias has been identified. Steele and Aronson (1995) coined the term "stereotype threat" to describe the phenomenon of black students performing differently on the same test depending on the test's name. The theory is that blacks, reacting to society's alleged stereotype that they are unintelligent, naturally perform worse when the same test is called an "intelligence test" rather than a "skills" test.

However, stereotype threat does not account for the black-white test score gap—it can only make the gap larger than what is normally observed (Sackett et al. 2004).

"**Mean scores on intelligence tests are rising steadily….No one is sure why these gains are happening or what they mean."** Herrnstein and Murray called the rise in test scores the Flynn effect, naming it after the man who is most responsible for bringing attention to it (Flynn 1984; Flynn 1987). The Flynn effect, which cumulatively has amounted to over 1 standard deviation since World War II, is not the result of one particular socioeconomic or ethnic group making gains on another, although part of the trend has been ascribed to improved early education and nutrition amongst the very poor (Lynn 1990). Much of the Flynn effect is like a rising tide lifting all the boats. Explanations such as the growth of a more cognitively challenging culture are, like nutrition, incomplete at best according to the APA. Similarly, Jensen (1998, 323-324) casts doubt on Brand's (1987b) suggestion that improved guessing ability is behind the Flynn effect. The real cause remains a mystery.

But the secular increase in IQ test scores does *not* prove that people are getting significantly smarter. Remember that IQ and *g* are not the same thing, so that improved performance on IQ tests could be due to gains in the non-*g* components of the tests. Indeed, Wicherts et al. (2004) found that IQ tests are not "measurement invariant" over time, meaning that the relationship between each subtest and *g* changes somewhat depending on the cohort that takes the overall battery. This means that IQ test scores are still fine approximations of *g* within cohorts, but that the tests should be frequently re-standardized over time to keep scores comparable. The issue may be becoming less important, however, because new evidence suggests the Flynn effect is now slowing or even reversing (Teasdale and Owen 2008; Flynn in press).

Summary. Like all sciences, the study of mental ability is fraught with ongoing disputes and controversies. However, most psychometricians have come to agree on a core set of findings that define the mainstream of their field. Among those core findings are that IQ tests reliably measure a trait known as general intelligence or ability, that scores on such tests arise from gene-environment interactions, that score differences between ethnic groups are not due to test bias, and that scores have risen largely independent of *g* throughout the twentieth century.

IQ Outside Psychology

Much of the science reviewed so far, treated as uncontroversial by the APA, may seem surprising to non-specialists. This unusually large discrepancy between expert knowledge and the conventional views held by educated laypeople is documented in Snyderman and Rothman (1988). They write:

> …the literate and informed public today is persuaded [wrongly] that the majority of experts in the field believe it is impossible to adequately define intelligence, that intelligence tests do not measure anything that is relevant to life performance, and that they are biased against blacks and Hispanics, as well as against the poor. It appears from book reviews in popular journals and from newspaper and television coverage of IQ issues that such are the views of the vast majority of experts who study questions of intelligence and intelligence testing. (250)

The discrepancy developed mainly because IQ can be an uncomfortable topic in a liberal democracy. The reality of innate differences between individuals and groups is often difficult to accept for those with an aversion to inequality. For this reason, journalists and academics in other fields are naturally attracted to scholars who downplay the role of genes in determining IQ, even if these scholars are a distinct minority. For example, media reports often approvingly cite iconoclasts like Leon Kamin, usually giving the false impression that their anti-heredity work reflects a widely-held viewpoint. At the same time, a more mainstream scholar like Arthur Jensen is portrayed as the defender of a marginalized group of hereditarians (247).

Even more troubling is the frequent citation of *The Mismeasure of Man* (1981), paleontologist Stephen Jay Gould's anti-IQ polemic written for a popular audience. In *Mismeasure*, Gould dismisses psychometrics as a pointless, invalid discipline used mostly to pursue racist agendas rather than to understand anything about mental ability. The book makes for a good case study of how the media are willing to embrace an apparently appealing message even as experts roundly reject it. To highlight this gaping difference of opinion, Davis (1983) contrasted the rave reviews of *Mismeasure* in the popular press with its negative reception in technical journals such as *Science, Nature*, *Contemporary Education Review*, *Intelligence*, *Contemporary Psychology*, and the *American Journal of Psychology*. The closer the reviewer was to pyschometrics the more severely he panned it. For example, the late John Carroll, one of the foremost experts on the factor analytic basis of *g*, said of Gould: "Some have called his exposition masterful, but I would call it masterful only in the way one might use that word to describe the performance of a magician in persuading an audience to believe in an illusory phenomenon" (1995, 125).

The book itself contains many claims about IQ—in particular, that *g* is a meaningless mathematical artifact (ch. 6)—that the APA report flatly contradicts. Gould also pokes fun at the poor methodology used by some early intelligence researchers, in an attempt to depict the whole field of psychometrics as a pseudoscience practiced by cranks. But it is hardly reasonable to lump dubious early work on intelligence with modern psychometrics, treating the whole history of IQ research as an unbroken line of fraudulent science. As Davis writes, this is analogous to condemning the medical profession by penning "…a tendentious history of medicine that began with phlebotomy and purges, moved on to the Tuskegee experiment on syphilitic Negroes, and ended with the thalidomide disaster…" Gould contributed essentially nothing to the science of IQ, but his influence among laypeople regrettably remains.

THE HISTORY OF IMMIGRATION AND IQ RESEARCH

Surprisingly little work has been done on immigration and IQ in the modern era, but the topic was analyzed in some detail in the early twentieth century. Once again, the facts are at odds with the conventional wisdom in the media. The typical history—Kamin (1974) and Gould (1981) are good examples—usually contains some or all of the following myths: early psychometricians developed IQ tests in order to show the ethnic supremacy of northern European "Nordics," testing at that time "proved" this point, and this proof led directly to the 1924 immigration restrictions that favored Nordics over other types of Europeans. In fact, none of these things is true. IQ tests were developed to help identify children with learning disabilities. Testing was seen as a much more efficient method for determining which children needed different types of curricula and extra help (Thorndike and Lohman 1990, 21-25). Later, intelligence tests became useful to large organizations, particularly the U.S. Army, which needed quick ways to assess aptitude and trainability.

It is true that some psychometricians, just like many educated Americans at the time, held views on race that are considered unacceptable today. But Kamin, Gould, and other critics used highly selective evidence to portray the entire field as hopelessly obsessed with proving racial differences. There certainly were some dubious IQ studies based on ethnicity and national origin, the most prominent of which (Brigham 1923) is discussed below. But a healthy debate within psychometrics was being waged in the 1920s about ethnicity and IQ. There was hardly any consensus at all about the topic—witness the numerous critical reviews of Brigham's racialist work by contemporary social scientists like E.G. Boring, Kimball Young, Percy Davidson, and William Bagley. Even Robert Yerkes and Lewis Terman, usually seen as sympathetic to Brigham's racial views, cautioned against his sweeping conclusions (Snyderman

and Herrnstein 1983). Like all fields, psychometrics was in the process of maturing as a science. In fact, Brigham (1930) eventually rejected his own methodology.

The Immigration Act of 1924. Concerned that the changing ethnic mix was altering the country's culture, Congress in 1924 severely restricted further immigration. National origins quotas were imposed, aimed at preserving the ethnic balance of the U.S. as of the 1890 census. Probably because there was no agreement about the science, IQ testing did *not* significantly influence this debate on immigration in the 1920s. In fact, an analysis of the Congressional debate on the act reveals almost no discussion of IQ. During those rare times when the mental ability of immigrants was mentioned at committee hearings, it was almost always to criticize the science as inconclusive or unsupportable. Debate on the floor of Congress showed even less concern for intelligence testing—just one instance in over 600 pages from the Congressional Record. Furthermore, no major IQ researchers were called to testify, and the final bill made no mention of testing (Snyderman and Herrnstein 1983).

Brigham. Although its viewpoint was hardly typical, it is still instructive to review Carl Brigham's *A Study of American Intelligence* (1923), the IQ research most explicitly associated with anti-immigration sentiment. Some of the book's methodological and interpretive problems were already noticeable in the 1920s, and they are glaring today. Brigham analyzed army intelligence testing used during World War I to compare the intelligence of officers versus draftees, whites versus blacks, and white natives versus immigrants (80-86). The group performance differences in standard deviations, often referred to as *d*'s, were 1.88, 1.08, and 0.60, respectively.

The army tests were crude by today's standards—they overemphasized test-taking speed, lacked the ability to differentiate people on the lower tail of the bell curve, and were put together in an ad-hoc fashion. Part of the "beta test," the version given to illiterate recruits, was particularly odd—it required recruits to interpret hand movements and suggestive facial

expressions just to understand the test directions. Brigham also did not offer the reader many of the psychometric properties of the intelligence test that researchers expect to see today, such as loading on *g*, the subtest intercorrelation matrix, and measures of reliability.

Brigham insisted that the native-immigrant test score difference reflected a real difference in intelligence. He explained this result by borrowing a racial theory (Grant 1916) that seems bizarre to the modern reader. Dividing Europe into three racial categories, he argued that Nordics were intellectually superior to people from the Alpine and Mediterranean regions of Europe. American natives, who were mostly of English and German descent, outscored early twentieth century immigrants who were from southern and eastern Europe. Based on this result, Brigham strongly hinted that non-Nordic immigration should be ended. Although he did not explicitly call for a race-based policy, his condemnation of interracial marriage and unrelenting focus on race clearly suggested what type of immigration program he would favor (197-210).

The most obvious problem with an ethnically exclusionary immigration policy is that it would be unnecessarily restrictive. According to Brigham's own results, there were thousands of Alpines and Mediterraneans who outscored the average Nordic, even if the mean group differences were valid. There would be no reason to exclude them purely on the basis of their group membership.

The other problem with Brigham's conclusions is that they were based on assumptions that we now know to be false. Although small differences are always possible, there is no modern evidence of substantial IQ differences among American whites of different national backgrounds. As mentioned above, Asian-white-Hispanic-black group differences certainly do exist in the U.S., but, with one important exception, intra-European differences do not. The only Americans from a European ethnic group that score consistently higher than the white

average are Jews, who did not come from a single nation. Ironically, Brigham was wrong about the one European ethnic group that actually is more intelligent than the average white, when he claimed that his numbers "…tend to disprove the popular belief that the Jew is highly intelligent" (190).

So where did Brigham go wrong? It appears that his beta test, the one that did not require English literacy, probably still suffered from bias. It is quite likely that people having no experience at all with the types of questions on IQ tests could be at a disadvantage, particularly in tightly-timed settings. This is especially true for Brigham's era, when high school graduation in the U.S. was rare, and some immigrants had no schooling at all. It is not that schooling necessarily imparted specific information that gave educated people an advantage—it is the fact that people in school were more familiar and comfortable with IQ test questions. This may be why the officer-draftee *d* of 1.88 was so high. Although the officers were almost certainly smarter than raw recruits, most officers had extensive schooling, while many draftees had little to none.

Interestingly, Brigham had contrary evidence in front of him. He reported that immigrant IQ scores rose with time of residency in the United States. In fact, immigrants who had been in the U.S. for twenty years or more had the same average IQ as natives! With just a static snapshot of America, it was impossible to know whether residency in the U.S. raised test scores or whether immigrant quality had simply become lower. Brigham chose the latter interpretation. His evidence was that greater proportions of non-Nordics were present among the most recent immigrants. But this was assuming what he was trying to prove, which was that non-Nordics were less intelligent. He also argued that even scores on the non-biased beta test rose with time of residency, meaning residency could not impart any experiences that were

advantageous on the test. Again, however, it is unknown whether the beta test was actually unbiased.

Obviously, Brigham's work is not the kind of science that should be emulated. This study differs from Brigham's in at least three important ways. First, the science of IQ was still in its infancy at the time of Brigham's writing. It is easy to parody early intelligence researchers who—just like early chemists, biologists, and geologists—made many assumptions that we now know to be untrue. As this chapter has hopefully demonstrated, the study of IQ is now a mature science with a well established empirical foundation. This study draws on the most up-to-date sources and materials from the psychometric world, a body of literature that is vastly larger and superior to what was available to Brigham. Second, I account for test bias against immigrants using several different datasets, a variety of techniques to evaluate test validity, statistical controls for education where necessary, and second generation data to look for test score convergence.

Finally, as I emphasize throughout the whole text, nothing in this study suggests that immigrants should be treated on the basis of their group membership. Although the next chapter presents some facts about how IQ varies across countries and ethnic groups, immigrants—and, indeed, all people—should be considered purely as individuals whenever possible. Unlike Brigham's *A Study of American Intelligence*, there is no racial or ethnic policy agenda here. One can deal frankly and soberly with group IQ differences while still subscribing to the classical liberal tradition of individualism.

MODERN RESEARCH

Immigration became a non-issue for most social scientists after the 1924 restrictions and the Great Depression made coming to the U.S. more difficult and less beneficial. But significant liberalization of immigration laws after 1965 revived interest in the topic. After the doors were

opened to Asian and Latin American immigrants, social science research on nearly all aspects of immigration policy eventually followed. However, unlike during the previous great wave, immigrant IQ has been largely excluded from the academic discussion, and with little justification. As this chapter has demonstrated, IQ has not been proven illegitimate or useless; on the contrary, modern research has cemented its standing as a measure of a fundamental human trait.

In the United States. The most relevant research in the U.S. has not focused on the broader implications of immigrant IQ. Instead, researchers have emphasized the more narrow issue of possible language biases faced by Hispanics and non-native speakers on psychological tests. As discussed above, no such bias exists for native speakers, but it may be present among those who speak English only as a second language. It is obvious that people who speak little to no English will not get a meaningful score on an English-language IQ test—that is certainly not in dispute. The more interesting question is how meaningful IQ scores become for non-native speakers with moderate to high proficiency in English—the typical immigrants studied in the next chapter.

One way to answer that question is to examine test scores on school admissions tests, since it would be unusual for a non-English speaker to apply to a school that conducts classes in English. Pennock-Roman (1992) surveyed studies of non-native speakers, particularly Hispanics, who took the SAT, ACT, and LSAT. In virtually all of the studies she cites, the ability of the tests to predict school grades did not significantly differ for non-native speakers compared to natives, or for Hispanics compared to non-Hispanic whites. Even specifically adding a measure of English proficiency added little to the accuracy of the predictions, and the verbal and mathematics sections of the SAT were roughly equal in their predictive power.

Since language difficulty could simultaneously affect test scores and college grades, external validity alone does not prove the complete absence of bias. Indeed, other test difficulties have been reported. For example, younger Hispanic children usually perform significantly better on non-verbal tests compared to verbal ones (Munford and Munoz 1980; Whitworth and Chrisman 1987). Converting English language tests to Spanish can introduce score anomalies (Valencia and Rankin 1985), and non-native speakers have a statistically significant disadvantage on mathematics tests, although its magnitude is tiny (Abedi and Lord 2001). Clearly, the testing of non-native speakers has problems that must be addressed through careful bias checking. However, the existing evidence shows that language difficulties are not an insurmountable problem, and that test results of non-native speakers are interpretable.

In the Netherlands. Dutch psychologists have been more willing to study the IQ of immigrants compared to their peers across the Atlantic. Although immigrants to Western Europe tend to be from the Middle East and South Asia rather than Latin America, the potential language and cultural biases they may face are comparable to the Hispanic experience in the U.S. Indeed, most of the Dutch research on immigrants conforms to the American findings on non-native speakers—although particular items and subtests show bias, most standardized testing is valid (te Nijenhuis and van der Flier 1999). For example, one study of Dutch immigrants (te Nijenhuis and van der Flier 2003) using the General Aptitude Test Battery found that the vocabulary subtest contained several biased items, but the other subtests showed little bias. Wicherts (2007, ch. 2) has suggested that the magnitude of the bias on certain subtests has been underestimated, but other subtests do not appear biased at all. Although they have conducted more empirical studies of immigrant IQ than Americans, the Dutch have similarly avoided a major discussion of its consequences.

SUMMARY

Although IQ research features controversies like any other scientific field, psychologists have come to a broad-based consensus on its foundations. There exists a general, partially-hereditary, physiologically-based intelligence factor called *g*. Standard IQ tests are reliable, unbiased approximations of this *g* factor, but mean IQ scores are not the same across ethnic groups or over time. In modern times, only a small number of researchers in the U.S. and Europe have analyzed immigrant IQ, and none has addressed its broader implications. The rest of this study begins that work, starting with the most important question—what is the average IQ of current immigrants?

Part Two:

THE IMMIGRANT IQ DEFICIT

Chapter 2 : IMMIGRANT IQ

Immigrants living in the U.S. today do not have the same level of cognitive ability as natives. Using a variety of datasets, this chapter presents evidence that the average IQ of current immigrants is substantially lower than the native white average. The deficit is roughly one half of one standard deviation, and it will likely persist through several generations. I first present a table summarizing the overall findings, and then detail the methodology used to derive an IQ score from each dataset. This chapter and the next are empirical accounts of immigrant IQ. The chapters following them explore the possible causes of the deficit and its implications.

Table 2.1 summarizes immigrant IQ estimates from several different sources. Although no single dataset can definitively settle the question—they inevitably vary in test quality, sample representativeness, and year of testing—a substantial IQ deficit exists in each dataset examined.

Table 2.1

Summary of Immigrant IQ Estimates by Broad Regional Background

Immigrant Origin	Fraction of Immigrants in 2006	National IQ (various years)	AFQT Math (1980)	PIAT-R Math (1997)	Digit Span (2003)
Europe	14.6%	98.0	96.9	102.2	99.1
Mexico	31.8%	88.0	86.9	80.5	82.4
Other Hispanic	24.5%	81.7	91.1	91.3	84.5
Eastern and Southern Asia	23.2%	94.0	105.1	102.6	106.9
All		**88.9**	**93.3**	**91.9**	**93.3**

Notes: IQ estimates are normed to the white native distribution of intelligence, with a mean of 100 and a standard deviation of 15. All estimates come from sample sizes of 40 people or more; see text for details.

Based on the available evidence, current immigrants have an average IQ in the low 90s, probably in the range of 91 to 94, with white natives at 100. The following sections address the quality of the data used to derive this estimate, including issues of test bias and measurement error.

LYNN AND VANHANEN'S NATIONAL IQ SCORES

A metastudy of worldwide IQ by Lynn and Vanhanen (2002), whose updated 2006 data is used in this study, finds that countries differ dramatically in their average IQ, with East Asian countries ranked the highest and sub-Saharan African nations placed at the bottom. The study has been criticized for sometimes using small and unrepresentative samples, or using unreasonable assumptions to impute data (Barnett and Williams 2004). Reviewers have also balked at the sheer size of the IQ differences between countries (Nechyba 2004), which are over 3 standard deviations in some cases. But while their exact numbers can be questioned, Lynn and Vanhanen (LV) have drawn attention to real cognitive differences that exist worldwide. They used "culture fair" IQ tests—tests shown to exhibit the same predictive and internal validity for different ethnic and cultural groups—whenever possible, and they adjusted older test scores upward to account for the Flynn effect. They also showed that multiple tests within one country correlate at over 0.9, countering criticism that single tests in some countries are too unreliable.

Furthermore, the high correlation between national IQ and economic success supports the validity of LV's data. Dickerson (2006) has found that IQ can account for 70% of the variance in GDP across nations, assuming an exponential relationship between the two variables. This IQ-wealth relationship is not due to very low IQ scores from the world's poorest countries. In fact, the IQ-wealth correlation is essentially unchanged—it is stronger, if anything—when low IQ countries are discarded (Whetzel and McDaniel 2006). The predictive value of LV's dataset, not only in terms of national wealth and economic growth, but also as a positive correlate of

educational success, nonagricultural ways of life (Barber 2005), and even suicide rates across countries (Voracek 2004), is strikingly robust.

Are LV's IQ numbers just proxies for some other factor, such as education, nutrition, or free markets? Initially, results were mixed when researchers attempted to answer this question. Weede and Kampf (2002) found a consistently significant and independent effect of IQ on economic growth, while Volken (2003) made the effect disappear by adding certain educational variables. The debate was resolved with the publication of Jones and Schneider (2006), which used the most technically sophisticated methodology on the subject. Jones and Schneider employed a version of the "I just ran two million regressions" method of Sala-I-Martin (1997), in which the significance of a particular variable is tested in thousands of potential growth models. Jones and Schneider found that IQ is a statistically significant predictor of growth in 99.8% of those models.[1]

Relationship to U.S. Immigrants. The relevant question for this study is whether national IQ scores say anything about immigrants to the U.S. If we follow LV by assigning a Chinese immigrant an IQ of 105, and an Iranian immigrant an IQ of 84, do these numbers translate to observable outcomes, such as earnings differences? The answer is yes.[2] In their 2006 book, LV list six of the best attempts by economists to link IQ with the earnings of

[1] Jones and Schneider speculate that their conflict with Volken is due to data differences—they discarded imputed IQ data and tests with low sample sizes, while Volken retained all of Lynn and Vanhanen's data. They do not offer any empirical evidence that LV's imputed data is weak or inaccurate. In fact, LV were able to test their imputed data in their 2006 updated study, after they had acquired real tests for 25 countries with previously imputed IQ scores. The new measured IQ scores correlated at 0.91 with the imputed scores (55). In explaining the Jones and Schneider disagreement with Volken, it is more likely that Jones and Schneider's analytic technique is simply superior.

[2] What follows in this paragraph is a modified version of the same analysis performed in an earlier, unpublished version of the Jones and Schneider paper.

American males[3] (table 3.3). In particular, these studies ask what percentage increase in earnings is expected for every one standard deviation increase in IQ. The answers vary from 11% to 21%. These studies use IQ scores directly measured by testing the individuals. What if immigrants in the United States are simply assigned an IQ score based on their national background? Would the same 11% to 21% increase in earnings per standard deviation of IQ be observed? To find out, I performed a simple regression of log earnings on age and national IQ score for the immigrants in the 2006 March CPS, similar to the reduced form wage equations used in the studies cited by LV. The earnings increase corresponding to a one standard deviation increase in national IQ was 19.2%, in line with estimates using American natives with individual IQ scores.[4]

The reduced-form wage equation lacks controls for education quality, home environment, and neighborhood effects, which are inevitably correlated with IQ. Introducing those controls would attenuate the predictive power of IQ, but the point here is that when individual American IQ scores are used to measure skill, the economic return to that skill is essentially the same as when immigrants in the U.S. are assigned IQ-by-country estimates. This indicates the remarkable predictive validity of LV's data.

Immigrant IQ Estimates. IQ scores are relative. Although the distribution of intelligence in a population is always bell-shaped, the practice of assigning an IQ value of 100 to the population mean is simply a convenience. In their dataset, LV chose not to set the worldwide mean IQ at 100; instead, a score of 100 on their scale is equivalent to the average IQ

---

[3] Women tend to have lower labor force attachment for reasons unrelated to their skill—i.e., they have children, and some stay home to raise them. That is why only men are used in the wage equations.

[4] The regression is the log of total wage and salary earnings on age and national IQ, restricted to men ages 18 to 64 with nonzero earnings.

in Britain in 1979. The British mean of 100 is also the mean for American whites, whereas the American population as a whole has an average IQ of 98. In this study, the white American average is set at 100 to conform to LV's scale.

Table 2.2

Immigrant IQ Estimates by Regional Background Using National IQ Data

Immigrant Origin	Fraction of Immigrants in 2006	Average IQ
Europe	14.6%	98.0
Northeast Asia	8.9%	105.5
Southeast Asia	9.0%	89.3
South Asia	5.2%	82.3
Western Asia / Middle East	3.4%	85.8
North Africa	0.7%	81.4
Sub-Saharan Africa	1.6%	69.7
Mexico	31.8%	88.0
Central America / Caribbean	17.5%	79.7
South America	7.0%	86.6
Pacific Islands	0.2%	85.1
All	100.0%	**88.9**

Notes: IQ estimates are normed to the white native distribution, with a mean of 100 and a standard deviation of 15. People with unknown or ambiguous birthplaces are excluded.

The LV data allow for a simple initial calculation of immigrant IQ. The 2006 CPS March supplement gives the place of birth of a representative sample of the American population. The sample includes 24,492 immigrants, defined as U.S. residents who are either

naturalized citizens or non-citizens. Applying LV's national IQ scores in proportion to the national background mix of these immigrants yields an estimate of 88.9, over 11 points lower than American whites. As table 2.2 indicates, immigrant groups coming from outside of Europe and East Asia are even lower than the overall immigrant average. In contrast, immigrants from Northeast Asia score significantly higher than the native average. For more detail, Appendix A contains a full list of national IQ scores, describes which nations are in which regions, and discusses some miscellaneous technical issues.

Given the predictive power of LV's data, these estimates should be taken seriously. Still, the dataset does not account for selection. Perhaps the United States attracts the smartest immigrants from each of these countries, so that national IQ scores are lower than actual immigrant IQs. The next step then is to examine datasets with individual immigrant IQ scores. The first to be examined is the 1979 National Longitudinal Survey of Youth.

THE 1979 NLSY

The National Longitudinal Survey of Youth (NLSY) is a panel dataset that began interviewing a nationally representative sample of American young people about education, work, and family life in 1979. A unique facet of the NLSY is that in 1980 valid scores on the Armed Forces Qualification Test (AFQT) were obtained from 11,878 of the NLSY respondents, representing about 94% of the sample. The AFQT is a subsection of a larger battery of tests known as the Armed Services Vocational Aptitude Battery (ASVAB) that the military uses to assess intelligence, aptitude, and vocational skill. The AFQT itself is composed of four subtests—mathematics knowledge, arithmetic reasoning, word knowledge, and paragraph comprehension. Although the ASVAB contains numerous tests of knowledge and skill in specific fields—such as in electronics, automobiles, and general science—the AFQT subsection is much like the SAT. It requires some knowledge of English and algebra, but it is designed to

test intellectual ability, not merely acquired skill. The AFQT results from the NLSY-79 are the main subjects of this section.

The AFQT and Intelligence. An important initial question is whether the AFQT can truly be considered an intelligence test. Herrnstein and Murray (1994, 607) show that the AFQT test battery is highly *g*-loaded, with each subtest correlated at over 0.8 with *g*. Although this fact is not in dispute, some critics of Herrnstein and Murray have claimed that intelligence is not the only trait that the AFQT measures. According to Heckman's (1995, 1103) critique, the "AFQT is an achievement test.... Achievement tests embody environmental influences: AFQT scores rise with age and parental socioeconomic status."

All measures of cognitive ability, including the AFQT and full-scale IQ tests, show a substantial correlation with parental socioeconomic status (SES), but it does not follow that the tests are measuring achievement. Parental SES is not exogenous to the IQ of parent or child (Scarr 1997). In other words, genes that help determine the intelligence of both parent and child also affect the environment that the parent provides. We cannot say that high SES causes high test scores, because both could be independently caused by genes. To see this most clearly, imagine a world in which intelligence is 100% genetic, meaning children's IQ is determined entirely by genes and unaffected by environment. Since intelligent parents create better environments for their children, an SES correlation with children's IQ tests would still exist, even though we know by definition that SES does not cause higher IQ in this hypothetical world.

Although the positive correlation between AFQT and parental SES is inevitable, all IQ tests do have certain baseline requirements of education and mental maturity. The AFQT was designed for seventeen- and eighteen-year-olds who speak English and have taken algebra. As Neal and Johnson (1996, 890-891) have shown, age does not fully control for exposure to these

baseline requirements, because strict school-entry cutoff dates mean a student's grade level can be a full year less than another student of comparable age. To minimize this problem, I normalize the scores around "expected grade level" rather than age, using August 30 as the typical school entry date.[5]

Respondents Born Abroad (First Generation Immigrants). The NLSY-79 did not ask about citizenship status until 1990, when many of the original respondents were not sampled. Therefore, an immigrant in the NLSY is defined to be a foreign-born person with at least one foreign-born parent.[6] As the comparison group, I use non-Hispanic white natives, which avoids interpretive difficulties that arise from group test score differences among native ethnic groups.[7] Each subtest score is the residual of a weighted regression of the raw scores on

---

[5] More explicitly, a child's expected grade level is his age minus 5 if he was born between January 1 and August 30, and age minus 6 if born between September 1 and December 31.

[6] The requirement on the parent ensures that the foreign-born respondent was not simply born on an overseas military base to American parents, as several apparently were. Legally, whether or when a foreign-born child with one American-born parent and one non-American-born parent is an "immigrant" has changed repeatedly over the years (Weissbrodt and Danielson 2005, 411-418). If the stricter requirement of two foreign-born parents is imposed on immigrants, then the immigrant test score deficit is actually slightly larger than reported in this section.

[7] There are a few reasons for using whites as the comparison group. First, the racial and ethnic composition of the native population has changed dramatically since the 1960s, mostly as a result of immigration. If a substantial immigrant IQ deficit exists, it would be partially masked by comparing immigrants to a native population that contains lower-IQ second generation immigrants. Second, white IQ has been more stable over time. There is some evidence (see chapter 4) that black IQ scores have been rising relative to whites, at least through the 1970s. Measurements of the native-immigrant difference at different time periods would be affected by the instability of black IQ. Third, whites are the historical founding population. For better or for worse, most of America's institutional, political, and social culture is the product of European Americans, which makes them the natural standard by which immigrants might be compared.

expected grade level dummies. The subsequent group differences are expressed in standard deviations.[8]

Table 2.3

Unadjusted ASVAB Immigrant - White Native Differences (in SDs)

	White Native (N=6,560) subtracted from…				
Immigrant Group -->	All (N=684)	European (N=114)	Mexican (N=283)	Other Hispanic (N=199)	Asian (N=47)
General Science (GS)	-1.02	-0.50	-1.72	-1.02	-0.76
Automotive Information (AI)	-0.95	-0.45	-1.36	-1.10	-0.93
Mechanical Comprehension (MC)	-0.78	-0.27	-1.22	-0.90	-0.73
Electronics Information (EI)	-0.85	-0.25	-1.50	-0.95	-0.68
Numerical Operations (NO)	-0.49	-0.03	-1.15	-0.53	0.00
Coding Speed (CS)	-0.62	-0.13	-1.30	-0.66	0.10
Arithmetic Reasoning (AR)	-0.66	-0.24	-1.23	-0.68	-0.20
Mathematics Knowledge (MK)	-0.47	-0.12	-1.08	-0.43	0.10
Word Knowledge (WK)	-1.06	-0.52	-1.91	-0.83	-0.87
Paragraph Comprehension (PC)	-0.96	-0.48	-1.89	-0.78	-0.36
AFQT (AR+MK+WK+PC)	-0.88	-0.37	-1.72	-0.77	-0.35

Notes: Each group difference in the table is an immigrant group's average score minus the white native average score. Negative differences indicate a native advantage. Scores are normed to "expected grade level" at the the time of the test; see text for details.

Table 2.3 shows the raw results before any further adjustments are made. There are large differences between white natives and each immigrant group, with even European and

[8] The formula for calculating the difference in standard deviations between two groups is: $d = (\overline{X}_I - \overline{X}_N) / \sqrt{(N_I \sigma_I{}^2 + N_N \sigma_N{}^2)/(N_I + N_N)}$, where I represents immigrants and N is natives.

Asian immigrants performing poorly on the verbal tests. These results cannot be taken seriously, however, because the data need to be adjusted for several potential artifacts.

Statistical Adjustments: First, it is clear from the table that a significant language bias probably exists. Immigrants do comparatively worse on the verbal components of the AFQT, WK and PC, than they do on the math components, AR and MK. This pattern holds for each immigrant group. To analyze the situation more closely, separate AFQT Math and AFQT Verbal scores will be displayed in the next table. Those scores are calculated by averaging the two relevant raw score tests rather than all four. AFQT Math then becomes the main score of interest.

Though focusing the analysis on these two subtests helps to reduce language bias, it does introduce another problem, which is the comparability of the AFQT Math with a full-scale IQ score. As discussed in chapter 1, subtests have different correlations with *g*. If two groups primarily differ in general intelligence, their score differences will be smaller on tests with smaller *g*-loadings. Therefore, an estimated full-scale IQ is provided in the next table, calculated by dividing *d* by the *g*-loading of AFQT Math before conversion to the N(100, 15) scale (te Nijenhuis et al. 2004). Formally, full-scale IQ = 100 + $d/g * 15$. Obviously, this technique has limited usefulness when the test in question has a very low *g*-loading, but it provides a decent estimate of IQ when a full test battery is unavailable or unreliable.

The next adjustment addresses the problem of "give-ups" and random guessing. In 1980 the AFQT was a strictly paper-and-pencil test. Each test-taker was confronted with 105 multiple choice questions, with four possible answer choices in each question. Neal (2006) has pointed out a high number of zero or near-zero scores. Since there was no penalty for guessing, randomly filling in answers should have given the average guesser about 26 correct out of 105.

A quick application of the binomial theorem indicates that the chances of getting fewer than even 10 questions correct when randomly guessing on the AFQT is less than 1 in 10,000.

It is obvious that some combination of frustration or exhaustion caused some test-takers to give up, failing to even make random guesses. The result is that guessers and non-guessers, despite having essentially the same level of ability, get very different scores. To combat this problem, anyone getting fewer than one quarter of the answers correct in each subtest of the AFQT has his scored bumped up to one quarter of the total. Since those who have their scores raised are still ranked at the bottom of the distribution, the adjustment compresses the variance without changing rank order.[9]

The final adjustment on the AFQT test is for educational attainment. As discussed in the introduction to this section, the AFQT is a good IQ test, *assuming* the test-taker has the appropriate academic background. Unlike purely abstract intelligence tests like Ravens' Matrices, the AFQT assumes a basic knowledge of English and algebra at an early high school level. The AFQT cannot be a particularly good measure of IQ when the person taking the test does not have that basic knowledge. So why not simply control for grade level rather than "expected grade level"? The reasoning behind using expected grade level is that a person's intelligence is strongly correlated with educational attainment. Smarter people are likely to stay in school longer. If AFQT scores are normed to actual grade level, an 18-year-old who dropped out after

[9] One problem that cannot be directly addressed is that AFQT questions, unlike those on the SAT, were not ordered by difficulty in each section. The thinking behind the SAT ordering is that if someone gives up halfway into the test because the questions are too hard, it is highly unlikely that person would have answered any of the later (harder) questions correctly even if he was trying. There is no such protection on the AFQT from give-ups. Someone who gives up could be skipping over very easy questions. The adjustment described above equalizes the scores of guessers and non-guessers, but nothing can be done about a person who starts guessing blindly in the middle of the test. If one group has less ability than another, the poorer performing group might be more likely to give up in the middle out of frustration, thus causing the group difference to appear larger than it is. That being said, there cannot be a "give up" bias without an actual group difference in the first place.

tenth grade would be compared against a 16-year-old tenth grader rather than his own peers. This would artificially raise his IQ.

One could think of adjusting for educational attainment as having the same problems as "controlling for occupational status." Doctors are surely smarter on average than truck drivers, and we would want any good IQ test to reveal that difference.[10] But comparing doctors against doctors and truck drivers against truck drivers would have the effect of throwing out all the variation across occupations. In much the same way, controlling for educational attainment compresses the IQ distribution, eliminating important differences between grade levels. However, *not* controlling for education can inaccurately widen the variance in IQ scores by comparing academically prepared people with those who are not. People may drop out of school for a variety of reasons, only one of which may be low intelligence. Consider the counterfactual situation in which the average high school dropout actually stays in school for another year. He will not do as well as his peers on the AFQT, but he will probably do somewhat better than he would have as a dropout.

Thus, we have a situation in which controlling for education makes IQ differences too small, and not controlling for education makes differences too large. In this situation, simply using a different IQ test, one with a lower knowledge requirement, is usually the best option, but that is not possible here. Since the purpose of this chapter is to demonstrate an immigrant IQ deficit, it is better to bias the results against that conclusion; if the deficit still remains, the conclusion is strengthened. Therefore, the adjusted NLSY results are controlled for educational attainment, not merely for expected grade level, but with one exception—educational attainment is top-coded at 12 years. The AFQT does not require any college-level knowledge.

---

[10] See Gottfredson (1986) for an interesting analysis of IQ and occupation.

Table 2.4

Immigrant - White Native ASVAB Group Differences (in SDs)

	White Native (N=6,528) subtracted from….				
Immigrant Group -->	All (N=619)	European (N=111)	Mexican (N=228)	Other Hispanic (N=193)	Asian (N=46)
General Science (GS)	-0.76	-0.47	-1.06	-0.91	-0.49
Automotive Information (AI)	-0.72	-0.42	-0.76	-0.96	-0.80
Mechanical Comprehension (MC)	-0.57	-0.26	-0.71	-0.79	-0.55
Electronics Information (EI)	-0.60	-0.24	-0.86	-0.82	-0.50
Numerical Operations (NO)	-0.22	0.00	-0.54	-0.43	0.39
Coding Speed (CS)	-0.34	-0.11	-0.63	-0.54	0.41
Arithmetic Reasoning (AR)	-0.44	-0.23	-0.74	-0.60	0.08
Mathematics Knowledge (MK)	-0.25	-0.09	-0.63	-0.33	0.41
Word Knowledge (WK)	-0.78	-0.52	-1.18	-0.71	-0.62
Paragraph Comprehension (PC)	-0.70	-0.47	-1.28	-0.68	0.02
AFQT Math (AR+MK)	-0.36	-0.17	-0.72	-0.49	0.26
AFQT Verbal (WK+PC)	-0.80	-0.54	-1.34	-0.76	-0.31
AFQT (AR+MK+WK+PC)	-0.62	-0.37	-1.09	-0.67	0.00
Full-Scale IQ (estimated from AFQT Math)	**93.3**	**96.9**	**86.9**	**91.1**	**105.1**

Notes: Each group difference in the table is an immigrant group's average score minus the white native average score. Negative differences indicate a native advantage. Scores are normed to highest grade completed, topcoded at 12 years; see text for details.

Results: The adjusted results are shown in table 2.4 above. Asians outscore natives, Europeans score slightly below natives, and Mexicans and other Hispanics score well below natives. The overall immigrant IQ estimate is 93.3. Group differences are slightly smaller in

most cases, owing to the adjustments described above. The full-scale IQ estimates, derived from the AFQT Math scores, are similar to the LV data.

The addition of separate math and verbal AFQT scores brings the possibility of language bias into better focus. Relative to native whites, immigrants of all backgrounds do significantly better on the mathematics sections than on the verbal sections. The immigrant math-verbal differences on the AFQT suggest that non-native speakers are at a disadvantage. How large is this disadvantage? The overseers of the NLSY will not release individual AFQT question data, so we cannot know the degree of bias with much certainty. However, what bias exists is not likely to change the primary conclusion derived from these data—immigrants have lower IQs than white natives. The immigrants in the NLSY are not "just off the boat." They immigrated at a young age and attended American school for varying numbers of years before taking the AFQT. Only 85 Hispanics requested the optional Spanish language instructions, and Hispanics with the least English proficiency are likely not to have participated at all (Bock and Moore 1986, 171 and 73). Moreover, the fact that immigrants, and Mexicans in particular, still lag far behind natives on mathematics tests, even when controlling for years of education, suggests that a substantial IQ deficit exists, even if it cannot be estimated precisely.

The Psychometric Properties of Results for the First Generation: Because the ASVAB is a battery of several varied cognitive tests, it is possible to analyze its factor structure and isolate the impact of *g* on each subtest. The purpose is to determine whether the ASVAB's factor structure is the same for immigrants and natives, and then to analyze the degree to which *g* itself is responsible for the subtest variation in group differences. Table 2.5 shows the results of a principal factor analysis of the adjusted test results for natives and for each immigrant group. The first principal factor is *g*, the general intelligence factor that accounts for the largest proportion of score

variance on a good IQ test. The ASVAB is highly *g*-loaded, as *g* explains most of the subtest score variance for each group, with the exception of European scores on Coding Speed.

The *g*-loadings of the individual subtests are, with a few exceptions, similar for each group. The congruence coefficient, a type of correlation measure, is a formal measure of factor similarity. A congruence greater than 0.95 indicates that the factor structures are the same (Jensen 1998, 374). The coefficient of congruence of white native factor structure with each immigrant group's structure is given in the second to last row of the table. All are uniformly high. Given the similarity of factor structure, it may be concluded that the ASVAB functions as an IQ test in the same manner for immigrants as it does for natives. If a large language or cultural bias were affecting immigrant scores, the explanatory power of the *g* factor would be attenuated.

The next step is to examine whether it is variation in *g* that explains the various group differences reported on each subtest. Jensen (1998) has repeatedly confirmed what he calls "Spearman's hypothesis," the prediction that white-black differences on IQ tests will be greatest on the most *g*-loaded tests. The implication is that the group differential reflects a difference in general ability rather than merely test-specific factors The same hypothesis can be tested here on the native-immigrant difference.

The technical procedure is described in detail in Appendix B, but the sense of the method is to correlate the group differences and *g*-loadings on each subtest. A high, statistically significant correlation is confirmation of the hypothesis. Table 2.5 lists the correlations for each immigrant group along with tests of significance. The results are ambiguous. All the correlations, except in the Other Hispanic category, are positive and moderately large, but none exceed the 0.56 threshold for statistical significance at the 95% level.

Table 2.5

ASVAB Subtest *g*-Loadings by Immigrant Group

Test		Immigrant Group				
	White Natives	All Immigrants	European	Mexican	Other Hispanic	Asian
General Science (GS)	0.8094	0.865	0.8746	0.8582	0.8728	0.8487
Automotive Information (AI)	0.5352	0.6842	0.666	0.7023	0.7308	0.6253
Mechanical Comprehension (MC)	0.7171	0.7541	0.7407	0.8179	0.7194	0.8099
Electronics Information (EI)	0.7217	0.7878	0.7672	0.7808	0.8105	0.8123
Numerical Operations (NO)	0.5497	0.5778	0.5333	0.6274	0.5996	0.4113
Coding Speed (CS)	0.4185	0.4995	0.2911	0.5993	0.6471	0.1294
Arithmetic Reasoning (AR)	0.8398	0.8179	0.8131	0.8184	0.838	0.6874
Mathematics Knowledge (MK)	0.7898	0.7968	0.7915	0.8461	0.8349	0.7421
Word Knowledge (WK)	0.7849	0.8126	0.8478	0.8585	0.8131	0.8394
Paragraph Comprehension (PC)	0.692	0.7411	0.6866	0.7772	0.7777	0.7162
congruence coefficent: factor similarity with white natives	-	0.998	0.996	0.996	0.994	0.984
Spearman correlation between *g*-loadings and group differences	-	0.45	0.42	0.37	0.16	0.52

Notes: The congruence coefficient (a type of correlation) measures the similarity of subtest *g*-loadings on the ASVAB between white natives and the immigrant comparison group. The Spearman correlation measures the relationship between the subtest *g*- loadings and the absolute value of the immigrant-native group differences given in the previous table. Significance levels of insignificant correlations are not shown.

The test of significance for a rank-order correlation is quite stringent, as it depends only on the number of subtests in the battery. The best interpretation of these results is that subtest differences have some *g*-component for all groups except non-Mexican Hispanics. Nevertheless, the varying language requirements on the subtests, which would make some subtest differences larger than predicted by their *g*-loadings, is probably masking the full effect of *g*. Spearman's hypothesis will be revisited with second generation immigrants in the next section.

NLSY Respondents Who Were the Children of Immigrants (Second Generation Immigrants). The previous sections have shown significant native-immigrant score differences on the ASVAB, due in part to actual differences in intellectual ability rather than language or cultural biases. The next question is whether subsequent generations of immigrants in the NLSY show the same cognitive deficit. Since parent and child IQ are positively correlated, the children of low-IQ immigrants are likely to be below average as well. However, perhaps there is an indirect, environmentally-driven positive effect on IQ scores from living in the U.S.

Recall the Flynn effect from chapter 2, which describes how IQ scores have gone up consistently since World War II, at least until recently, while *g* likely has not. If the Flynn effect, or something like it, has been inflating native scores independent of *g*, the scores of recent immigrants may not get the same cumulative boost. With the Flynn effect leaving them behind, immigrants could score lower than natives, even on a completely culture-fair test, without differing from natives nearly as much in *g*. Since the Flynn effect itself does not yet have a widely accepted explanation, this kind of ad hoc explanation for low immigrant IQ does not have much of a theoretical basis. Nevertheless, the theory can be tested by examining second generation immigrant IQ scores broken down by ethnic origin. Do second generation immigrants, born and raised in the U.S., close the gap with white natives?

As mentioned earlier, an immigrant is defined for NLSY purposes as someone who was born in a foreign country and has at least one foreign-born parent. A second generation immigrant was born in the U.S. but has at least one parent who was born elsewhere.[11] A third generation or higher immigrant, which I designate as the "3+ generation," is native-born and has

[11] The stricter definition of second generation, born in the U.S. with *both* parents born abroad, results in a rather small number of observations in the NLSY, partially due to missing parent birth data. If the stricter definition is used anyway, second generation IQ is slightly lower.

parents who were both born in the U.S. This section looks at the second and 3+ generation immigrants in the NLSY.

It is important to make clear that these second generation immigrants are not the children of the immigrants who were previously examined. They are the same age as NLSY immigrants, but they were born in the U.S. Because of their American roots, the NLSY second generation respondents provide some clues about how immigrants may perform on the AFQT with the benefits of an American upbringing, including an earlier and more immersive English experience.

Table 2.6 shows the difference between 3+ generation whites and second and 3+ generation immigrants by ethnic origin. The second and 3+ generation samples also present another opportunity to test Spearman's hypothesis; the results appear in table 2.7.

Despite going down substantially, the Mexican and other Hispanic IQ deficits are still quite large. The difference between Hispanic math and verbal scores is now much smaller, suggesting that language bias has been mitigated. But even with an American upbringing, Hispanics still lag behind native whites. Furthermore, third generation Mexican and other Hispanic IQ is actually lower than the second generation. (European 3+ generation "immigrants" are not included because they cannot be distinguished from the native white control group.) There is no evidence here that Hispanic IQ will converge with whites. In fact, with less distortion due to language difficulties, the *g* component of Hispanic IQ differences with whites becomes much more evident. Even though the deficits are smaller, the correlations of *d* and *g* are larger and more significant for Mexicans in the second and 3+ generations compared to the first. Non-Mexican Hispanics differences are still not related to *g* in the second generation, but the 3+ generation, which features a much larger sample of Hispanics, does show a strong relationship.

Table 2.6

ASVAB Ethnic Group Differences by Immigrant Generation (in SDs)

	Second Generation Immigrants				3+ Generation Immigrants	
	3+ White Native (N=6,106) subtracted from…				3+ generation White Native (N=6,106) minus…	
"Immigrant" Group -->	All (N=736)	European (N=277)	Mexican (N=291)	Other Hispanic (N=108)	Mexican (N=435)	Other Hispanic (N=482)
General Science (GS)	-0.11	0.12	-0.87	-0.21	-0.86	-0.74
Automotive Information (AI)	-0.17	-0.03	-0.50	-0.30	-0.64	-0.64
Mechanical Comprehension (MC)	-0.11	0.07	-0.66	-0.24	-0.71	-0.63
Electronics Information (EI)	-0.15	0.06	-0.80	-0.14	-0.85	-0.65
Numerical Operations (NO)	-0.12	-0.01	-0.48	-0.11	-0.41	-0.64
Coding Speed (CS)	-0.02	0.04	-0.19	-0.07	-0.23	-0.39
Arithmetic Reasoning (AR)	-0.15	0.02	-0.68	-0.21	-0.77	-0.68
Mathematics Knowledge (MK)	-0.05	0.09	-0.55	-0.08	-0.70	-0.55
Word Knowledge (WK)	-0.16	0.06	-0.84	-0.22	-0.86	-0.81
Paragraph Comprehension (PC)	-0.15	0.01	-0.68	-0.22	-0.68	-0.73
AFQT Math (AR+MK)	-0.11	0.06	-0.65	-0.15	-0.77	-0.65
AFQT Verbal (WK+PC)	-0.17	0.03	-0.82	-0.24	-0.83	-0.83
AFQT (AR+MK+WK+PC)	-0.15	0.05	-0.79	-0.21	-0.88	-0.80
Full-Scale IQ (estimated from AFQT Math)	**98.0**	**101.2**	**87.8**	**97.2**	**85.6**	**88.2**

Notes: A second generation immigrant was born in the US to at least one parent who was foreign-born. A 3+ generation person is a native with two native parents. Each group difference in the table is a second or 3+ generation "immigrant" group's average score minus the 3+ white native average score. Negative differences indicate a "native" advantage. Scores are normed to highest grade completed, topcoded at 12 years; see text for details.

Despite the lagging scores of Hispanics, overall the second generation is much closer in IQ to native whites than the first generation, and Europeans have closed the gap entirely. All three ethnic groups—there were too few Asians in the second and 3+ generations—make gains. Does this mean the second generation always improves drastically? Maybe, but remember the caveat from a previous paragraph. The difference between the second generation and the actual immigrants is that the second generation had parents who immigrated earlier enough so that their children were born in the U.S. If both generations are of similar ability and background,

the second generation may be a good indicator of how successful the actual immigrants' children will be.

Table 2.7

ASVAB Subtest *g*-Loadings by 2nd and 3+ Generation Group

Test		2nd Generation Group				3+ Generation Group	
	3+ White Natives	All	European	Mexican	Other Hispanic	Mexican	Other Hispanic
General Science (GS)	0.8071	0.8672	0.8421	0.8427	0.836	0.7985	0.859
Automotive Information (AI)	0.535	0.5996	0.5472	0.5987	0.7109	0.6169	0.6742
Mechanical Comprehension (MC)	0.7152	0.781	0.7569	0.681	0.7519	0.6925	0.7757
Electronics Information (EI)	0.7228	0.7648	0.7089	0.7482	0.8063	0.7562	0.7669
Numerical Operations (NO)	0.5479	0.574	0.584	0.5558	0.5746	0.5438	0.6595
Coding Speed (CS)	0.4107	0.4909	0.5248	0.3946	0.3511	0.3707	0.5334
Arithmetic Reasoning (AR)	0.8425	0.8363	0.8102	0.7923	0.8534	0.8155	0.854
Mathematics Knowledge (MK)	0.7903	0.8059	0.7868	0.7566	0.7995	0.7592	0.7985
Word Knowledge (WK)	0.7825	0.8323	0.8209	0.7935	0.8074	0.8205	0.8442
Paragraph Comprehension (PC)	0.6877	0.7397	0.7313	0.7108	0.6601	0.7051	0.7687
congruence coefficent: factor similarity with 3+ whites	-	0.999	0.998	0.999	0.997	0.999	0.997
Spearman correlation between *g*-loadings and group differences	-	0.13	0.45	0.79***	0.05	0.66**	0.62*

Notes: The congruence coefficient (a type of correlation) measures the similarity of subtest *g*-loadings on the ASVAB between white natives and the second and 3+ generation comparison group. The Spearman correlation measures the relationship between the subtest *g*- loadings and the absolute value of the "immigrant"-native group differences given in the previous table. Significance levels of insignificant correlations are not shown.

However, the assumption that each generation is comparable is dubious. NLSY respondents were born between 1957 and 1964, and immigration policy was changed to favor lower-skill immigrants after 1965. Approximately 75% of NLSY immigrants came to the U.S. after 1965, meaning the difference between the first and second generation may just reflect changes in policy rather than intergenerational intelligence gains. A better way to examine how immigrant IQ scores change over time is to examine the actual children of the immigrants in the NLSY-79.

Children of NLSY First Generation Immigrants. Since 1986 the biological children of NLSY-79 respondents have been profiled on a biennial basis, allowing researchers to examine how the socioeconomic characteristics of one generation pass on to the next. The NLSY Children dataset contains several cognitive measures, including Peabody Individual Achievement Tests in math, reading comprehension, and reading recognition, the Peabody Picture Vocabulary Test (PPVT), and the digit span from the WISC-R. Completion rates for these tests have ranged from about 85% to 95% in any given year. Many of the same children were eligible for testing in multiple years, meaning some children who were missed in one wave have valid scores in another. When multiple scores are reported for an individual, the median is used. All scores are age-adjusted.[12]

Table 2.8 shows test score differences between the children of the white natives in the NLSY-79 and the children of the immigrants.[13] The results are similar to the second generation immigrants from the previous section. The children of European immigrants score higher than the children of white natives, while the children of Mexican and other Hispanic immigrants score much lower. Mexicans and other Hispanics score especially poorly on the PPVT, but this is probably due to many of the children speaking only Spanish at home. Since the PPVT was given to children as young as three, a language bias is probably inflating the difference, although many of the children with language barriers were not tested. The most informative score is on the math test, in which second generation Mexicans and other Hispanics trail whites by almost as much as their parents did on the AFQT Math.

---

[12] There is no need to adjust for education, because almost all of the children are too young to have dropped out of school.

[13] Note that the ethnic origins in the table are determined by the mother's ethnicity given in the NLSY-79, not the child's ethnicity. The distinction makes very little difference in the results.

Table 2.8

Achievement Test Group Differences (in SDs):
Children of Immigrants minus Children of White Natives

	Children of White Natives minus…			
Children of Immigrant Group -->	All	European	Mexican	Other Hispanic
Peabody Math (ages 5-14)	-0.45 (N=509)	0.04 (N=45)	-0.83 (N=287)	-0.45 (N=140)
Peabody Picture Vocabulary (ages 3-18)	-0.84 (N=524)	0.15 (N=48)	-1.43 (N=297)	-1.22 (N=142)
Peabody Reading Comprehension (ages 5-14)	-0.33 (N=488)	0.27 (N=42)	-0.71 (N=270)	-0.45 (N=139)
Peabody Reading Recognition (ages 5-14)	-0.20 (N=509)	0.24 (N=45)	-0.57 (N=286)	-0.24 (N=141)
Digit Span (age 7+)	-0.24 (N=474)	0.26 (N=37)	-0.63 (N=271)	-0.09 (N=130)
Full-Scale IQ (estimated from Peabody Math)	**91.5**	**100.8**	**83.7**	**91.5**

Notes: Each group difference in the table is an immigrant group's average score minus the white native average score. Positive differences indicate an immigrant advantage. Scores are normed to age. The number of cases in the white native comparision group are, from top to bottom, 3246, 3302, 3145, 3248, and 3023.

Conclusion. In summary, there are substantial native-immigrant differences on the ASVAB, including the highly *g*-loaded AFQT. The differences are largest for Mexicans and other Hispanics, and they are smaller for Europeans, consistent with the LV data. In the second and third generations, the native-European difference on the AFQT either goes away or switches sign, but Hispanics still trail native whites by a considerable margin. Assessing the degree of language bias on the ASVAB subtests is an imprecise science, because individual

question data are not available to be examined. However, there are four reasons to believe that real intelligence differences are responsible in large part for the differences in test scores. First, most of the immigrants in the NLSY are young people who have attended American schools. Second, natives score well above immigrants on mathematics tests, even when controlling for years of education. Third, factor analysis shows that the *g*-loadings of the subtests are essentially the same for immigrants and natives. Fourth, there is a positive correlation between subtest *g*-loading and native-immigrant *d* for most ethnic groups.

PIAT-R MATH FROM THE NLSY-97

A new NLSY sample was selected in 1997. The NLSY-97 is similar in design and content to its predecessor, and it includes the results of a computerized version of the AFQT. Initial results from the 1997 AFQT appear to show the immigrant-native difference at about one quarter of a standard deviation, but severe non-response bias makes the result impossible to interpret. In 1980, 94% of respondents took the AFQT, and the NLSY contains a special weight to correct for what little non-response bias existed. However, in 1997 over 20% of the sample chose not to participate. Non-responders included 29% of immigrants, and 33% of Hispanic immigrants. A comparison of test-takers with non-test-takers reveals significantly lower parental SES in the latter category. At this time, no adequate weight exists to adjust for this problem.

The interpretable test scores from the NLSY-97 come from the revised Peabody Individual Achievement Test in Mathematics (PIAT-R Math), a test similar to the mathematics knowledge subtest of the AFQT, with a *g*-loading of 0.70 (Markwardt 1998, 73). Unlike the AFQT in 1997, the PIAT-R received a good response rate of over 95% of the targeted sample. Table 2.9 compares the scores of natives and immigrants who are matched on education.

Table 2.9
Immigrant - White Native Differences on
1997 PIAT-R Math

Immigrant Group	Initial *d*
All (N=706)	**-0.39**
European (N=78)	0.09
Mexican (N=343)	-0.92
Other Hispanic (N=188)	-0.42
Asian (N=60)	0.14

Notes: All scores are adjusted for educational attainment. The comparison group is 2,837 white natives.

These results show a pattern similar to the AFQT Math in 1980—a substantial IQ deficit, with Mexican immigrants exhibiting the largest difference with white natives. There were too few Asian immigrants in NLSY-79 to meaningfully evaluate, but here they slightly outperform white natives, as do European immigrants.

As was the case with the AFQT for the NLSY-79, potential biases must be examined. Unlike the AFQT, the PIAT-R can be analyzed question-by-question thanks to new data released in 2008. Individual questions can be assessed by checking for differential item functioning (DIF), a general term meaning group differences that are independent of the ability measured by the test.[14]

Checking for DIF. An item is a single question on a test. When two groups perform differently on a particular item, psychometricians do not automatically assume the item is biased, because the performance difference could be due to underlying ability differences between the two groups. To check for true item bias, groups must first be matched on ability. If a

[14] Bias, which connotes an *unfair* advantage for one group (Donoghue and Allen 1993), is actually a subset of DIF.

significant group performance difference still exists on the item, then the item may be said to exhibit DIF.

Psychometricians have developed several advanced techniques to detect DIF. One of the more popular is the Simultaneous Item Bias test (SIBTEST) procedure (Shealy and Stout 1993), which I use here. Each test subject is assigned an overall ability level θ based on his total score on the PIAT-R Math, which contains 100 items. SIBTEST compares the probability of a correct answer on a given item by the reference group (white natives) versus the probability for the focal group (immigrants), when each group is matched on θ. For each item *i*, this difference B_i is given by

$$B_i(\theta) = P_{Ri}(\theta) - P_{Fi}(\theta),$$

where P is a probability and R and F indicate the reference and focal groups, respectively. The total theoretical DIF β_i is B_i weighted according to ideally-smooth distributions of ability in the reference and focal groups. SIBTEST uses the estimator $\hat{\beta}_i$ to approximate β_i based on the actual number of reference and focal group members at each ability level. Conceptually, $\hat{\beta}_i$ is the observed advantage in probability of a correct answer on item *i* for the reference group over the focal group when ability levels are matched. The null hypothesis tested for each item is $\beta_i = 0$.

One of the strengths of SIBTEST is that it provides both a test of the significance of the DIF (based on the asymptotically normal distribution of $\hat{\beta}_i$) *and* a measure of its magnitude. Roussos and Stout (1996) adapted a system used by the Educational Testing Service to classify the severity of DIF on each item. An "A-level" item has significant DIF but with inconsequential magnitude ($\left|\hat{\beta}_i\right| < 0.059$). A "B-level" item has significant DIF, but its

magnitude is within a specific range ($0.059 \leq \left|\hat{\beta}_i\right| \leq 0.088$) that makes it moderately acceptable if no other items are available. The least desirable item is "C-level," which has DIF that is both statistically significant and large ($\left|\hat{\beta}_i\right| > 0.088$).

SIBTEST Results. Individual SIBTEST runs were performed for each immigrant subgroup and for immigrants as a whole. Table 2.10 shows both the significance and magnitude of bias on the PIAT-R Math items, where the reference group is white natives and the focal group is Mexican immigrants, who experienced the greatest amount of DIF of any subgroup. When the DIF reaches statistical significance, the item is classified as A-, B-, or C-level, in accordance with the rules set out above.

Theoretically, some items could be biased against white natives. Whenever two groups of substantially different backgrounds are compared, each will likely have some built-in advantages, even if one group has many more than the other. Immigrants who speak Spanish may be advantaged on certain items that use difficult English words with close Spanish cognates (Schmitt 1988), for example. However, the purpose here is to determine whether bias against immigrants explains part of the test score deficit with white natives. Therefore, all of the significance tests are one-tailed. This makes each item more likely to be flagged for bias against immigrants, and it effectively disregards any DIF against natives as statistical noise.

As the table indicates, there was enough variation in scores to find a meaningful $\hat{\beta}_i$ on 84 of the 100 items. Of those 84 items, 10 items showed statistically significant DIF. However, 9 of those items were A-level, meaning negligible in magnitude. Only item number 64 showed large DIF. The same analysis performed on the other immigrant subgroups showed even less DIF. This indicates that the PIAT-R Math is free of any large internal bias against immigrants.

Table 2.10

Analysis of DIF with SIBTEST: White Natives versus Mexican Immigrants

Item	Beta-hat	Std. Error	p-value	DIF Level	Item	Beta-hat	Std. Error	p-value	DIF Level
8	-0.001	0.001	0.825		59	-0.012	0.016	0.770	
9	-0.001	0.001	0.821		60	0.003	0.015	0.415	
14	-0.001	0.001	0.821		61	-0.017	0.017	0.839	
19	-0.001	0.001	0.730		62	0.008	0.018	0.328	
20	**0.002**	**0.001**	**0.022**	**A**	63	0.000	0.016	0.504	
22	0.000	0.001	0.328		**64**	**0.094**	**0.017**	**0.000**	**C**
23	-0.001	0.001	0.813		65	0.000	0.020	0.499	
24	0.001	0.001	0.217		66	-0.019	0.020	0.829	
25	0.000	0.002	0.477		67	-0.015	0.018	0.796	
26	0.002	0.002	0.252		68	-0.009	0.023	0.655	
27	0.002	0.002	0.086		69	0.015	0.013	0.130	
28	-0.002	0.002	0.820		70	-0.053	0.021	0.994	
29	0.001	0.002	0.307		**71**	**0.046**	**0.021**	**0.015**	**A**
30	-0.002	0.002	0.779		72	-0.003	0.014	0.579	
31	**0.005**	**0.003**	**0.037**	**A**	73	-0.006	0.026	0.597	
32	**0.007**	**0.003**	**0.006**	**A**	74	-0.007	0.023	0.619	
33	0.003	0.003	0.157		75	-0.004	0.023	0.567	
34	-0.001	0.003	0.636		76	-0.067	0.021	0.999	
35	**0.013**	**0.006**	**0.010**	**A**	77	0.005	0.020	0.404	
36	-0.001	0.004	0.621		78	-0.032	0.021	0.939	
37	0.003	0.004	0.249		79	-0.026	0.026	0.842	
38	0.003	0.004	0.210		**80**	**0.040**	**0.020**	**0.025**	**A**
39	0.008	0.005	0.057		81	-0.005	0.024	0.579	
40	-0.007	0.006	0.862		82	0.010	0.020	0.301	
41	0.008	0.008	0.147		83	0.024	0.022	0.139	
42	-0.004	0.002	0.942		84	-0.026	0.019	0.911	
43	-0.006	0.006	0.813		85	0.031	0.024	0.095	
44	0.009	0.007	0.098		86	-0.030	0.026	0.878	
45	-0.009	0.006	0.926		87	0.007	0.029	0.403	
46	0.002	0.008	0.384		88	-0.005	0.033	0.562	
47	**0.012**	**0.006**	**0.017**	**A**	89	0.025	0.029	0.196	
48	-0.009	0.007	0.908		90	-0.037	0.027	0.914	
49	-0.009	0.011	0.783		91	0.024	0.019	0.101	
50	-0.011	0.012	0.804		92	-0.003	0.026	0.539	
51	-0.012	0.010	0.893		93	-0.032	0.023	0.913	
52	0.009	0.009	0.165		94	-0.011	0.028	0.649	
53	-0.018	0.011	0.945		**95**	**0.033**	**0.019**	**0.041**	**A**
54	-0.003	0.012	0.583		96	0.005	0.020	0.406	
55	0.013	0.012	0.157		97	-0.042	0.026	0.949	
56	-0.005	0.014	0.643		**98**	**0.024**	**0.010**	**0.010**	**A**
57	-0.022	0.015	0.928		99	0.003	0.009	0.392	
58	-0.003	0.006	0.708		100	-0.020	0.011	0.963	

Notes: Positive values of beta-hat indicate bias against Mexican immigrants. The p-values are one-tailed. Items not appearing in the table had too little variation between groups to generate meaningful data. "A" is neglible DIF, "B" is moderate, and "C" is large; see text for details.

Adjusted Scores. But how much do the observed DIF items affect total scores? The question can be answered by eliminating the biased questions and recalculating total scores.

Table 2.11 shows immigrant–white native differences in SDs on the PIAT-R Math both before and after the DIF items, even the A-level items, are eliminated. The unadjusted results show a pattern similar to the AFQT Math in 1980—a substantial IQ deficit, with Mexican immigrants exhibiting the largest difference with white natives. After the bias adjustment there is very little difference in scores. The immigrant–white native difference moves only from -0.39 SDs to -0.38. The observed DIF on the eliminated items is not large enough to meaningfully affect group differences. These results confirm what was asserted in the AFQT section—there is some detectable bias against immigrants on standardized tests, but it is not nearly large enough to nullify the IQ deficit observed.

Table 2.11

Immigrant - White Native Differences on 1997 PIAT-R Math With Bias Adjustment

		number of deleted items at…				
Immigrant Group	Initial *d*	A-level	B-level	C-level	Bias-adjusted *d*	Full-Scale IQ
All (N=706)	**-0.39**	**7**	**1**	**0**	**-0.38**	**91.9**
European (N=78)	0.09	2	2	1	0.10	102.2
Mexican (N=343)	-0.93	9	0	1	-0.91	80.5
Other Hispanic (N=188)	-0.42	4	1	1	-0.40	91.3
Asian (N=59)	0.12	0	0	1	0.12	102.6

Notes: The bias adjustment is an elimination of test items that fail the SIBTEST criterion for non-bias. There were 100 items on the test initially. All scores are adjusted for educational attainment. The comparison group is 2,837 white natives.

Full-scale IQs are equivalent to $100 + d/0.7 + 15$, since the *g*-loading of the Peabody Math is 0.7. The approximate IQ scores from the Peabody show the same pattern as the AFQT, though Europeans score somewhat higher on the Peabody compared to the AFQT, and Mexicans score somewhat lower.

Some Caveats. Although the SIBTEST procedure is one of the more popular methods of DIF detection, it is not perfect. Like all internal validity checks, it can detect only bias that

varies from item to item. If there were a uniform bias affecting every item identically, SIBTEST would not see it. This could be a problem on a test of immigrant verbal skills, where lack of English knowledge could conceivably push down immigrant scores compared to native scores, even as the relative difficulty of each item remains the same for both groups. However, this is far less likely on a math test, in which the verbal content of an item is unrelated to the difficulty of the mathematical concept being tested. When language bias affects a math test, its impact will almost certainly vary by item.

SIBTEST can also be used to test bundles of items at one time for DIF (Douglas et al. 1996), rather than just individual items as in this section. The theory is that undetectable bias at the item level may be amplified and significant at the bundle level. Unfortunately, evaluation of every possible bundle on a 100-item test is not feasible. Without the text of the items on the PIAT-R Math, it is not possible to argue even informally that certain bundles are more suspect than others. Nevertheless, a preliminary investigation of some bundles—e.g., the first quarter of the test—has not revealed anything substantial.

DIGIT SPAN FROM THE 2003 NEW IMMIGRANT SURVEY

The New Immigrant Survey (NIS) collects detailed information from a representative sample of legal and newly-arrived immigrant families, including over 2,000 children. Although the children were given several cognitive tests, only one is clearly free of culture and language bias—the digit span test.

Digit Span and Intelligence. Digit span is administered in two parts, forward and backward. Forward digit span is essentially a test of memory. The tester reads aloud a sequence of digits, and the subject must repeat back the sequence in order. Forward digit span is not highly *g*-loaded—it requires little more than verbal repetition and short-term memory. The backward digit span, however, has a significantly higher *g*-loading (Prokosch et al. 2005). A

quick self-test should make it easy to understand why repeating a sequence backward is much more mentally taxing, and hence more *g*-loaded, than repeating it forward. The backward digit span requires the subject to memorize the sequence in order, and to keep that order in short-term memory while manipulating and verbalizing the reverse sequence. It is a deceptively difficult task. The average adult can repeat about 7 digits forward but only 5 digits backward (Jensen 1998, 263n22).

This section will consider only the results from backward digit span, since it taps into *g* more effectively than the forward span. However, it should be emphasized that digit span tests—whether forward, backward, or combined—are not stand-alone measures of intelligence. The combined digit span's overall *g*-loading of 0.47 for children means that it is a useful but rough approximation of intelligence (Kaufman 1979, 110). Its major virtue is its lack of cultural content. It requires only that subjects are familiar with the digits from one to nine. Because of its simplicity and cultural neutrality, the digit span has been used for, among other things, predicting entrepreneurial ability in poor countries (Djankov et al. 2005; de Mel et al. 2007). Even language is not an issue here, because the NIS conducted the digit span tests in the preferred language of the immigrant children, with seemingly no limits on exoticism. In fact, three children were read numbers in Amharic, an Ethiopian dialect.

NIS Respondents Born Abroad (First Generation Immigrants). The NIS uses the version of the digit span from the revised Wechsler Intelligence Scale for Children (WISC-R), which was standardized in 1972. It is the successor to the original 1949 WISC, but since then both the 1991 WISC III and the 2003 WISC IV have become available. The Flynn effect has little impact on digit span scores (see Appendix B), but it is still advisable to compare immigrants to native norms that are as recent as possible. The backward portion of the digit span is

administered slightly differently in the WISC-IV, which means the most appropriate normative sample of natives comes from the WISC-III.

For each age level, Wechsler (1991) gives the mean and standard deviation of the longest string of digits that could be repeated backward by a cross-section of American children, including non-whites, in 1991. The immigrants from the NIS are compared to those standards in table 2.12. The first column shows the immigrant-native *d*, where the native comparison group includes both whites and non-whites. The second column gives an estimated full-scale IQ score for each ethnic group based on *d*. Each *d* is divided by the correlation of backward digit span with *g*, which is approximately 0.5 (Jensen 1985, 208). The larger *d* is then converted to the standard scale used in this chapter, with an average American whole-population IQ of 98. The following formula illustrates the calculation used: $IQ = 98 + 15 * d / 0.5$.

Table 2.12

Immigrant - Native Digit Span Group Differences

Immigrant Group	N	Proportion of sample	*d*	Full Scale IQ estimate
Europe	119	12.3%	0.04	99.1
Northeast Asia	56	5.8%	0.26	105.8
Southeast Asia	96	9.9%	0.21	104.4
South Asia (India)	72	7.4%	0.46	111.9
Sub-Saharan Africa	54	5.6%	-0.30	89.0
Mexico	106	10.9%	-0.52	82.4
Central America / Caribbean	96	9.9%	-0.51	82.6
South America	41	4.2%	-0.39	86.3
All	**971**	**100.0%**	**-0.16**	**93.3**

Notes: Each group difference is the immigrant mean minus the native mean. Positive differences indicate an immigrant advantage. Natives include all races, not just whites. Regional groups with fewer than 40 people are not shown but are included in the total.

The results tell a familiar story about the immigrant IQ deficit, with Mexicans at the bottom and other Hispanics low as well. The large NIS sample size allows finer-grained ethnic analyses than previous datasets. According to these digit span results, high immigrant Asian IQ is not just the product of Northeast Asians, as the LV national IQ numbers might have implied. The IQ of Indian immigrants is also high, which suggests that the United States enjoys positive selection from that part of the world. The IQ of sub-Saharan Africans is similarly much higher than the LV data would predict, though it is still low by native standards. The impact of selection pressure on immigrant IQ will be discussed in more detail in the next chapter.

NIS Respondents Born in America (Second Generation Immigrants). Table 2.13 shows the results for the American-born children of the NIS immigrants, though with a smaller sample of second generation children only a few ethno-regional groupings are large enough to give meaningful estimates. The overall IQ estimate is much lower in the second generation than in the first, but this is due to children with Latin American parents accounting for a much larger proportion of the sample.

Table 2.13

Second Generation Immigrant - Native Digit Span Group Differences

Immigrant Group	N	Proportion of sample	*d*	Full Scale IQ estimate
All Asia	41	5.9%	0.23	105.0
Mexico	285	41.2%	-0.53	82.1
Central America / Caribbean	228	33.0%	-0.27	89.8
All	**691**	**100.0%**	**-0.33**	**88.0**

Notes: Each group difference is the "immigrant" mean minus the native mean. Positive differences indicate an immigrant advantage. Natives include all races, not just whites. Regional groups with fewer than 40 people are not shown but are included in the total.

The ethnic breakdown is fairly consistent with the first generation. The scores of American-born children with Mexican-born parents are almost identical to Mexican-born

children's scores. Asian scores are also similar to the first generation. Central American and Caribbean scores are higher, but overall there is not much evidence of improvement in the second generation on this culture-fair test.

Some Caveats. A study of Welsh speaking children (Ellis and Hennelly 1980) suggested that the average number of syllables in a language's words for each digit can affect scores on the digit span. Only one digit between 1 and 9 in English has two syllables (the number 7), but several digits in Welsh are disyllabic. The added difficulty for Welsh speakers was theorized to have caused lower scores on the digit span compared to the scores of English speakers. But research on other European and Asian languages (Hoosain 1979; Valencia and Rankin 1985; Stigler et al. 1986; Olazaran et al. 1996) has reproduced the effect of syllable count mostly or exclusively on the *forward* digit span, which was not used in this section. Another study (da Costa Pinto 1991) suggests that the syllable problem is exaggerated, since people use abbreviated forms of the digits in their minds. No cognitive test will have perfect cross-cultural validity, but digits backward appears to come close.

There are two other potential drawbacks to the NIS, which have ambiguous effects on the IQ estimates. First, the NIS surveyed only legal immigrants, who have a somewhat different demographic profile compared to immigrants overall. A second concern is that the NIS interviewed a representative sample of *new* immigrants, meaning recently arrived. Acculturation and education can help raise IQ scores of children, but they probably offer little benefit on the digit span. One of the hypothesized causes of the Flynn effect is increasing familiarity with IQ test questions, yet, as discussed in Appendix B, little to no Flynn effect appears to exist on the digit span. It is a test that is so simple in form, even familiarity may not be of much help. As with the other datasets examined in this chapter, the NIS digit span is not completely ideal, but

the IQ estimates are consistent with the other data presented here, showing a significant immigrant IQ deficit.

CONCLUSION

This chapter has shown that today's immigrants do not merely lack native education and income levels. They also lack the average cognitive ability that natives possess, and there is little evidence that the difference will go away after a few generations. Estimates of immigrant IQ inevitably depend on a variety of data-specific factors, but the results in this chapter are generally consistent across different datasets.

Each of the datasets considered in this chapter has had strengths and weaknesses. The LV national IQ data were culture-fair tests with strong predictive validity, but they could not account for immigrant selection. The NLSY data feature an excellent representation of young immigrants in 1980 who took the ASVAB, but language bias is hard to measure precisely. The PIAT-R can be effectively stripped of internal bias, but as a single test it cannot be subjected to factor analysis as the ASVAB was. Unlike the ASVAB and the PIAT-R, the digit span has a very low knowledge requirement, but it is not as *g*-loaded as the other tests.

Despite individual weaknesses, the datasets complement each other. For example, although language bias cannot be directly measured on the ASVAB, it *can* be isolated on the PIAT-R Math, and the result is similar to the ASVAB. Similarly, we do not know if the *g*-loading is the same for immigrants as it is for natives on the PIAT-R, but we *do* know the *g*-loadings are essentially the same on the ASVAB, and the result is similar to the PIAT-R. None of these datasets alone is dispositive, but their consistency shifts the burden of proof. The contrarian would need to cite a highly *g*-loaded test on which representative samples of white natives and immigrants score the same. No such test exists to my knowledge.

Chapter 3: HISPANIC IQ

The IQ disparity between Hispanics and non-Hispanic whites has major implications for immigrant IQ. Over 56% of immigrants living in the U.S. in 2006 were Hispanic—that is, born in either Mexico (32% of total immigrants), Central American and the Caribbean (17%), or South America (7%). And while a few Hispanics have roots in the southwest going back centuries, nearly 75% of Hispanic Americans in 2006 were first or second generation immigrants.[15] An accurate measure of IQ among Hispanic Americans is thus a useful proxy measure for the IQ of Hispanic immigrants.

Hispanics are not a monolithic group either ethnically or culturally, but the category still has real meaning. Hispanics can be of any race, but they are most often "Mestizo"—a mixture of European and Amerindian background. Mexico, for example, is 60% Mestizo (LV 2006, 241). Hispanics also share ethno-cultural tendencies that are different from the majority Anglo-Protestant culture of the United States (Huntington 2004, 253-255). Most come from Spanish-speaking nations with cultures heavily influenced by Catholicism. And many Hispanics choose to identify themselves as such, as the existence of groups like the Hispanic Chamber of Commerce, the National Council of *La Raza* ("the race" or "the people"), and the Congressional Hispanic Caucus readily demonstrates.

HISPANIC IQ ESTIMATES

We have seen from LV's data that Hispanic countries tend to have lower national IQs compared to East Asian and European countries, and Hispanic immigrants to the U.S. do poorly as well. The same result is apparent for Hispanic Americans regardless of generation. A 2001 meta-analysis by Roth et al. surveyed 39 separate studies that attempted to measure Hispanic IQ. They found an average white-Hispanic IQ difference of 0.72 standard deviations, suggesting a

[15] Source: 2006 CPS March supplement.

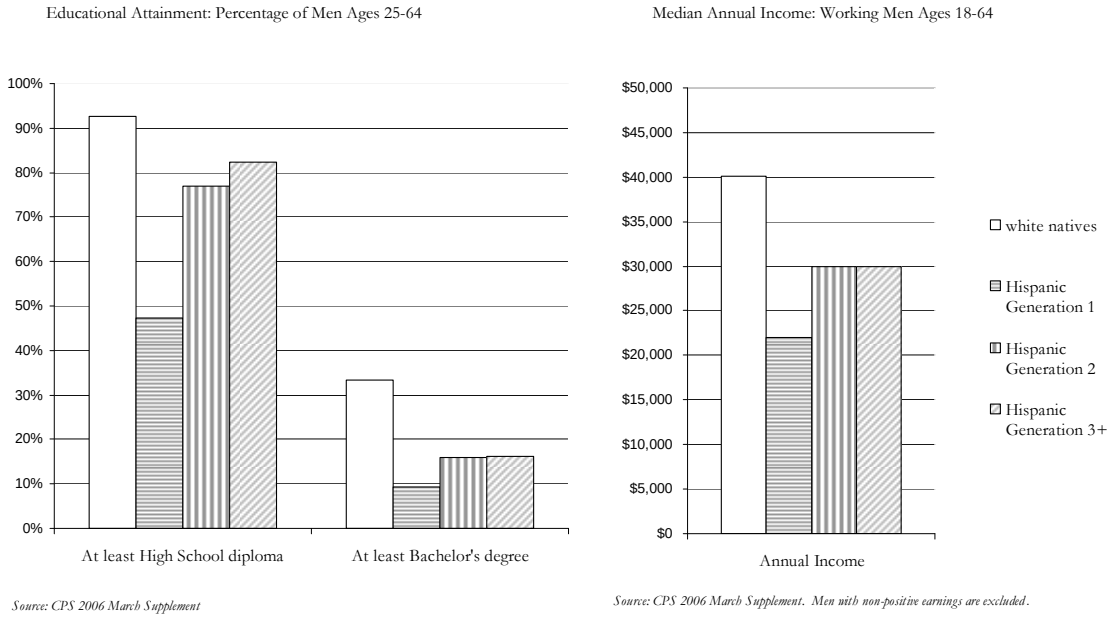
Hispanic-American IQ of 89.2. Since the Hispanics studied were not exclusively immigrants, one could expect fewer problems with language bias—recall from chapter 1 that test bias is essentially nonexistent for native English speakers, regardless of ethnicity.

When Roth et al. separate their IQ results into verbal versus non-verbal tests, the white-Hispanic gap shrinks while still remaining substantial. Here is the magnitude of that difference, in standard deviations, on the verbal versus non-verbal portions of the SAT, ACT, and GRE, respectively: 0.70 versus 0.69, 0.61 versus 0.35, and 0.60 versus 0.51. The differences are still large. Furthermore, as Roth et al. describe, their meta-analysis is consistent with previous attempts to estimate the white-Hispanic difference. Gottfredson (1988) puts the difference at 0.5 standard deviations, while Sackett and Wilk (1994) estimate the difference is between 0.6 and 0.8. Herrnstein and Murray (1994, 275) suggest 0.5 to 1. Finally, the APA's 1995 report stated that "the mean intelligence test scores of Hispanics typically lie between those of blacks and whites."

HISPANIC INTEGRATION BY GENERATION

Another way of examining Hispanic American IQ is to look at socioeconomic outcomes, which are related to intelligence. Figure 3.1 compares Hispanics of several generations to white natives on measures of educational attainment and income. On all three measures, Hispanic natives outperform Hispanic immigrants. However, progress stalls after the second generation, and Hispanics remain well behind whites economically. Even Hispanics whose parents were born in America (the 3+ generation) make only 75% as much annual income as whites. As for education, Hispanics are close to whites in high school graduation rates, but whites are more than twice as likely to hold bachelors' degrees.

Figure 3.1



Source: CPS 2006 March Supplement

Source: CPS 2006 March Supplement. Men with non-positive earnings are excluded.

Other Data on Integration. Some scholars have extended the generational analysis even farther. Samuel Huntington (2004, 230-243) has summarized how specifically Mexican economic and social integration has lagged even into the fourth generation. Huntington cites a 1990 study showing that the percentage of Mexican households with incomes greater than $50,000 rises from 7% in the first generation to 11% in the second. But the statistic in the third and fourth generations stays right at 11%, at a time when the national rate (excluding Mexicans) was 25%. 41% of fourth generation Mexican-Americans also lacked a high school degree in 1989 and 1990, compared to 24% of all other Americans.

A recent book-length study of Mexican-American integration comes to similar conclusions. Telles and Ortiz (2008) revived a 1960s era cross-sectional survey of Mexican Americans by re-interviewing many of the original respondents more than forty years later. By adding information about the children of the respondents in the second survey wave, the authors were able to construct a longitudinal dataset that extends to fourth-generation Mexican Americans. The results show that, relative to whites, the educational attainment of fourth

generation Mexican Americans is no better than the second or third generation. In the words of Telles and Ortiz: "At best, given the statistical margin of error, our data show no improvement in education over the generation-since-immigration and in some cases even suggest a decline" (116). The economic story for Mexican-Americans is no different: "Our findings show a consistent lack of economic progress across generations-since-immigration" (155). For example, Mexican Americans in poverty in 2000 were 17%, 14%, and 21%, respectively, of generations 2, 3, and 4+ when the children of the original respondents were considered (141).

Huntington blames the lack of socioeconomic assimilation on cultural differences, while Telles and Ortiz cite inadequate education. As I discuss in chapter 5, both may be confusing symptoms with the underlying problem. Neither mention low average IQ in the Mexican and other Hispanic populations, which appears to be a key factor. Alternative explanations for the failure of Hispanics to close the socioeconomic gap must point to a phenomenon that differentially affects certain ethnic groups, causes low test scores, and prevents economic assimilation. One cannot simply cite poverty or racial discrimination, since many other groups, especially Asians (Taylor 1992, 109-113), have experienced a large amount of both before becoming successful.

Comparison to Previous Immigration Waves. Low IQ and socioeconomic status has persisted among Hispanics through several generations since 1965, with few signs of improvement. This invites comparison to early twentieth century immigrants from Europe, who were also thought by some to have inferior intelligence levels compared to natives. Today the descendants of those European immigrants are highly similar to the "founding stock" on most measures. The optimistic view of post-1965 immigration is that Hispanic IQ will rise as environments improve, and assimilation will take place much as it did for those Europeans who came a century ago. Unfortunately, this view is misguided for several reasons.

First, European immigrant IQ in the early part of the last century is difficult to ascertain. It was certainly not as low as Brigham and others claimed. The army tests, as chapter 1 explained, were not good measures of intelligence. Quality IQ tests were not used widely until the 1920s, and datasets with valid immigrant IQ scores from that era are hard to come by. There is no doubt that Italians and Poles and others had inferior academic achievement in the first couple of generations, but their abstract reasoning ability compared to the founding stock was not well known.[16] The size of the IQ deficit with natives eventually closed by European ethnic groups is likely much smaller than the one facing Hispanics today.

Second, European ethnics made steady socioeconomic gains, and their assimilation was largely complete after three generations. In comparison, Hispanic assimilation has stalled after the second generation. Among Mexican Americans, for whom we have the most data, even the fourth and fifth-generations do no better than the second.

A third reason that optimism about immigrant IQ is unwarranted is that a sizable number of Mexicans actually did immigrate at the same time as the Southern and Eastern Europeans, and many were in the U.S. even earlier. Unlike the Europeans, they failed to assimilate. Consider Thomas Sowell's (1978) collection of twentieth century IQ data summarized in table 3.1. Jews had high IQ scores dating back to the 1920s. Italians and Poles caught up to the white average by the 1950s, but for Mexicans there was no clear upward trend, just as there is no upward trend today. The quality of Sowell's dataset is questionable, since it was patched together from a variety of tests given to not-necessarily representative subpopulations. However, at a minimum we know that Italians and Poles improved their measured cognitive skills over time, while Mexicans showed little if any increase.

---

[16] Sowell's (1978) claim that groups like the Italians and Poles had poor abstract reasoning ability as well as poor academic performance is not well substantiated.

Table 3.1

Average Ethnic IQ Scores By Decade

Decade	Jewish	Italian	Polish	Mexican
1920s	112	92	91	x
1930s	104	93	95	x
1940s	104	95	99	83
1950s	102	99	104	83
1960s	x	103	107	82
1970s	x	100	109	87

x = too few observations

Source: Sowell (1978, table 1 and table 6)

The same story is true for earnings and education. Borjas (1994b) found that ethnic differentials in earnings and education among immigrant groups in 1910 still existed in 1980 among the third generation. However, excluding Mexico from his analysis made the intergenerational relationship statistically insignificant (Alba, Lutz, and Vesselinov 2001; Borjas 2001). European ethnic groups largely converged in earnings and education over three generations, while Mexican Americans remained well behind.[17] Since Mexicans who have roots in the U.S. going back over a century have not assimilated, and post-1965 Mexican and other Hispanic immigrants have not assimilated over several generations either, it is difficult to be optimistic about their chances in the future.

The fourth reason to be pessimistic is that chances for immigrant advancement are probably greater today than they were for the Europeans a hundred years ago. In the early twentieth century school quality varied enormously, high school graduation was unusual, travel was relatively difficult, and universities and employers were free to ethnically discriminate. Today all but the worst inner-city schools are well-funded, high school graduation is expected,

[17] The remaining intra-European correlation is probably due to high-performing Jewish immigrants, who have made Americans of Russian, Romanian, and Austrian heritage consistently more successful than other European groups.

traveling around the country to look for work is easier, and elaborate affirmative action programs give school- and work-related preferences to Hispanics. Despite these advantages over their European counterparts, many Hispanics have failed to climb the economic ladder.

Today's immigrants do face some comparative disadvantages. The rise of multiculturalism in schools (Krikorian 2008, ch. 1) may discourage many Hispanics from developing an American identity. There are also fewer blue-collar manufacturing jobs in the modern economy, and educational differences between today's natives and today's Mexicans are larger than any native-immigrant difference a century ago (Jencks 2001). Nevertheless, Cubans in Miami have demonstrated that Americanization is not required for economic success, and Asian immigrants have shown that doctors and engineers can emerge from humble roots.

Finally, it is worth asking "how long is too long?" when it comes Hispanic assimilation. No one knows whether Hispanics will ever reach IQ parity with whites, but the prediction that new Hispanic immigrants will have low-IQ children and grandchildren is difficult to argue against. From the perspective of Americans alive today, the low average IQ of Hispanics is effectively permanent.

SUMMARY

The persistently low IQ of Hispanic Americans helps to corroborate the immigrant IQ estimates from the previous chapter, showing that the intelligence of immigrants is a much more valid concern today than it was 100 years ago. The immigrant IQ deficit is a reality that needs to be confronted. The proceeding chapters explore what might be causing the deficit, discuss the importance of IQ generally, and detail some of the deficit's more pressing implications.

Chapter 4: CAUSES OF THE DEFICIT

A natural question to ask about the immigrant IQ deficit is, simply, "What is causing it?" This brief chapter discusses the relevant research, before warning that, in terms of social policy, the *persistence* of the IQ deficit is much more important than its causes. A full treatment of the literature on the causes of group IQ differences is beyond the scope of this study, but readers are encouraged to investigate for themselves the sources in the text and in the note for more information.[18]

SELECTION

One explanation for the IQ deficit is that the United States attracts people from the low side of the skill distributions in poorer countries. Borjas (1987) applied the Roy selection model to the movement of workers between countries. He theorized that the decision to leave one's native country and come to the United States depends on the relative wage distribution in each nation. Countries with compressed wage distributions, where there is a lower relative return to general skill, are likely to send higher-skill immigrants to the United States, where incomes are more spread.

On the other hand, countries with wage distributions even more dispersed than the U.S. will encourage the immigration of lower-skill people who do not wish to be so far below the average wage. Relative to the U.S., the distribution of wages in Western Europe is highly compressed, and the distributions in Latin America and much of the third world are highly

---

[18] Probably the best summary is the exchange between Rushton and Jensen (2005a) and three sets of critics in *Psychology, Public Policy, and the Law* volume 11, number 2. Elsewhere, Herrnstein and Murray (1994) offer a balanced account, and Jensen (1998) is a strong brief for the hereditarian position. The APA statement (1995) has a good outline of environmentalist positions. All of these sources are accessible to non-specialists.

dispersed. This is one reason why, Borjas suggested, third world immigrants in the U.S. have had lower earnings than first world immigrants, even when controlling for education level.[19]

It is easy to accept the central premise of a selection story, which is that people who immigrate are demographically different from the people who stay. The complex economic and social factors that influence the migration decision make that obvious. Nevertheless, there are also good reasons to doubt that selection in and of itself could cause such large IQ disparities, since other factors could overwhelm the effect of wage distributions. In order to be the primary cause of the IQ deficit, Roy-type negative selection must not be outweighed by cognitively challenging requirements like raising money for the trip across the border or the ocean, making one's way in a foreign country, and holding a job without proper documentation (Chiquiar and Hanson 2005).

Even more importantly, the LV data show large differences in IQ across nations, which means no negative selection is necessary to explain low-IQ immigration from low-IQ countries. If anything, the U.S. enjoys positive selection from Southeast Asia, South Asia, Central America, and the Caribbean, according to the results from the previous chapter. There may be a moderately negative selection of Mexicans, but the effects are small. In short, immigrants do not have low IQs because of negative selection. They have low IQs because they come mostly from low-IQ countries. Although selection surely has some effect on immigrant quality, a more parsimonious explanation of group differences recognizes national variation in average IQ.

MATERIAL ENVIRONMENT

If selection cannot fully explain the deficit, the next question is why nations themselves vary in intelligence. The most common explanation is that low-IQ nations suffer from poverty

[19] One could tell a similar story about the generosity of social welfare. Relatively speaking, Europe is more generous than the United States, which is more generous than most poor countries. Therefore, low-skilled Europeans have no reason to come to the U.S., but low-skill people from poor countries do have such an incentive.

and disease that retard the intellectual development of the population. As discussed in chapter 1, the development of cognitive skills is influenced at a young age by environmental factors, as even the strictest hereditarian acknowledges. The national IQs of impoverished nations, particularly in sub-Saharan Africa, could be raised by improved nutrition, healthcare, and early schooling (LV 2006, 244).

Still, there is little evidence that low-IQ countries can fully close the deficit with Europe and East Asia through environmental intervention. As seen in the previous chapter, the immigrant IQ deficit shrinks but does not go away in the Hispanic-American population, even after two generations born in the U.S. Since IQ gains through environmental improvement seem to stall, the real debate is over how much the material environment can affect IQ development after a certain environmental threshold has been met. In the midst of real deprivation, there is no doubt that improving nutrition and cognitive stimulation can raise IQ. But in developed countries where the basic needs of nearly every citizen are met, can environmental interventions still make a difference? The question is particularly acute given the persistence of the Asian-white-Hispanic-black IQ rank order in the United States.

I will not attempt a full treatment of the vast literature on attempts to raise IQ through environmental intervention, but Herrnstein and Murray (1994, 389) sum it up well: "Raising intelligence is not easy…. For the foreseeable future, the problem of low cognitive ability are not going to be solved by outside interventions to make children smarter." Heckman (1995, 1103), in an otherwise critical review of *The Bell Curve*, agreed that "efforts to boost IQ substantially are notoriously unsuccessful."

In order to be considered a success, an intervention must show a statistically significant IQ test gain between a treatment and control group, demonstrate IQ gains across a wide variety of tests, and prove that the effects are long-lasting. Many programs show temporary IQ gains,

but those gains usually shrink or disappear completely as the retest effect loses its impact (Jensen 1998, 334). Initial IQ gains from Head Start, for example, disappear by sixth grade[20] (Herrnstein and Murray 1994, 403).

Still, it is wrong to assume that persistent IQ gains are impossible. A highly intensive early intervention known as the Abecedarian project has produced a 4.4 point IQ difference at age 21 between treatment and control groups (Campbell et al. 2002). The program is not without its critics, who charge that the treatment and controls did not have initially equal ability (Spitz 1992). Abecedarian was also exceedingly expensive, costing $18,000 per child per year for the first five years (Duncan et al. 2007). The Infant Health and Development Program (IHDP) was a similarly intense intervention with a much larger sample size compared to Abecedarian, although it was conducted over a shorter time span. IHDP resulted in no IQ difference between the experimental and control groups by age 5 (Brooks-Gunn et al. 1994).[21] Another intense intervention, the Perry Preschool Program, could not maintain its IQ gains either (Herrnstein and Murray 1994, 404-5). The modest, tentative success of Abecedarian should encourage further research, but a strong dose of realism about raising IQ is needed.

In summary, it is clear that environmental factors significantly affect IQ development when the environment is dire. Immigrants from lower-IQ nations would certainly bring better developed cognitive ability to the U.S. if the material environment in their home countries were

---

[20] This is not to say that Head Start or any other intervention inherently lacks value. Some programs may help children make non-cognitive gains in educational achievement and reduce their chances of committing crimes. These programs should be evaluated, using proper cost-benefit analysis, with all their strengths in mind, even if IQ enhancement is not one of them.

[21] The designers of IHDP report a 4 point increase for the children who were not low birth weight (LBW). LBW children actually saw a decrease in their scores, which averages to no difference in the full sample. Since the designers had originally intended to test the effects of intervention on LBW infants, it is hard to interpret the study as a success. The gains to non-LBW children are as modest as those from the Abecedarian project (Murray 2008, 175-178).

improved. It is much less clear that environmental improvement is effective in developed nations. The evidence on early intervention programs in the United States shows that improving IQ, if it is possible at all, requires a very large resource investment that produces only modest gains. The difficulty occurs because cognitive returns to environmental improvement seem to rapidly diminish after a certain threshold is reached. This is consistent with the findings in the previous chapter, in which immigrant IQ improved over two generations without fully closing the gap with natives. It appears that the material environment is responsible for some but not all of the immigrant IQ deficit.

CULTURE

A subset of environmental explanations for IQ differences is one based on culture rather than on specific material goods. The cultural theory posits that parents or peer groups who are uninterested in education themselves will not provide a cognitively enriching environment for young children. Portes and Zhou (1993), who found that immigrant group culture is related to success, can be considered support for this theory. They found that Sikh immigrant families in California maintained a far more productive ethic compared to the Mexican Americans in their study, and these striking differences in cultural attitudes could help explain IQ differences.

Although not about immigrants, some work on the culture of black Americans is also relevant here. The sociologist John Ogbu (2003) theorized that black underachievement in school and on IQ tests is due to cultural differences with whites. In an ethnographic study of Shaker Heights, Ohio—a racially-mixed, relatively affluent suburb—Ogbu characterizes as "dismal" black parental involvement in their children's education at both home and school (261). Self-report surveys of black attitudes often contradict Ogbu's ethnographic findings (e.g., Ferguson 2001), and it is unclear which type of study is more reliable. In any case, Ogbu's

argument is consistent with an argument put forth by Sowell (2005, 31), that a "redneck" culture transplanted to black ghettos is responsible for low black IQ.

The moderate success of adoption as an "intervention" to raise IQ also can also support cultural arguments. Although it is difficult to identify specific environmental factors that depress IQ in rich countries, adoption can transfer the small, unobservable series of environmental effects that culture entails to disadvantaged children. Indeed, adoption of poor children into middle- or upper-class homes has been a modest but statistically significant success (Jensen 1998, 339-340). One famous study of children adopted into white homes shows small IQ gains, although the magnitudes of the adopted children's IQs still follow a clear hierarchy, with whites highest, blacks lowest, and biracial children in the middle (Weinberg et al. 1992; Levin 1994).

The explanatory power of the culture argument is analyzed in the next chapter in the context of the Hispanic underclass. In short, it is difficult to distinguish the arrow of causation—does culture cause low IQ, or does low IQ influence culture?

GENETICS

Unlike the previous three explanations, a partial genetic theory of group differences in intelligence tends to provoke outrage in the general media,[22] but the theory as applied to black-white differences actually has the support of a plurality of experts (Snyderman and Rothman 1988, 128).[23] The APA report notes, correctly, that no direct genetic evidence for group

---

[22] Recently, Nobel laureate James Watson, the co-discoverer of the double-helix DNA structure, caused uproar when he suggested that Africans have a low average IQ. Watson was excoriated by various scientific academies and public figures, and he retired from his research laboratory amid the firestorm. His treatment is not unique.

[23] I say "plurality" rather than "majority" because some experts did not respond to the question. Here is the full breakdown of the response to Snyderman and Rothman's survey question

"Which of the following best characterizes your *opinion* of the heritability of the black-white difference in IQ?"

differences in IQ exists. However, substantial indirect evidence does exist (Murray 2005). Hereditarians, as supporters of a partial genetic explanation for group differences are often called, start with the observation that controlling for basic environmental indicators does not close the IQ gaps among races, nor do systematic attempts to raise IQ through intervention. They further note that poor environmental quality among some groups could be as much a result, rather than a cause, of low IQ. The incompleteness of environmental factors alone as an explanation for IQ differences suggests genetics could be an underlying cause.

Hereditarians also claim that socioeconomic hierarchies correlate consistently with race all across the world, not just in the United States. Whether the multi-racial region in question is North America, the Caribbean, South America, or Southeast Asia, economic achievement follows familiar racial lines, with East Asians the most successful and sub-Saharan Africans the least (Lynn 2008). When explaining racial differences in achievement, hypotheses that involve slavery, colonialism, and racial oppression have some explanatory power within certain countries and regions. However, none of these local explanations can account for the consistent, global racial differences always observed in societies that have featured reasonable levels of economic freedom. There are no countries, for example, in which ethnic Chinese are less successful than Amerindians, even in places like the Caribbean where the Chinese are a tiny, historically-oppressed minority. When the same racial differences emerge regardless of historical context, genetic differences in ability are implicated.

The difference is entirely due to environmental variation: 15%
The difference is entirely due to genetic variation: 1%
The difference is a product of both genetic and environmental variation: 45%
The data are insufficient to support any reasonable opinion: 24%
(no response): 14%

Among actual respondents, a majority cite genetics as a partial cause.

The hereditarian case is buttressed by a large amount of data showing physiological differences across races—in brain size, rate of maturation, rate of twinning, sex ratio at birth, and many others (Rushton 2000, 9). The racial rank order of these differences is almost always the same, with whites intermediate and Asians and blacks at the extremes. For example, magnetic resonance imaging has shown that Asians have larger brains than whites, who have larger brains than blacks. As discussed in chapter 1, brain size is well correlated with *g*. On other measures, the same physiological rank order emerges. Blacks mature faster than whites, who mature faster than Asians. Blacks also have more twins than Asians, again with whites in the middle. Far from being fringe science, these findings have been replicated by numerous researchers (Gottfredson 2005). They indicate that race is more than "skin deep," meaning genetic differences in intelligence are not at all implausible.

The hereditarians have their critics, of course. For one thing, the white-black IQ gap may have narrowed over the past half century, which is also positive news for the native-immigrant deficit, but the degree and persistence of the narrowing is under intense empirical dispute (Dickens and Flynn 2006; Murray 2006; Rushton and Jensen 2006; Murray 2007b). One could also use a very optimistic read of the Abecedarian Project and adoption studies to attack the hereditarian hypothesis.

But perhaps the most intriguing evidence against heredity is blood group analysis cited by Nisbett (2005). Two different studies from the 1970s (Scarr et al. 1977; Loehlin et al. 1973) used blood groups to estimate the European heritage of black Americans. They found no correlation between European ancestry and IQ. As Rushton and Jensen (2005b) point out, we can now use DNA testing to determine racial heritage far more accurately than blood group analysis. However, assortative mating—the tendency for parents to have similar traits, including comparable IQs—makes any result based on racial admixture difficult to interpret (Jensen 1998,

478-481). The totality of the evidence suggests a genetic component to group differences in IQ, but the extent of its impact is hard to determine.

THE NATURE-NURTURE DEBATE IN PERSPECTIVE

There are several plausible answers to the question of why immigrants and natives differ in IQ. Whole books could be written on just this topic, so the discussion here has been necessarily cursory, and the conclusion that all suggested causes have some truth to them is intentionally vague. Furthermore, much of the research on group differences has compared only blacks and whites. Immigrants, and Hispanic immigrants in particular, have received significantly less attention. More research beyond the black-white dichotomy is needed to draw more definitive conclusions. But regardless of how this research turns out, there are three important points to keep in mind.

Nature versus Nurture is Not an Either-Or Proposition. The previous sections treated environments and genes as distinct causes of IQ differences in order to make the best case for each. However, both causes are intertwined in complicated ways. For example, if someone is genetically predisposed to take a keen interest in mathematics, and that active interest subsequently boosts his mathematical ability, is it biology or the environment that deserves credit? Genes need good environments to exploit, and environments need good genes to enrich. The two interact in ways that make an "either-or" approach to the causes of group differences quite simplistic.

Group Generalizations Are Not Necessary to Immigration Policy. If enough individual data are available, generalizations about group differences, genetic or otherwise, are irrelevant. This applies to all judgments about individuals, but it is particularly important when it comes to immigration policy. It would make little sense to tell an immigration applicant, for example, "Poor people like you tend to have low IQs, so you cannot be admitted," or "Sorry,

people from your ethnicity usually don't score high on IQ tests." As long as each applicant for immigration is considered individually, group generalizations are not necessary.

The Persistence of IQ Differences is Key. Lastly, the equating of environment with malleability and genes with permanence is mistaken.[24] For one thing, genetic disadvantage can often be overcome. A simple example is a nearsighted person who wears glasses. Poor eyesight is usually caused by genes, but the problem can be quickly corrected with a trip to the eye doctor. This is not to suggest that technological compensation for low IQ is as easy, but minor examples already exist—e.g., McDonald's picture-based cash registers for illiterates. Though it may take decades, advances in gene research and brain science are likely to produce future "treatment" for low IQ through direct genetic alterations.

At the same time, environmental disadvantage is not necessarily changeable. We do not know precisely what environmental factors (beyond basic needs) are critical for cognitive development, and few interventions, if any, have been able to permanently raise IQ above a control group. Since nourishing environments for IQ are likely a combination of many small and diverse factors, we may never know how to conduct environmental interventions cost-effectively (Jensen 1998, 344). Because these small environmental factors are also embedded in group cultures, the problem is even more difficult to grasp. How do we go about changing the whole culture of some Americans? Is that even desirable, when the same set of traits can be helpful in some ways and damaging in others?[25]

The degree to which IQ differences are due to environment versus genes does not imply anything about how long the differences will continue. The reason the immigrant IQ deficit is disturbing is not because there may be some genetic component to its causes. The primary

[24] Herrnstein and Murray (1994, 313-315) offer a similar discussion of this point.

[25] Murray (2005) points out that the same fighting spirit that made the Scots-Irish in America such effective pioneers probably also made them prone to violence.

concern for immigration policy is that the differences are *persistent*—for whatever reason. We have seen from the previous chapter that immigrant groups from Europe in the early twentieth century quickly caught up to natives in earnings and academic achievement, while Mexican immigrants persistently lagged behind. Newer waves of Mexicans also continue to underperform natives. Would knowing that intractable cultural differences are preventing Mexican assimilation make the situation any better than discovering intractable genetic differences?

Once again, it is the fact that immigrant IQ differences have persisted that should make policymakers worry, since we have no way to eliminate these differences at this time. Although it is highly unlikely, imagine it were suddenly proven that there are no genetic differences between ethnic groups that could affect IQ, *or* that IQ deficits are entirely genetic in origin. Neither fact would raise anyone's intelligence, and the continuing immigrant IQ deficit would be no less of a problem in either case. The next two chapters discuss the social and economic consequences of this continuing deficit.

Part Three:

CONSEQUENCES AND SOLUTIONS

Chapter 5: THE SOCIOECONOMIC CONSEQUENCES

As the previous chapter argued, the gene-environment debate is much less important than the continued existence of the IQ deficit. This chapter now explores some of the consequences of a continuing deficit. I first discuss the myriad socioeconomic outcomes with which IQ is correlated among individuals, arguing that many of these correlations are causal. I then present in detail two specific areas in which the persistence of the IQ deficit has important implications—the growing Hispanic underclass, and the impact of ethnic diversity on social capital.

Table 5.1
Correlates of IQ

Positive Correlates		**Negative Correlates**
achievement motivation	memory	accident-proneness
altruism	migration (voluntary)	acquiescense
anorexia	military rank	aging effects
artistic ability	moral reasoning	alcoholism
craftwork	motor skills	authoritarianism
creativity	musical ability	conservatism of social views
dietary preference for less sugar and fat	myopia	crime
educational attainment	occupational status	delinquency
eminence, genius	perceptual ability	dogmatism
emotional sensitivity	practical knowledge	impulsivity
extra-curricular attainment	psychotherapy, response to	infant mortality (IQ of parent)
health, fitness, longevity	reading ability	lying
height	social skills	obesity
humor, sense of	SES of parent	psychoticism
income	SES achieved	racial prejudice
interests, breadth and depth of	spelling	reaction time
leadership	supermarket shopping ability	smoking
logical ability	talking speed	truancy

Source: Brand (1987a)

IQ AND INDIVIDUAL SOCIOECONOMIC SUCCESS

IQ is related to a host of socioeconomic outcomes, from educational success, to occupational prestige, to income. In almost all cases, a higher IQ leads to the more desirable outcome. This means that bringing in a large number of immigrants who have lower intelligence

levels will, quite simply, result in more of the bad outcomes in American society and fewer of the good. This section offers a basic overview of IQ's socioeconomic correlates, beginning with table 5.1.

IQ and Socioeconomic Outcomes: Establishing Causality. Although some of the correlates listed in table 5.1 are only indirectly related to IQ (Jensen 1998, 299), others have more direct relationships. One of the most well known demonstrations of the causal relationship between IQ and socioeconomic outcomes is *The Bell Curve* (Herrnstein and Murray 1994). The authors used the NLSY dataset to link AFQT scores with poverty, schooling, occupational success, marriage, illegitimacy, welfare dependency, parenting quality, crime, and civility. By regressing each outcome on AFQT and parental SES, Herrnstein and Murray showed that AFQT score dominates SES as a predictor in almost all cases.[26] For example, the probability of a man in the NLSY who is of average age and SES ever being interviewed in prison goes from 12% to well below 1% as his IQ goes from two SDs below the mean to two SDs above. Conversely, the prison probability for a man of average IQ varies much less with SES—from just 3% to 1.6% as SES goes from -2 SDs below the mean to 2 SDs above (645). The same pattern held for most of the outcomes that Herrnstein and Murray examined.

Criticism of Herrnstein and Murray's method tended to involve the interaction of SES and AFQT, since the two are difficult to separate in practice. Hereditarian critics could charge that parental SES was a reflection of the genes passed from parent to child, so that *The Bell Curve* actually overestimated the role of SES. However, the more common criticism was that Herrnstein and Murray inadequately controlled for parental SES, making it look like a much weaker predictor compared to AFQT than it really is.

---

[26] These analyses were restricted to whites in order to avoid racial complications. They were also broken down by educational attainment where appropriate.

In a response, Murray (1995) asserted that his and Herrnstein's SES index was standard for the literature, and that *Bell Curve* detractors would have to reexamine their own SES variables as a result of their criticism. Two different studies, Fischer et al. (1996, ch. 4) and Korenman and Winship (2000), accepted Murray's challenge to better define the childhood environment, each with mixed results. Using their "better" estimate of SES, the critics needed to address two questions. First, does the power of AFQT drop significantly when SES is "properly" controlled for? And, second, does the power of the environment increase using the new SES measure when AFQT is controlled? The answers are an emphatic "no" to the first question, and a cautious "yes" to the second.

As for the power of AFQT with better controls for the environment, Korenman and Winship employed a clever strategy that ended up confirming Herrnstein and Murray's analysis. Since the NLSY contains hundreds of sibling pairs, the authors used siblings as the SES control. There is hardly a better way to match environments than to compare people who grew up in the same household. When Korenman and Winship did this, they found that Herrnstein and Murray's SES variable had not been inadequate. AFQT scores were still very significant predictors of socioeconomic success within families, just as they were within SES groups broadly defined. "Incredible as it may seem," wrote Korenman and Winship without sarcasm, the result confirmed the independent power of AFQT and the adequacy of Herrnstein and Murray's SES variable to isolate it.

The critics were more successful in arguing that the independent effect of the environment with AFQT controlled is actually larger than Herrnstein and Murray portrayed it. Korenman and Winship redid Herrnstein and Murray's regression analyses by loading the model with every additional environmental variable available to them—number of parents, urban versus non-urban setting, possession of a library card, magazine and newspaper subscriptions,

labor force status of the mother, number of siblings, age of the mother at the time of the respondent's birth, whether the respondent is the oldest child, and immigration status. The result was that environmental factors as a whole now had about the same independent power as AFQT scores. Fischer et al. performed roughly the same procedure and found the same result.

The potential problem with this approach is the one identified by the hereditarian critics—environmental variables partly reflect the intelligence of the parents and their children. The more these "SES" variables are piled on to the right hand side of a regression equation, the more IQ variation they could absorb from the actual IQ variable. Given this possibility, it is actually a testament to the power of IQ that it remained a significant predictor (Nielsen 1997). More controls do not always lead to better regression results—often, they lead researchers to miss the larger picture. For example, Korenman and Winship's results tell us that receiving magazines is a useful predictor of years of education even when AFQT scores are equalized. But what would happen in a controlled experiment that regularly sent copies of Newsweek and Scientific American to randomly selected homes? Would the magazines retain their value as predictors of achievement? Random placement could take away the primary source of the magazine variable's power, since it can no longer absorb part of the child's IQ measure. This is why overspecified models like those in Korenman and Winship and Fischer et al., having so many collinear regressors, are not always useful.

Regarding these methodological problems, Korenman and Winship have an answer to their own critics. If the additional environmental variables are absorbing power from AFQT, they reason, why does AFQT remain such a robust predictor? In fact, it appears that the enhanced SES variables add explanatory power to the regressions without diminishing AFQT. This is a surprising result, since AFQT is undoubtedly correlated with many of the new environmental variables. Nevertheless, it appears that Herrnstein and Murray's critics have

succeeded in establishing a larger role for the environment, without proving a lesser role for AFQT.

Overall, few can deny that the home environment is an important independent factor in child development and adult success. Material goods, family structure, and community culture are surely significant. However, the crucial point here is that IQ is also important, and it cannot be ignored in analyses of social inequality. The old view in qualitative sociology that IQ does not matter at all, whether stated explicitly or by simple omission of the topic, must be discarded.

Unfortunately, *Inequality by Design* by Fischer et al. aggressively endorses the environment-only viewpoint, billing itself as a thorough refutation of *The Bell Curve*. Although the book does succeed in showing, just as Korenman and Winship did, that the impact of the environment was probably underestimated by Herrnstein and Murray, its overarching theme is that nearly all outcomes in life are socially-determined, with no significant role for genes. Fischer et al. devote a whole chapter to the environmental determinants of intelligence itself, ignoring the substantial differences in AFQT scores between siblings, to assert that test performance simply reflects environmental quality.

The main thesis of the book, that social structure determines the level of a society's inequality, is a near tautology that the authors treat as a profound insight. One can think of any number of ways to structure society so that outcomes are equal—a complete redistribution of wealth comes to mind—but natural differences in ability can only be concealed by redistribution policies, not eliminated. The evidence for the biological heritability of IQ is overwhelming (see chapter 1), and any parent with more than one child knows that the same environment can produce very different people. Social scientists are right to examine the home environment, but they are not seeing the whole picture if they follow Fischer et al. by minimizing or ignoring IQ.

I give the last words on this debate to two extensive reviews of the recent literature. According to Bowles, Gintis, and Osborne (2001), the correlation between IQ and earnings is only about 0.15 when education is controlled. But *no* variable is a good predictor of earnings, which appears to depend on a variety of idiosyncratic differences in personality. Nevertheless, the authors state: "The independent importance of schooling and cognitive functioning [IQ] is uncontroversial" (1147).

A careful meta-analysis by Strenze (2007) demonstrates that the importance of IQ is much more evident in the literature when it is linked to education—average correlation of 0.56 in 59 different studies—and occupational prestige, with an average correlation of 0.43 in 45 studies. Strenze sums up: "Intelligence is an independent causal force among the determinants of success; in other words, the fact that intelligent people are successful is not completely explainable by the fact that intelligent people have wealthy parents and are doing better at school" (416). In short, IQ matters.

IQ as Probability of a Skill Set. But what does "IQ matters" actually mean? When comparing individuals, the effect of IQ differences is often small. A large number of personality attributes, many of which are unrelated to IQ, affect a person's ability to succeed in life. For that reason, an individual's IQ score is merely a probability of future success, not a prediction from a crystal ball. For example, a person's IQ affects his likelihood of completing college, but some college graduates are not very smart. Betting that an individual person with an IQ of 100 will complete more years of schooling than a person with an IQ of 95 is a risky gamble. The less intelligent person may be a very hard worker, while the smarter person could be lazy and unmotivated. However, if presented with two groups of 100 random Americans, one group with average IQ 95, the other group at 100, it is a virtual certainty that the smarter group will

have higher educational attainment. In this way, IQ scores can be thought of as individual probabilities that aggregate into certainties in large groups.

The first row of the following table shows the percentage of NLSY-79 respondents by IQ group who earned a four-year degree. College completion by people with below-average IQs is rare, and earning a degree is commonplace only among those with IQs above 115.

Table 5.2
Percentage of NLSY Respondents Earning a BA or BS by IQ Group

	<76	76-80	81-85	86-90	91-95	96-100	101-105	106-110	111-115	116-120	120-125	>125
among all NLSY respondents	0.0%	0.5%	1.4%	3.5%	5.0%	8.8%	22.8%	26.0%	43.1%	61.2%	75.9%	77.8%
only among those who enrolled in college	0.0%	2.0%	5.0%	11.7%	15.4%	19.4%	37.5%	40.5%	54.7%	69.0%	79.5%	80.8%

Many people do not attempt to complete any post-secondary education, but IQ helps determine college completion even when the sample is limited to those who try. The second row considers only NLSY respondents who enrolled in a college at some point after high school. The percentages with college degrees are higher in each IQ group, but the association with IQ is still strong. Among people with IQs in the 96-100 range who go to college, fewer than one in five will go on to earn a four-year degree.[27] Not everyone who goes to college intends to earn a BA or BS, but this indicates that college completion is not simply a matter of access—it is also a matter of IQ.

Going beyond educational achievement, Gottfredson (1997) has developed average skill profiles of people in various cognitive classes by linking results from the National Adult Literacy Survey (NALS) to IQ.

---

[27] These data are also restricted to whites. In order to qualify as college graduates, NLSY respondents needed to claim a BA or BS and have at least 15 years of schooling by 1990. Anyone in college between 1979 and 1990 counted as someone who had enrolled, although no one currently in school in 1990 was considered in the analysis. College enrollment data was missing in 1987.

Table 5.3
National Assessment of Literacy Scales and the IQ Distribution

Skill Level	Example Skills	IQ Range	Proportion in Each Skill Level: Whites	Proportion in Each Skill Level: Immigrants	White/Immigrant Ratio
5	• interpret a brief phrase fron a lengthy news article • summarize two ways lawyers may challenge prospective jurors • using information in a news article, calculate difference in times for completing a race • using a table comparing credit cards, identify the two categories used and write two differences between them		4.1%	1.4%	3.00
		126.1			
4	• contrast views expressed in two editorials on technologies available to make fuel-efficient cars • use table of information to determine pattern in oil exports across years • using information stated in a news article, calculate amount of money that should go to raising a child • explain difference between two types of employee benefits		21.6%	11.8%	1.83
		109.8			
3	• calculate miles per gallon using information given in a mileage record chart • use a bus schedule to determine appropriate bus for given set of conditions • using a calculator, determine the discount from an oil bill if paid within ten days • read a news article and identify a sentence that provides interpretation of a situation		36.1%	30.2%	1.19
		95.5			
2	• identify and enter background information on application for social security card • locate eligibility from table of contents • determine difference in price between tickets for two shows • calculate postage and fees for certified mail		25.0%	30.8%	0.81
		83.3			
1	• locate one piece of information in sports article • total a bank deposit entry • locate time of meeting on form • locate expiration date on driver's license		13.2%	25.8%	0.51

Notes: Assumes immigrant IQ of 93. The white IQ distribution is converted to N(100,15) from N(101.4, 14.7) in Gottfredson (1997, table 8).

The NALS identifies how many Americans fit into five different levels of competence in practical, everyday skills. Gottfredson describes how these skill levels closely match the American IQ distribution, with each successively more complex task providing a greater cognitive challenge. Table 5.3 describes some of the skills required for competency at each level, the range of IQs that correspond to those skills, and the percentages of people who fall within each range. I have contrasted the distribution of white American skill with hypothetical immigrant skill, assuming an immigrant mean IQ of 93. The difference in IQ distributions obviously results in substantial differences in practical skill, with the differences most pronounced at the tails of the distribution.

Note that these estimates are not based on empirical tests of immigrant literacy skills, which would surely be affected by language bias. These data represent the distribution of immigrants' skills if they were to acquire native proficiency in English, meaning the data overestimate their current ability level. In fact, actual immigrants in the NALS were 3.7 times as likely to appear in the lowest skill level as white natives, compared to only about twice as likely in the table above (Kirsch et al. 1993, table 1.1). Also, each skill listed in the chart is based on a probability. There are surely people in the lowest range of IQ who can calculate postage on certified mail, but that task is not *typically* a skill possessed by the average person in that cognitive class.

This brief review of the practical validity of IQ was meant to add context to the immigrant IQ deficit documented in chapter 2. IQ is significantly correlated with a large number of life outcomes, and this correlation survives controls for environmental advantages. A person's IQ helps determine not only his major life accomplishments, such as finishing school and choosing a career, but also the basic skills that allow him to function well in society on a day-to-day basis (Gordon 1997). People with high IQs have a high probability of graduating

from college, working a well-paying job, keeping their families intact, and avoiding crime. On the opposite end of the IQ spectrum, school achievement and occupational success are hard to find, and social pathologies like crime and illegitimacy are far more common.

Therefore, the overrepresentation of immigrants on the left side of the bell curve has substantial implications for the American economy and for society in general—so many, in fact, that listing them all may not even be possible. However, there are two specific implications of low-IQ immigration that are worth explicating in some length—first because they are prominent social problems, and second because IQ is rarely considered to explain them. They are the growing Hispanic underclass, and the negative effect of ethnic diversity on social capital.

THE HISPANIC UNDERCLASS

A broad but useful generalization is that there are two types of poor people—those that conform to middle class standards of behavior, and those who flout such standards. The former group is the working poor, a class of people who stay employed even at low-paying jobs, have children only when economically prepared for them, and contribute to civil society. The latter group is the underclass, a socially-isolated group of people for whom crime, welfare, labor force dropout, and illegitimacy are normal aspects of life (Wilson 1987, 7-8; Jencks 1992, 16).

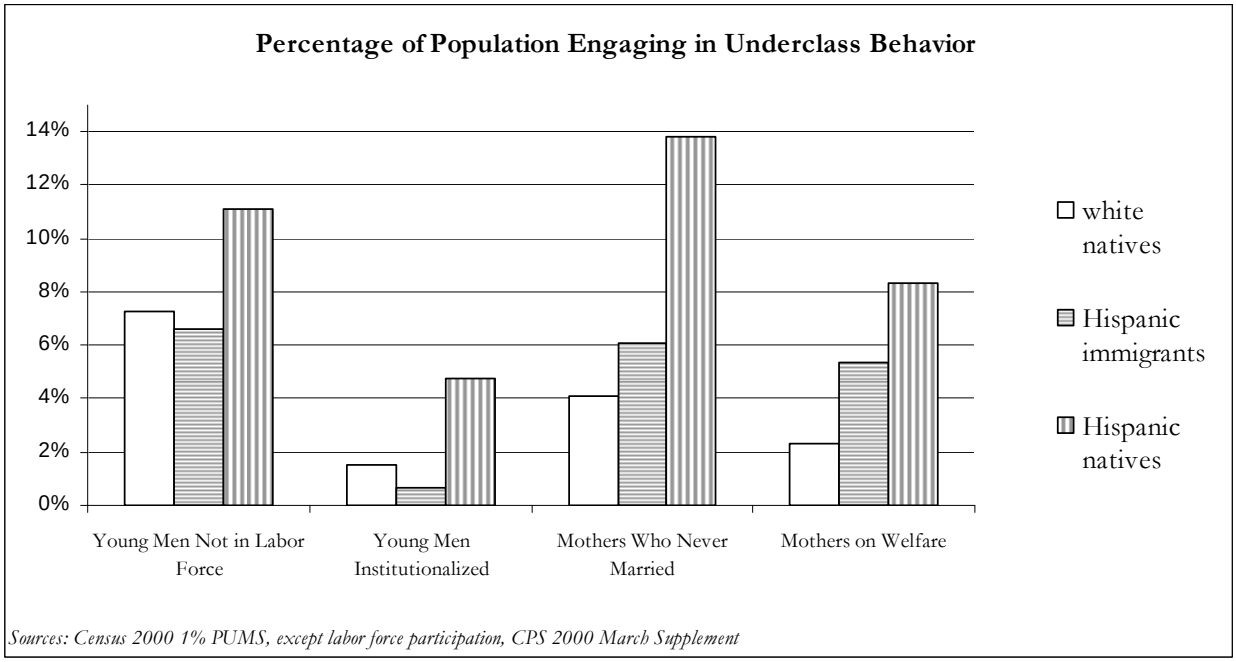
The differences between each group are often blurred at the margins (Jencks 1992, 202-203), but underclass behavior is a distinct social problem that grew to prominence after the 1960s. While the working poor must struggle to make ends meet, they are at least in a position to enjoy the basic satisfactions of life. The underclass, on the other hand, lacks access to strong families, enriching community associations, and safe neighborhoods, all of which contribute to a satisfying existence (Murray 1999, 36). Underclass behavior is also a particularly difficult problem due to its intractability. Expanded opportunities for employment and education have

helped the working poor, but they have done much less for the underclass due to cultural obstacles (Wilson 1996, 75-77).

This section discusses the growth of the Hispanic underclass in the United States. I first document how many second generation Hispanics slip from the working poor status of their immigrant parents into the barrio underclass. I then offer the hypothesis that IQ, a long ignored topic in the underclass literature, can account for this intergenerational phenomenon.

Underclass Behavior in the Hispanic Second Generation. Many Hispanics have taken full advantage of the opportunities the U.S. provides by getting educations and entering the middle class. At the same time, however, an underclass has developed among some Hispanic natives. Figure 5.1 compares white natives, Hispanic immigrants, and Hispanic natives on four of the most common indicators of the underclass. In each case, Hispanic immigrants are comparable to white natives, but Hispanic natives do much worse than either group.

Figure 5.1



Sources: Census 2000 1% PUMS, except labor force participation, CPS 2000 March Supplement

The first indicator is labor force participation. Anyone at work or actively seeking work is counted as a member of the labor force. The percentages shown in the table are men ages 16 to

24 who are out of the labor force—that is, not in school, not at work, and not looking for work. Most of these young men will get jobs later in life, but their youthful idleness will have prevented them from gaining the experience and training needed for higher-paying jobs (Murray 1999, 10). As the figure indicates, Hispanic immigrants come to work, but their children's labor force participation slips considerably.

The second indicator is the percentage of young men who are institutionalized, which is a proxy for imprisonment.[28] Perhaps surprisingly, Hispanic immigrants are less than half as likely to be institutionalized as white natives.[29] Institutionalization among Hispanic natives, however, is very high relative to the other two groups. The same story applies to mothers who never married and mothers on welfare. Each time, Hispanic natives do significantly worse than the comparison groups. The outlook is not all downhill for Hispanic natives, who do earn more and get better educations on average than their parents (see chapter 2). But superior performance on basic economic indicators is to be expected from the later generations, who go to American schools, learn English, and become better acquainted with the culture. Despite built-in advantages, too many Hispanic natives are not adhering to standards of behavior that separate middle and working class neighborhoods from the barrio.

Ethnographic studies confirm the development of countercultural attitudes characteristic of the underclass in the Hispanic second generation. Portes and Zhou (1993) observe that Mexicans and South Asians from immigrant families have distinctly different behavior regarding

---

[28] The Census classifies as institutionalized not only people in prison but also those who are in facilities for physical and mental disabilities. The categories cannot be separated in the Census as of 1990, but prisoners are easily the largest institutionalized group. A study using survey data in Chicago (Sampson et al. 2005, table 2) gives essentially the same results as the Census data, with Hispanic crime rates going up substantially in the second and third generations.

[29] The difference is not due to immigrants having a shorter stay in the U.S. (Rumbaut and Ewing 2007; figure 9) or being deported rather than imprisoned (Butcher and Piehl 2008).

assimilation. Mexicans often assimilate into the "barrio culture" of poor Mexican-Americans, featuring underclass attitudes counterproductive to advancement, whereas the South Asians in their study remained culturally aloof from the underclass and prospered.

Portes and Zhou find that negative attitudes toward work and school among Mexican immigrant families actually increase with assimilation into Mexican-American culture. The authors describe second generation "Chicanos" and "Cholos" as "locked in opposition with white society" (88). They are seen by their teachers as unmotivated and irresponsible, and they view "acting white" as disloyalty to their own group. In contrast, Portes and Zhou describe the success of Punjabi Sikhs in California, who had no Indian-American oppositional culture to absorb them. Unlike the Mexicans, the Sikhs developed a strong emphasis on English, math, and science, and they outperformed whites academically.

IQ and the Underclass. There can be little dispute that post-1965 immigration has brought a larger and increasingly visible Hispanic underclass to the United States, yet the underlying reasons for its existence cannot be understood without considering IQ. The standard theories offered to explain the underclass usually fall into two categories—the loss of good-paying manufacturing jobs in cities, and the expansion of the social welfare system. The first theory was developed fully by Wilson (1987), who argued that structural changes in the economy during the 1970s eliminated many manufacturing jobs, leaving some black inner city residents unemployed. The lack of good jobs led to a dearth of "marriageable men" for black women, which caused illegitimacy to rise. Eventually, chronic unemployment and illegitimacy, combined with the outmigration of middle-class blacks from the ghetto, helped create an underclass culture hostile to low-wage work and traditional marriage.

Regardless of its value in explaining the black underclass, this theory is not relevant to most Hispanics, who have been in a different economic situation compared to blacks. Hispanic

immigrants intentionally move to the parts of the United States where jobs are most available. The children of recent immigrants have not subsequently experienced the rapid deindustrialization that young blacks encountered in the 1970s, yet many still join the underclass culture that their less privileged parents avoided.

The welfare theory was prominently advanced by Murray (1984). He argued that government began to have a more permissive attitude toward the poor—primarily through less restrictive welfare benefits, but also via changes in bureaucratic regulations and elite attitudes—that made destructive long-term behavior appear attractive to the poor in the short-term. The government made it economically possible to have children out of wedlock and avoid undesirable work, so many took advantage of the situation, eventually weakening the social stigma against such behavior.

Using government transfers to turn illegitimacy and joblessness into attractive short-term decisions could certainly increase underclass behavior. However, a key question is left unanswered by the welfare theory—even if something looks like a good choice in the short-term, shouldn't most people understand that it is still a bad choice in the long-term, and then avoid it? One of the hallmarks of a high IQ is the ability to understand the long-term consequences of behavior (Wilson and Herrnstein 1985, 167). This includes setting and fulfilling future goals and making important decisions with the long-term in mind. When given the choice between a paycheck from a low-paying job and a welfare check, most intelligent people would realize that the welfare check offers them no potential for advancement. Low IQ people do not internalize that fact nearly as well. Indeed, Hymowitz (2006, 115) reports interviewing unwed teenage mothers who have dreamy beliefs about becoming doctors or lawyers someday, apparently unaware that single motherhood could be an impediment. This is not the fleeting

idealism of youth, but rather a lack of understanding about the investment of time and energy needed to live a normal adult life.

In order to explain the creation of the underclass, the welfare theory requires present-oriented recipients, a common trait in low-IQ populations. In fact, table 5.4 lists the rate of various underclass behaviors within cognitive classes. While rare for the cognitive elite, social pathologies are far more common at the lower tail of the IQ distribution.

Table 5.4
Percentage of White NLSY Respondents Exhibiting Underclass Behavior in Each Cognitive Class

	IQ Class					lowest:highest ratio
Underclass Behavior	<75	75-90	90-110	110-125	>125	
men not in labor force one month or more	22	19	15	14	10	**2.2**
women who gave birth to illegitimate baby	32	17	8	4	2	**16**
mothers on welfare after first birth	55	21	12	4	1	**55**
men ever interviewed in prison	12	7	3	1	1	**12**

Source: Herrnstein and Murray (1994, pgs 158, 180, 194, 248)

In addition to a low IQ population, the welfare theory also requires an oppositional culture. If welfare recipiency, illegitimacy, and joblessness met with strong social condemnation, whether or not people could make rational long-term calculations would be irrelevant. The social disapproval of such behaviors would prevent them from becoming widespread. Here the welfare theory is incomplete, because it treats cultural change only as a *result* of widespread bad decision-making rather than as an enabling factor. In fact, countercultural attitudes can be explained by IQ differences. The argument, in brief, is that Hispanics become less willing to play by the rules of the middle class when their low average IQ prevents them from joining it.

The detailed version of the story goes as follows. Poor and unskilled immigrants travel to the United States, seeking to earn a higher wage in the U.S. and give their children more opportunities than they had themselves. This first generation of immigrants does not belong to the underclass. The first generation works hard—why even bother to come if not to work?—

stays away from crime and drugs, and tries to advance. This is generally true of all immigrants regardless of origin, but the story begins to diverge with the second generation.

Hispanic immigrants and their children have a low average IQ, which prevents the second generation from reaching equality with the native majority. Parental expectations for their children are not met, because they cannot be, given the level of intelligence present in the community. The average Hispanic child inevitably lags behind the average white in high school achievement, in college admissions, and in job selection. The failure to achieve parity with natives then triggers a natural human response, which is to downplay the importance of things that one is not good at.

This might be called the "nerd-jock phenomenon." While some people are blessed with both academic and athletic talent, many people have just one or the other. In most cases, the "nerds" will consider their bookish pursuits to be far more important than, say, throwing a ball through a hoop, while the "jocks" will feel exactly the opposite way. This is a natural psychological mechanism that helps give people a sense of self-worth. In the case of some second generation Hispanics, it causes them to reject the basic cultural norms of the majority. Schoolwork becomes unimportant, college-prep is snobbery, and holding down a low-paying job means working for chump change.

An entirely different situation exists with most Asian immigrants, who generally possess the intellectual ability to not only compete but to out-compete natives in academic pursuits. The children of Asian immigrants—even when their parents are uneducated, as in the Sikh example—quickly realize that they can beat whites at their own game, so there is no alienation, no resentment of success, and no looking down upon hard work. It is the underlying ability of each immigrant group that affects not only their actual socioeconomic success, but also their cultural attitudes toward achieving success. This is how low IQ accounts for the negative

attitudes toward work that the welfare theory cannot fully explain. The frequency of failure causes people to turn away from conventional means of trying.

Reverse Causation. As mentioned in the previous chapter, some scholars have theorized that it is actually the oppositional culture that causes low IQ rather than the other way around. If Hispanic children are dissuaded from traditional work and school by parents and peers, perhaps their IQ scores are depressed as a result. As with other hypothesized causes of the IQ deficit, culture could certainly have some explanatory power. However, in this case it suffers from a fundamental flaw—IQ was low before oppositional culture took hold.

As stated above, it is natural for individuals to downplay the importance of skills they do not possess or tasks that they do not perform well. If many Mexican-Americans cannot succeed in school due to low IQ, they may develop opposition to schoolwork as a psychological defense mechanism. Portes and Zhou acknowledge the point about self-worth:

> …U.S.-born children of earlier Mexican immigrants readily join a reactive subculture as a means of protecting their sense of self worth. Participation in this subculture then leads to serious barriers to their chances of upward mobility because school achievement is defined as antithetical to ethnic solidarity. (89)

The authors blame the origin of this defensive culture not on low ability but on white racism and the immigrant parents' poverty. But that is an insufficient explanation in light of the Sikh example discussed above. The Sikhs were equally impoverished and subject to discrimination, yet they embraced education and hard work. Portes and Zhou claim the difference is that no Indian-American oppositional culture existed that might assimilate them. This is true, but how did the original negative subculture develop among Mexicans? Why did the first Mexican Americans and their children not succeed, when there was no subculture trying to assimilate them? Can everything be blamed on being "involuntary minorities" after the Mexican-American war, as Ogbu and others have suggested? This is a chicken-and-egg problem. Culture can affect intelligence, but intelligence surely affects culture as well.

One can see the problems with the culture-only argument by imagining what would happen if Hispanics suddenly had the same underlying distribution of IQ as whites. Hispanics would rapidly become competitive with whites in school. Equal proportions of whites and Hispanics would have the ability to earn academic honors and succeed in gifted classes. Oppositional culture would still push some down, but all that is needed is a critical mass of smart Hispanics who would work hard in school in order to earn top honors, go to prestigious universities, and get well-paying jobs. That kind of economic success would be difficult to resist. Once the goal was within reach, there would be little reason for other Hispanics to regard it as betrayal of their group. Similarly, imagine if Asians suddenly suffered a dramatic decrease in their intellectual ability. As Asian school achievement declined, would alienation not set in? Would near-obsessive devotion to study not be curtailed in order to protect self-esteem? The reality of IQ's effect on culture, and its subsequent role in underclass behavior, must be considered.

IMMIGRATION AND SOCIAL CAPITAL

Though he did not invent the concept, Robert Putnam helped make "social capital" one of the central concerns of economics and sociology with the publication of his essay "Bowling Alone" in 1995. Putnam defines social capital in simple terms: "social networks and the associated norms of reciprocity and trustworthiness." Like human capital (physical and mental ability) and physical capital (land, machines, etc.), social capital is an important factor in economic production functions. Building complex networks of friends and associates, trusting others to keep their word, and maintaining social norms and expectations all grease the wheels of business by enabling cooperation. But the importance of social capital goes beyond economics, straight to the heart of happiness itself. People living in areas with high social capital tend to have more friends, care more about their community, and participate more in civic

causes. All of these things are associated with happiness generally. Putnam sums up: "…Where levels of social capital are higher, children grow up healthier, safer and better educated, people live longer, happier lives, and democracy and the economy work better" (2007, 137-138).

Ethnic Diversity. Recently, Putnam encountered a finding that was disturbing to him—ethnic diversity is negatively associated with social capital, and no amount of statistical wrangling can make the relationship go away (2007). The places where people are most likely to say that they trust their neighbors—a key component of social capital—are homogenously white areas such as North Dakota, Montana, New Hampshire, and Maine. The least neighborhood trust exists in places like San Francisco and Los Angeles, where whites, blacks, Hispanics, and Asians live in close proximity to each other. Even when individual people rather than communities were the units of Putnam's analysis, more diversity was associated with less social trust. The problem this presents for immigration policy is obvious, since most immigrants to the U.S. are non-white. In this section, I develop an argument that IQ selection could partially mitigate the negative effect of diversity, making immigration more palatable without resorting to a race-based policy.

IQ and Social Capital. Do higher IQ communities have more social capital? Intuitively, it is not a stretch to believe that smarter people are better at organizing and maintaining networks, understanding the long-term benefits of cooperation, and internalizing their place within a community. Empirically, no one has directly examined its impact on social capital, but IQ has been separately linked to major components of social capital, such as altruism, trust, and cooperation. Herrnstein and Murray (1994, 253-266) devoted a chapter to AFQT scores and what they called the Middle Class Values (MCV) index. The MCV index is a binary variable coded as 1 for respondents in the NLSY if they meet all of the authors' criteria—graduating from high school, keeping out of jail, staying married to a first spouse, maintaining

employment, waiting until marriage to have children, etc. The MCV index is a quick way to measure "…ways of behaving that produce social cohesion and order." 74% of people in the highest cognitive class met the MCV criteria, while just 16 percent did in the lowest class. The relationship easily survived controls for parental SES.

Interesting as it may be, the MCV index is an indirect and somewhat simplistic measure of real social capital. It is probably true that people meeting the MCV criteria are largely the same people who go to PTA meetings and return lost wallets as Herrnstein and Murray assert. But there is more direct evidence linking IQ to social capital, starting with the work on impulsivity by de Wit et al. (2007). The more impulsive a person is, the more likely he is to discount future rewards in favor of immediate gratification. The authors of this study measured impulsivity by making a variety of hypothetical monetary offers to a group of 600 adults who had also taken an abbreviated IQ test. Each offer consisted of a lesser cash reward in the present versus a larger cash reward at some future date. Answers to these questions allowed the researchers to determine the degree to which each participant discounted the future.

The major finding was that higher IQ people are substantially less impulsive, even controlling for age, gender, race, education, and income. The large and diverse sample used by de Wit et al. makes this one of the best studies of its kind. The findings were soon bolstered by a meta-analysis (Shamosh and Gray 2008) that found a moderate mean correlation between IQ and "delay-discounting"—that is, the tendency to ignore the future—of -0.23.

Intuitively, smarter people should be able to internalize future rewards more easily. They are probably more future-oriented because they can better manipulate their surroundings, whereas incompetent people exert less control on their future, making it murky and unknown. Whatever the cause, the impulsivity of low-IQ people has serious implications for social capital. People in less intelligent populations will be less willing to set up networks for potential long-

term payoffs, make personal investments in the community, and follow basic norms of behavior with the expectation of future reciprocity.

An even more direct link between IQ and social capital was recently shown by Jones (2008) in a clever study of prisoner's dilemma games played on college campuses. The prisoner's dilemma is a well-known and much studied game theoretic situation. There are many variations, but the basic situation is as follows. You and an accomplice are accused of a crime that carries a maximum penalty of 10 years in prison. The police admit that if no one confesses, they will only have sufficient evidence to charge you each with a lesser crime, and you will both get 2 years in prison. If you both confess, the authorities will be lenient, and you will each have to serve 5 years in prison. So each of you is offered a separate deal. If you confess and your partner does not, you get just 1 year in prison, while your partner gets the full 10. If your partner confesses and you do not, then the payoffs are reversed (Mas-Collell et al. 1995, 236).

Obviously, neither person confessing is the best overall outcome for the prisoners. However, selfish prisoners will end up both confessing, because confessing always provides the better individual payoff. In order to achieve the socially optimal result, trust in your partner is required. Will he recognize the potential for cooperation by not confessing, and will you reciprocate by refusing to confess as well? People who trust each other more will usually achieve the best outcome. This is just one formalized example of how social trust can improve the way a society functions.

Prisoner's dilemma games have been played as experiments on college campuses to test all sorts of hypotheses over the years. The key insight made by Jones is that average SAT scores for each college are known. Although the Educational Testing Service does not describe it as an IQ test, the SAT is actually a good measure of *g* (Frey and Detterman 2004). Jones correlated the proportion of students who cooperated in the prisoner's dilemma at each college with the

average SAT score of the college. He found a substantial and robust correlation. To illustrate, colleges with SAT scores around the national average of 1000 cooperated about 30% of the time when faced with the prisoner's dilemma. Top-flight colleges with average SAT scores around 1450 cooperated about 51% of the time. Had IQ scores of individuals been available rather than just group averages, the relationship would likely have been even stronger. It is clear that more intelligent people are better at cooperating.

So far, IQ has been linked to possessing middle class values, having a future time orientation, and cooperating in competitive games—all components of social capital. Altruism is one last social value with which IQ may be associated, although the evidence is less definitive. An altruist endures a personal cost in order to help others, even when he gets no extrinsic reward for doing so (Rushton 1981).[30] Unlike the prisoner's dilemma game discussed above, in which each party stood to gain from cooperation, altruism is simply generosity. Intuitively, it is much less clear why intelligent people would be more purely generous—unlike mutual cooperation, there is no individual reward to enjoy.

Nevertheless, recent evidence suggests a positive relationship among adults.[31] Millet and Dewitte (2006) gave a group of undergraduates the Social Value Orientation measure, which presents a series of situations in which the respondent gets one amount of money and a stranger gets another amount. Respondents must rank their order of preference for each situation as the amounts of money change. Altruistic people were defined as those who preferred less money for themselves in order for a stranger to receive a higher amount. The most altruistic people

[30] This is Rushton's definition; Sorrentino favors a stricter standard. Rushton's definition relies on the behavioral aspect of altruism, and it ignores the possibility of intrinsic rewards enjoyed by the altruist. Technically, altruism may be a logical impossibility when the stricter definition is required—if a person *wants* to be an altruist, then his generosity is a product of self-interest, and altruism becomes a self-defeating concept.

[31] Rushton and Wiener (1975) found no relationship between IQ and altruism in young children.

scored nearly 8 points higher on an IQ test than the least altruistic people. Whether altruism serves as a costly signal of intelligence as the authors suggest, or intelligence gives people a broader social perspective, or some intervening variable is responsible for the relationship, is unknown.

In summary, higher IQ people exhibit greater valuation of and planning for the future, cooperate more easily when mutual benefit could occur, possess "middle class values" at higher rates, and may even be more given to altruism. These results are supported by both standard intuition and solid empirical evidence.

The Effect of Low-IQ Immigration on Social Capital. Since several components of social capital are intimately related to IQ, the level of trust and cooperation in a population will be partially determined by its intellectual strength. Even leaving aside the ethnic diversity issue for now, Americans can expect low-IQ immigrant neighborhoods to feature significantly less social capital, which will make them less pleasant places to live, work, and go to school. Indeed, there is now significant evidence that Hispanics, both at the individual and community level, are less trusting compared to whites. Putnam (2007) found that Hispanic ethnicity was associated with substantially lower levels of social trust, even when the relationship was tested in regression equations with a detailed set of control variables.

There is no consensus explanation from sociologists for this phenomenon, yet low average IQ has not yet been identified as a possible cause. Standard stories about poverty and crime will not suffice, since they are controlled for in the Putnam study. Wierzbicki (2004, 16) has suggested that Hispanics have too little time for socializing because they are working nearly constantly. If true, however, this explanation would not be adequate to explain low trust among blacks and native-born Hispanics, who have much lower labor force participation rates than Hispanic immigrants. Another idea listed by Wierzbicki is that disproportionate representation

in domestic jobs causes social isolation among Hispanic immigrants, but this theory lacks much empirical support. Lastly, Mahler's (1995) provocative thesis is that interclass jealousies and ruthless intragroup competition among immigrants cause trust to erode.

There are surely many reasons why some groups are less trusting than others, not all of which depend on IQ. For example, Rice and Feldman (1997) demonstrated substantially different levels of civic engagement across white American ethnic groups, even though each group has essentially the same average IQ. However, the individual-level relationship between social capital and IQ is too strong to ignore. In fact, it seems that high IQ is an insufficient but *necessary* condition for fostering highly cooperative and trusting communities in the modern world.

Mitigating Diversity with IQ Selection. As discussed above, the negative impact of diversity per se on social capital is difficult to dispute. Literally thousands of different model specifications used by Putnam failed to uncover a confounding variable that could make the relationship spurious. Nevertheless, it is also true that the *type* of diversity could help determine the extent of its undesirable effects. My hypothesis is that higher-IQ non-whites will have substantially less negative impact on social capital. People with higher IQs are more likely to build trusting communities themselves, and they could also find it easier to integrate into established high-trust neighborhoods. If this is true, then a major benefit of immigrant IQ selection is that it could make non-white immigration more tolerable in terms of maintaining social capital.

One testable prediction of the hypothesis is that the presence of Asians (who have a high average IQ) in a given neighborhood should cause less deterioration of trust among whites than the presence of blacks or Hispanics (who have comparatively lower average IQs). Table 5.5 displays the results of a regression of social trust among whites on the percentages of blacks,

Hispanics, and Asians living in their census tract. The dependent variable is the response to the survey question, "How much can you trust people in your neighborhood?" from Putnam's dataset. Respondents indicate their level of trust on a four-point scale. Column I shows that the presence of Asians decreases trust among whites by a substantially smaller amount than the presence of blacks or Hispanics.

Table 5.5

Effect of Ethnic Composition of Census Tract on Social Trust Among Whites

	(I) No controls	(II) With Controls
tract %black	-0.740*** (0.040)	-0.253*** (0.048)
tract %Hispanic	-0.741*** (0.052)	-0.254*** (0.096)
tract %Asian	-0.203** (0.086)	-0.247** (0.125)
constant	2.524*** (0.007)	1.152*** (0.212)
observations	20,356	18,271
r-squared	0.029	0.169

*** p<0.01, ** p<0.05, * p<0.1

Notes: *Dependent variable is "How much can you trust people in your neighborhood?" Control variables are the same as in Putnam (2007, table 3), including individual- and tract-level income and education variables, but excluding the diversity index.*

Column II shows how the coefficients on the ethnic makeup of the tract change when a large set of control variables are added, including individual- and tract-level measures of education and income. When census tracts are matched on these other variables, the impact of Asians on whites' trust of their neighbors becomes no different from the impact of blacks and Hispanics. If higher Asian IQ explains the results from column I, the effect of IQ is entirely

accounted for by other observables like income and education in column II. This indicates the difficulty of measuring the independent impact of IQ. Assuming high IQ causes high income and education to some degree, these results are consistent with the hypothesis, though more empirical work is needed to confirm that racial diversity's negative impact on trust can be mitigated with intelligent non-whites.

CONCLUSION

This chapter has shown how the immigrant IQ deficit will have a pervasive impact on society. Many people are tempted to downplay or ignore this uncomfortable reality, but the issue should be of serious concern to policymakers. The topic that tends to dominate discussions of group differences in IQ—whether their source is nature or nurture—is actually unimportant from a policy perspective. The salient policy issue is the well-documented persistence of the IQ deficit. Whatever its cause, the deficit will increase undesirable social outcomes, such as low academic achievement, underclass behavior, and reduction of social capital within communities.

The next chapter shifts away from social consequences and focuses on the economic impact of the IQ deficit, specifically on the labor market.

Chapter 6: THE LABOR MARKET CONSEQUENCES

The social consequences of low-IQ immigration are unambiguously negative; however, the effect solely on the labor market is not immediately clear. This chapter leaves aside all of the social costs identified in the previous chapter and discusses immigration's effects, both positive and negative, on the labor market. All immigrant workers, no matter how intelligent or physically skilled, theoretically generate some net benefits for natives as long as they are employed. Adding additional workers to an economy should lower the price of labor and make production less costly. This hurts native workers who directly compete with immigrants but benefits the native economy as a whole. Generally speaking, a "good" labor market effect from a national perspective is one that generates a large native surplus—that is, extra money accruing to natives because immigrants are in the workforce—while minimizing the adverse impact on low-skill native wages.

The important question is which type of worker benefits the labor market the most—those who are skilled or unskilled? It is clear that, if immigrants affect the prevailing wage at all, they will always hurt the natives with whom they directly compete. High-skill immigrants will lower the wage of high-skill natives, and low-skill immigrants will lower the wage of low-skill natives. Much less clear is which type of immigrant maximizes the total native surplus. The answer depends on the character of the economy, as discussed in the next section. From a policy perspective, if low-skill immigrants tend to create a larger native surplus, then policymakers have a difficult balancing act to perform—increasing total gains requires an increasing burden on the native poor. However, if high-skill immigrants create the largest surplus, the negative wage effects will fall only on high-skill Americans, and distributional effects will not be a major concern.

As discussed below, in the modern American economy there can be little doubt that skilled workers provide the greatest net benefit to natives. Higher-IQ workers are also the ones who are most skilled. This chapter details the opportunity cost of favoring low-IQ over high-IQ immigrants for the American labor market.

INTRODUCTION

After briefly discussing the economic theory of immigration and introducing a three factor model of the labor market, this chapter attempts to answer three major questions:

(1) How do the native surplus and the distributional effects under our current immigration system compare to the surplus and distributional effects when selecting for education or selecting for IQ?

(2) How well can IQ tests identify future skilled workers, even before they acquire the education and experience that will allow them to work at skilled jobs?

(3) Does selecting for IQ affect the skills of second generation immigrants?

The conclusions are that (1) selecting for IQ or education produces a greater native surplus and a smaller low-skill wage reduction compared to the current immigration system. (2) IQ tests are nearly equivalent to knowing how much education an immigrant will acquire in the future in predicting the surplus generated. And, (3) selecting the first generation on the basis of IQ generates second generation skill more reliably than education selection.

Datasets. In this chapter, two different datasets are used to estimate the effect of IQ selection. Part 1 uses the National Longitudinal Survey of Youth (NLSY), a project that initially interviewed approximately 12,000 young adults in 1979 about education, work, and family life. Each respondent was given the Armed Forces Qualification Test (AFQT), a good measure of IQ as discussed in chapter 2. The benefit of the NLSY is that individual IQ scores are known at a young age, so that IQ and early education can be correlated with labor market success twenty

years later. The downsides of the NLSY are that natives must be used as proxies for immigrants, and the restricted age range of the participants limits its applicability to the labor market as a whole. I will use the NLSY to answer questions (1) and (2).

Part 2 employs actual immigrant data from the CPS March 2000 Annual Demographic Survey, with national IQ scores from Lynn and Vanhanen (2006) assigned to each immigrant on a country-by-country basis. The benefit of this dataset is that actual immigrants (rather than native proxies) are used over a full working age range of 18 to 64. Additionally, second generation immigrants can be identified based on questions about parents' places of birth. The drawback is that IQ scores for each immigrant are based on national averages, creating a more noisy relationship between wages and IQ. Also, CPS immigrants cannot be tracked over long periods of time. The CPS data will offer answers to questions (1) and (3).

The Model. Finding an immigration policy that maximizes the immigration surplus accruing to natives is not necessarily as simple as merely bringing in high earners (Borjas 1994a). Immigration increases the supply of labor, a key factor in production. If this influx lowers the prevailing wage, then the cost of production goes down and natives benefit through lower consumer prices. If the wage is not reduced, then the cost of production remains the same, and natives cannot benefit. The wage impact is measured by the elasticity of factor price for labor e_L, which tells us the percentage change in the wage given a 1% increase in the labor supply. As e_L becomes larger in (negative) magnitude, the more the wage is lowered by immigration, and the more natives benefit.

Estimating factor price elasticities is difficult, but an exhaustive survey by Hamermesh (1993, ch. 3) indicates some consensus that the price elasticity of skilled labor, e_{SS}, is more negative than the elasticity of unskilled labor, e_{UU}. Reasonable estimates of these factor prices

range from –0.2 to –0.6 for e_{UU}, and –0.5 to –1.0 for e_{SS}. These numbers are also used in Borjas (1995).

The intuition here is that skill and capital have gone from substitutes to complements over time. In the early part of the last century, a clothing manufacturer could hire either a skilled artisan or an unskilled laborer using a sewing machine. Today, however, sophisticated capital such as a computer often requires skilled labor to be utilized effectively. Now that skill and capital exhibit complementarity, the price of skilled labor is more sensitive to supply shocks. Skilled immigrants reduce the market wage, and thus the cost of production, by a greater percentage than do unskilled immigrants. Now, unlike the economy of a hundred years ago, an immigration policy that brings in skilled rather than unskilled workers will generate more gains for natives. These gains come from *high*-skill (rather than low-skill) native wage reductions.

A major difficulty in analyzing the "skilled" versus "unskilled" labor market lies in the actual definition of those terms. Hamermesh surveys papers that variously define the skill dichotomy as production versus nonproduction workers, blue collar versus white collar, educated versus uneducated, and low-wage versus high-wage. In this chapter I define skill using wages, with alternate models assuming 50% and 75% of the workforce is skilled. The fact that the definition of skilled is vague makes exact calculations of immigration's labor market impact impossible, but that should not prevent an investigation using reasonable estimates.

The model I use here is liberally borrowed from Borjas (1995). It is a three factor production model consisting of capital (K), skilled labor (L_s) and unskilled labor (L_u):

$$Q = f(K, L_s, L_u)$$

If we let b and β represent the fraction of skilled workers among natives (N) and immigrants (M) respectively, then:

$$Q = f(K, bN + \beta M, (1-b)N + (1-\beta)M)$$

Since *M* is essentially the change in labor supply caused by immigration, we differentiate *Q* to obtain the change in output:

$$\Delta Q = (K\frac{\partial r}{\partial M} + bN\frac{\partial w_s}{\partial M} + (1-b)N\frac{\partial w_u}{\partial M})M$$

Some algebraic manipulation leads to the following equation, where $m = M/(L_s + L_u)$:

$$\frac{\Delta Q}{Q} = -\frac{s_s e_{ss}\beta^2 m^2}{2p_s^2} - \frac{s_u e_{uu}(1-\beta^2)m^2}{2p_u^2} - \frac{\beta(1-\beta)m^2(s_s e_{su} + s_u e_{us})}{2p_s p_u}$$

s_s and s_u are the shares of national income held by skilled and unskilled workers, respectively. The variables p_s and p_u are the shares of the native workforce that are skilled and unskilled, respectively. In the last term, e_{su} and e_{us} are the cross-price elasticities of skilled and unskilled labor.

The companion formula for the percentage change in the low-skill wage is derived in Borjas (1998). It is:

$$\frac{s_u e_{us}\beta m}{2p_s} - \frac{s_u e_{us}\beta(1-\beta)m^2}{2p_s p_u} + \frac{s_u e_{uu}(1-\beta)m}{2p_u} - \frac{s_u e_{uu}(1-\beta)^2 m^2}{2p_u^2}.$$

PART 1: NLSY AND THE AFQT

This section uses the AFQT scores of respondents in the NLSY to generate a hypothetical class of highly intelligent immigrants. The fractions of skilled and unskilled immigrants are applied to the model above to calculate the immigration surplus and wage impact that would result.

Method. The main method used in the NLSY portion of the paper is relatively simple. First classify respondents in the NLSY as skilled, unskilled, or out of the labor force using wage data from the year 2000. Then take the top 10% of scorers on the AFQT and examine what fraction of these respondents fits each skill classification. Then plug into the above model the fractions of skilled and unskilled people in the top 10% of AFQT. The result is the immigration

surplus that would accrue to natives if immigrants had been limited to people with top-decile AFQT scores. Repeat the process by selecting the top 10% by education, and compare the resulting surplus against the AFQT method.

Prior Education or Eventual Education? I define prior education level as the number of years of education that immigrants have when they first enter the U.S. Eventual education is the amount of education they end up with after attending school in the U.S. The distinction is crucial, because people with greater cognitive ability are likely to pursue more education in order to gain the credentials needed for high-wage jobs. A major benefit of selecting for IQ is that immigrants without a solid prior education can acquire one in the receiving country. It makes little sense then to analyze the top 10% of immigrants in the NLSY by *eventual* education in comparison to the top 10% in IQ. After all, an immigrant's eventual education, unlike his IQ, is unknowable when he is first admitted to the U.S. This is why the ideal dataset would contain every immigrant's prior educational level. However, the NLSY has a limited age range. It consists of young Americans who were between the ages of 14 and 22 in 1979, which means that most of the immigrants in the sample already have at least some years of American education when they are first interviewed.

The method I employ here is to abandon the use of NLSY immigrants, who are too few in number in recent years to analyze properly. Instead, I examine the education level in 1980, the same year the AFQT was administered, of an unweighted cross-section of natives ages 15 through 23. I use these respondents as proxies for immigrants. Their IQ in 1980 is known, but their eventual educations are not. Young adults at this age range face an uncertain educational future. Some may drop out of high school, some may get a diploma, and some may go on to college. Much like immigrants entering the country for the first time, their education level may or may not change. The point is that we know very little about their eventual education in 1980,

but a lot about their IQ. Can IQ as measured in 1980 predict future wages as well as *future* education levels predict wages? The answer is "yes."

Other Data Issues. As mentioned above, AFQT and prior education are measured in 1980. Each is age-adjusted. Wages are measured in the year 2000, when the economy was about $9.8 trillion in size. Using the census figure of approximately 250,321,000 natives in the year 2000, along with the result from the CPS that about 46.25% of those natives are actively employed civilians, yields an estimated 115,773,500 natives in the workforce. According to the Census, there were 24.8 million immigrants ages 18-64 living in the US in 2000. The number of hypothetical immigrant workers in each simulation is calculated by multiplying 24.8 million by the predicted immigrant labor force participation rate, depending on the selection criteria.

Skilled laborers are defined in two different ways—as the top half of wage earners, and as the top three quarters. A skilled worker is defined as one with an hourly wage rate of at least $13 per hour or $8.65 per hour, for the 50% and 75% skilled assumptions, respectively. Skilled labor's share of national income in these cases is 52% and 63% respectively, using CPS data and Borjas's (1995) assumption of 70% of national income going to labor in general. Three different pairs of wage elasticities are used, as discussed in the literature review. Finally, although they make very little difference in the results, e_{su} and e_{us} are assumed to be 0.02 and 0.01 respectively. Hamermesh (1993, ch. 3) suggests these cross-elasticities are nonnegative and of small magnitude. The values themselves are adapted from Borjas (2003, 1367).

Results. The calculations that follow are meant to answer a hypothetical question—*if* the 24.8 million working-age immigrants living in the U.S. in 2000 had been selected by AFQT or education, what would the native surplus and wage effects have been? Table 6.1 first gives the skill profiles of hypothetical immigrants depending on the selection method.

Table 6.1
Skill Profile of Hypothetical Immigrants by Selection Method, NLSY data

Fraction of natives who are skilled:	ROW	COLUMN -> Select by:	I %skilled	II %unskilled	III %out of labor force
0.5	1.	eventual education	80.1%	10.8%	9.2%
	2.	AFQT	77.5%	12.0%	10.4%
	3.	prior education	65.2%	21.5%	13.3%
	4.	actual immigrants	60.1%	30.4%	9.5%
0.75	5.	eventual education	86.4%	3.5%	10.1%
	6.	AFQT	84.5%	5.1%	10.4%
	7.	prior education	78.8%	7.9%	13.3%
	8	actual immigrants	75.2%	15.3%	9.5%

Notes: Estmates are for a hypothetical immigrant population that is between 35 and 43 years old in the year 2000. Actual immigrants refer to NLSY immigrants, not a cross-section of immigrants in 2000.

The table looks complicated, so let us examine it in smaller pieces. Rows 1-4 are estimates using the assumption that half of the native workforce is skilled, while rows 5-8 are identical calculations assuming three-quarters are skilled. The rows represent hypothetical selection methods—the top 10% of eventual education by 2000, top 10% by AFQT score in 1980, and top 10% by "prior education" measured in 1980. The "actual immigrants" rows refer to all the immigrants who were originally interviewed in the NLSY, without any further selection criteria. In order to help with interpretation, take the number 10.8% in row 1, column II. This number means that 10.8% of immigrants would hold unskilled jobs in 2000 if they were all selected from the top decile of educational attainment, and half the native workforce is considered skilled. It is clear from the table that any of the three selection methods produces a

more skilled workforce than the actual immigrants observed, with education and AFQT significantly better than prior education as predictors.[32]

As discussed in the literature review, a more skilled workforce does not necessarily translate into a greater benefit for natives. To estimate the actual surpluses, we need to plug the skill profiles from table 6.1 into the labor market model discussed above. Table 6.2 shows the results. It is similar in structure to table 6.1, except now the columns are different possible wage elasticities that affect how skills translate into surpluses.

Table 6.2
Estimated Immigration Surplus Accumulating to Natives, NLSY data (year 2000 dollars, in billions)

			COLUMN ->	I	II	III
				Wage Elasticities: (unskilled, skilled)		
Fraction of natives who are skilled:	ROW	Select by:		(-0.2, -0.5)	(-0.4, -0.75)	(-0.6, -1.0)
0.5	1.	eventual education		105.2	158.3	211.4
	2.	AFQT		98.9	148.9	198.9
	3.	prior education		71.0	107.6	144.1
	4.	actual immigrants		60.2	92.1	123.9
0.75	5.	eventual education		66.2	99.4	132.7
	6.	AFQT		63.2	95.1	127.0
	7.	prior education		55.4	83.6	111.8
	8	actual immigrants		49.7	75.7	101.6

Notes: Assumes a $9.8 trillion economy and 31 million immigrants, with a hypothetical immigrant population that is between 35 and 43 years old in the year 2000. Actual immigrants refer to NLSY immigrants, not a cross-section of immigrants in 2000.

[32] It may be surprising to see that even actual immigrants outperform a cross-section of natives, who by definition are only 50% or 75% skilled depending on the assumptions. But keep in mind that NLSY respondents are in their prime working age when measured in 2000, while the working population as a whole is between the ages of 18 and 64. Also, due to dropouts from the survey, the immigrants in 2000 were significantly smarter on average than those measured in 1980. As stated in the text, the absolute numbers in the tables are much less important than the relative comparisons.

Again to help with interpretation, look at row 2, column III. The number there means that the native surplus in 2000 would be $198.9 billion dollars if immigrants had come from the top decile of the AFQT distribution, assuming 50% of natives were skilled, the unskilled wage elasticity was –0.6, and the skilled elasticity was –1.0. Similarly, the surplus would be $144.1 billion if all of the same assumptions held true, except that immigrants had been selected on the basis of their prior education rather than by their AFQT score. Rows 4 and 8 represent the surplus that would be created by immigrants who have the skill profile of the actual immigrants aged 35-43 living in the U.S. in the year 2000.

Table 6.3

Aggregate Change in *Unskilled* Native Wages (year 2000 dollars, in billions)

NLSY data, year 2000 dollars in billions

			COLUMN ->	**I**	**II**	**III**
				Wage Elasticities: (unskilled, skilled)		
Fraction of natives who are skilled:	**ROW**	Select by:		(-0.2, -0.5)	(-0.4, -0.75)	(-0.6, -1.0)
0.5	**1.**	eventual education		-0.7	-1.9	-3.0
	2.	AFQT		-0.9	-2.2	-3.5
	3.	prior education		-1.9	-4.1	-6.4
	4.	actual immigrants		-2.7	-5.8	-8.8
0.75	**5.**	eventual education		-0.1	-0.2	-0.3
	6.	AFQT		-0.1	-0.3	-0.4
	7.	prior education		-0.2	-0.4	-0.7
	8	actual immigrants		-0.4	-0.8	-1.3

Notes: Figures refer to total amount of wealth transferred from low-skill natives to immigrants and native employers, not percentages.

Finally, table 6.3 shows how unskilled natives are affected by each immigrant selection method. Looking at column III, the total wage losses (in billions) suffered by unskilled natives

would be about $3.5 billion with AFQT selection, but $8.8 billion under the current system. Clearly, more unskilled immigrants lead to greater losses for unskilled natives.

An important caveat is that these calculations assume all 24.8 million immigrants have the same work habits as people between the ages of 35 and 43. This is not entirely realistic, as many immigrants will have more or less work experience compared to that group. The reason for the assumption is the limited age range of the NLSY, but it should not be viewed as a fundamental weakness.

The purpose here is to generate comparisons across selection methods, not to examine absolute amounts. I could have chosen any number of immigrants in the simulation to facilitate comparisons. 24.8 million, being the actual number of working-age immigrants in 2000, was simply used for convenience. One can think of the estimates above as the surplus *if* the 24.8 million working-age immigrants in the U.S. were all replaced by adults ages 35-43 who were selected for their education or IQ.

The major takeaway from these results is that selecting for eventual education is only marginally superior to selecting for AFQT, while using prior education as a selection criterion is significantly inferior to AFQT. It appears that nearly the same surplus can be achieved through IQ selection as can be predicted by the eventual education of immigrants. Any of the three selection methods creates a larger surplus (and smaller wage reductions for the unskilled) than actual immigration.

Can Natives Really be Treated as Proxies for Immigrants? One of the major assumptions made is that immigrants and natives with the same talents will have the same success in the labor market. Is this realistic? Not in extreme cases. For example, an illiterate 50-year-old peasant from an impoverished country probably will not come to the U.S. and immediately acquire a skilled job, regardless of how high his IQ is. On the other hand, a very

smart and energetic 20-year-old immigrant could quite plausibly learn English, acquire useful training, and take on a skilled job within a short time. The analysis in this paper is more relevant to the latter case, when immigrants come to the U.S. at a young age and gain education and work experience. The question I can test here is whether young immigrants (those in the NLSY) will have the same skill profile as natives with the same ability.

Table 6.4 is the same as table 6.1, except now the actual immigrants from the NLSY are used rather than the proxy natives, and selection criteria is increased to the top 25% to create a larger sample. For example, selecting by AFQT means evaluating the skill profile of only immigrants who are in the population's top quarter in AFQT. The table reports the percentages of natives that are skilled in each category subtracted from the percentages of skilled immigrants. For example, row 2, column I indicates that the fraction of skilled immigrants is 2.43 percentage points higher than the fraction of skilled natives when selecting for AFQT. Similarly, natives exceed immigrants by 0.34 percentage points in the fraction that are unskilled.

Table 6.4
Immigrant - Native Difference in Skill Profile, in percentage points

		COLUMN ->	I	II	III
Fraction of natives who are skilled:	ROW	Select by:	skilled	unskilled	out of labor force
0.5	1.	eventual education	12.58	-8.79	-3.79
	2.	AFQT	2.43	-0.34	-2.09
	3.	prior education	5.65	-2.97	-2.68
0.75	5.	eventual education	8.06	-4.27	-3.79
	6.	AFQT	3.95	-1.86	-2.09
	7.	prior education	2.93	-0.24	-2.68

Note: Estmates are for a hypothetical immigrant population that is between 35 and 43 years old in the year 2000.

Clearly, immigrants are actually *more* skilled and more likely to be in the labor force than comparable natives in the NLSY. The gaps are quite substantial when selecting for eventual education. These exact numbers should not be taken entirely seriously, because there is only a small sample of immigrants that can be used. Table 6.4 simply shows that there is no prima facie evidence that immigrants underperform natives of comparable talent and experience.

PART 2: RESULTS WITH THE CPS AND IQ-BY-COUNTRY ESTIMATES

This section re-answers question (1) with different data, and then it suggests an answer to question (3). As mentioned in the introduction, I use actual immigrant wage data from the CPS, and each immigrant is assigned an IQ score based on his place of birth. The national IQ scores are from Lynn and Vanhanen (2006), discussed in depth in chapter 2, and the complete list of the countries and their corresponding IQs used in this chapter can be found in Appendix C. When re-answering question (1) with the LV data, this method sacrifices an exact IQ score in exchange for the benefit of using real immigrants with a more realistic age range.

Immigrant Results. Table 6.5 compares selecting immigrants from countries with average IQs higher than the U.S. median to the actual surplus generated by current immigrants. As the table indicates, selecting for IQ still creates a substantially more skilled group of immigrants compared to the present class. Unfortunately, the national IQ range is too small, and high-IQ countries are too few, in order to break down the IQ selection into smaller groups. Additionally, since the CPS is not longitudinal, there can be no discussion of prior versus eventual education. Nevertheless, these CPS data affirm the NLSY answer to question (1). Table 6.6 converts the skill profiles from table 6.5 into the dollar value in billions of the native surplus produced, and Table 6.7 shows the impact on unskilled natives.

Table 6.5
Skill Profile of Immigrants by Selection Method, CPS Data

Fraction of natives who are skilled:	ROW	Select by:	COLUMN -> I %skilled	II %unskilled	III %out of labor force
0.5	1.	IQ > U.S. median	38.5%	27.4%	34.1%
	2.	all immigrants	27.3%	38.2%	34.5%
0.75	3.	IQ > U.S. median	51.7%	14.2%	34.1%
	4.	all immigrants	41.9%	23.6%	34.5%

Notes: Estimates are for actual immigrants ages 18 to 64 living in the US in 2000. IQ is based on Lynn and Vanhanen's 2006 IQ-by-country estimates.

Table 6.6
Estimated Immigration Surplus Accumulating to Natives, CPS data (year 2000 dollars, in billions)

Fraction of natives who are skilled:	ROW	Select by:	COLUMN -> I Wage Elasticities: (unskilled, skilled) (-0.2, -0.5)	II (-0.4, -0.75)	III (-0.6, -1.0)
0.5	1.	IQ > U.S. median	27.7	42.9	58.2
	2.	all immigrants	16.2	26.5	36.9
0.75	3.	IQ > U.S. median	25.8	39.6	53.3
	4.	all immigrants	18.0	28.6	39.2

Notes: Estimates are for the 24.8 million actual immigrants ages 18 to 64 living in the US in 2000. IQ is based on Lynn and Vanhanen's 2006 IQ-by-country estimates.

Table 6.7

Aggregate Change in *Unskilled* Native Wages by Immigrant Selection Method
CPS data, year 2000 dollars in billions

			COLUMN -> I	II	III
			Wage Elasticities: (unskilled, skilled)		
Fraction of natives who are skilled:	ROW	Select by:	(-0.2, -0.5)	(-0.4, -0.75)	(-0.6, -1.0)
0.5	1.	IQ > U.S. median	-2.7	-5.6	-8.5
	2.	all immigrants	-3.7	-7.5	-11.4
0.75	3.	IQ > U.S. median	-0.4	-0.8	-1.3
	4.	all immigrants	-0.6	-1.3	-1.9

Notes: Figures refer to total amount of wealth transferred from low-skill natives to immigrants and native employers, not percentages.

Second Generation Results. The CPS data also identify second generation immigrants, people who were born in the U.S. but have at least one parent who was born in a foreign country. The second generation is important to any immigrant selection system, because the acceptance of a single immigrant means accepting several subsequent generations of people as well. If skills fail to transfer from one generation to the next, the gains from any selection system could quickly evaporate. To examine how selection could influence the skills of the second generation, I assigned each second generation immigrant in the CPS parental IQ and parental education scores. Parental IQ is based on the national IQ of the country where the parent was born.

Parental education is assigned in a similar fashion. Immigrants from the 1970 census are likely to be the parents of the second generation in the 2000 CPS. I used the average educational level by country of origin of first generation immigrants in the 1970 census to assign a parental education value to the second generation in the 2000 CPS. (See Appendix C for a list of average education and IQ by country.) For example, if a second generation individual in the CPS has a

Chinese-born parent, then his parental IQ score would be the Chinese IQ given in Lynn and Vanhanen, and his parental education score would be the average education of Chinese immigrants in 1970.[33]

Tables 6.8, 6.9, and 6.10 show the skill profile, surplus, and wage impact, respectively, of second generation immigrants based on parental selection. Row 1 of table 6.8 shows the skill profile of second generation immigrants who have an immigrant parent from a higher IQ country. Row 2 shows the skill profile if the selection system is changed to parents with higher education countries. Finally, the last row shows the actual 10.5 million second generation immigrants living in the U.S. in the year 2000.

Table 6.8
Skill Profile of *Second Generation* Immigrants by Selection Method

		COLUMN ->	I	II	III
Fraction of natives who are skilled:	ROW	Select by:	%skilled	%unskilled	%out of labor force
0.5	1.	parental IQ > U.S. median	44.6%	25.2%	30.2%
	2.	parental education > U.S. median	39.6%	27.1%	33.4%
	3.	all second generation immigrants	36.1%	32.1%	31.8%
0.75	4.	parental IQ > U.S. median	57.0%	12.7%	30.2%
	5.	parentel education > U.S. median	48.4%	18.3%	33.4%
	6.	all second generation immigrants	50.7%	17.5%	31.8%

Notes: Estimates are for actual second generation immigrants ages 18 to 64 living in the US in 2000. Parental IQ is based on Lynn and Vanhanen's 2006 IQ-by-country estimates. Parental education is measured post-migration.

Clearly, second generation immigrants whose parents possessed high IQ continue to show substantially higher levels of skill than the second generation as a whole. Even more interestingly, parental education appears to transfer skills to the next generation less reliably than

[33] If the second generation individual has one immigrant and one native parent, only the immigrant IQ and education scores are counted. If the individual has two immigrant parents from different countries, the higher IQ or education parent is used.

parental IQ. While selecting for either eventual education or IQ can generate benefits, only IQ selection substantially maintains those benefits into at least one more generation.

Table 6.9

Estimated *Second Generation* Immigrant Surplus Accumulating to Natives (year 2000 dollars, in billions)

		COLUMN ->	I	II	III
			Wage Elasticities: (unskilled, skilled)		
Fraction of natives who are skilled:	ROW	Select by:	(-0.2, -0.5)	(-0.4, -0.75)	(-0.6, -1.0)
0.5	1.	parental IQ > U.S. median	7.5	11.6	15.6
	2.	parental education > U.S. median	6.1	9.4	12.7
	3.	all second generation immigrants	5.2	8.2	11.1
0.75	4.	parental IQ > U.S. median	6.5	9.9	13.2
	5.	parental education > U.S. median	4.8	7.4	10.0
	6.	all second generation immigrants	5.2	8.0	10.8

Notes: Estimates are for actual second generation immigrants ages 18 to 64 living in the US in 2000. Parental IQ is based on Lynn and Vanhanen's 2006 IQ-by-country estimates. Parental education is measured post-migration.

Table 6.10

Aggregate Change in *Unskilled* Native Wages Due to *Second Generation* Immigrants
CPS data, year 2000 dollars in billions

		COLUMN ->	I	II	III
			Wage Elasticities: (unskilled, skilled)		
Fraction of natives who are skilled:	ROW	Select by:	(-0.2, -0.5)	(-0.4, -0.75)	(-0.6, -1.0)
0.5	1.	parental IQ > U.S. median	-1.2	-2.5	-3.7
	2.	parental education > U.S. median	-1.3	-2.7	-4.0
	3.	all second generation immigrants	-1.5	-3.1	-4.8
0.75	4.	parental IQ > U.S. median	-0.2	-0.4	-0.5
	5.	parental education > U.S. median	-0.3	-0.5	-0.8
	6.	all second generation immigrants	-0.2	-0.5	-0.7

Notes: Figures refer to total amount of wealth transferred from low-skill natives to immigrants and native employers, not percentages.

CONCLUSION

This chapter has used a three-factor model of the U.S. labor market to compare the native surplus and wage reductions due to immigration under different selection criteria. I find that selecting for AFQT and eventual education produce substantially greater total gains for natives than selecting for prior education. Additionally, all of the three selection methods lead to more overall native gains and smaller wage reductions for the unskilled compared to the actual immigrant cohort from the NLSY. Even when using IQ-by-country estimates for each individual immigrant, IQ selection still produces a much larger surplus than the status quo.

Most significantly, I find that a test of IQ taken twenty years prior to measuring wages is nearly as good a predictor of labor market success as the eventual education of the worker. This finding suggests that immigrants with high IQs but only modest schooling can, given a period of work experience and training in the U.S., become as productive as the most educated citizens. Finally, the superior skills of high-IQ immigrants appear to transfer well to the second generation. By taking in lower-IQ immigrants instead of more intelligent people, the U.S. misses out on many economic gains, and low-skill Americans suffer more.

Chapter 7: IQ SELECTION AS POLICY

The dissertation began in Chapter 1 by summarizing the science of IQ, using a statement by the American Psychological Association as the framework for the discussion. Chapter 2 analyzed a variety of datasets that included scores on *g*-loaded tests from representative samples of immigrants. The immigrant population was found to have an average IQ somewhere in the low 90s, below the native white average of 100. Chapter 3 used the experience of Hispanic Americans to confirm that today's immigrant IQ deficit is not ephemeral or illusory as it was for European immigrants in the early twentieth century. Chapter 4 explored the possible causes of the IQ deficit, which likely involves a complex interplay between environmental deprivation and genetic differences. Chapter 5 discussed the causal role of IQ in helping to determine myriad life outcomes, warning in particular that low immigrant IQ has helped create a new underclass and could undermine social trust. Chapter 6 used an economic model to contrast the labor market impact of high-IQ hypothetical immigrants with other selection methods and with the status quo.

My contribution has been to identify the immigrant IQ deficit using several different tests, and to discuss the effects, some obvious and some more subtle, of the deficit on the economy and on society. But identifying a problem and discussing its impact is in some sense the easy part of public policy research. Finding a practical solution is the harder step. This concluding chapter is not a formal policy analysis or even a detailed proposal. Instead, the chapter simply explores the proposition that immigration policy should select for IQ, and it discusses the ethical, legal, and political issues raised by such a policy. It is the beginning of a needed discussion. The argument I advance in this chapter is that, recognizing the many practical difficulties that would have to be overcome, selection for IQ could in theory make our

immigration policy both beneficial *and* charitable, fulfilling two goals that are often considered conflicting.

THE PHILOSOPHICAL PERSPECTIVE

I begin with a short review of the philosophy of immigration. The literature on immigration is immense, but it can be summarized by briefly examining four of the most popular ethical systems.

Utilitarian. A global utilitarian would assert that everyone in the world is entitled to equal consideration of interests. From that perspective, any kind of immigration restriction is based on the morally irrelevant factor of nationality. This implies that a Nigerian has the same right to move to New York as a Pennsylvanian does, but some utilitarians regard that analysis as too simplistic. Family members and neighbors relate to each other more readily; therefore, it makes more sense for communities to favor their own members to some degree (Singer 1993, 233).

Libertarian. Now consider libertarianism in the tradition of Robert Nozick. In a libertarian world, the government can legitimately act only as a "night watchman," doing nothing other than protecting property and keeping the peace. Although Nozick does not directly discuss immigration in his classic *Anarchy, State, and Utopia* (1974), other philosophers have extended his reasoning to a global scale. Since international migration does not impinge on any individual's freedom, they reason, a libertarian government cannot legitimately restrict it. In fact, restriction implies collective ownership of property by the state, a notion that libertarians like Nozick reject (Carens 1987).

This open-borders view is disputed by other libertarians, most notably Hans Hermann Hoppe. Hoppe argues that our current immigration system amounts to both forced exclusion *and* forced integration (2001, 142). The reason lies in the nature of public property. Regardless

of who is admitted, some natives will object to immigrant presence on public property (forced integration), and other natives will wish different immigrants could arrive (forced exclusion). In the ideal libertarian world where all property is private, landowners would carefully monitor and evaluate people wishing to enter their territory, eschewing open borders for a selection system.

Rawlsian. Liberal egalitarians in the Rawlsian tradition are similarly conflicted over the immigration issue. Rawls' veil of ignorance, behind which no one can see his own natural talents and life circumstances, tends to induce risk aversion. Under a Rawlsian system, the way we feel about public policy when behind this veil is a more just approach to setting up societies. Like Nozick in his magnum opus, Rawls does not discuss immigration in detail in his *A Theory of Justice* (1971), but other philosophers have applied Rawls' thinking to justice across nations (Carens 1987). If a person were going to be born in a random country, this argument goes, the real possibility of subsistence living in a remote African jungle might compel him to support open borders. This implies that immigration would be unrestricted.

But Rawls himself in a later work rejects applying his original position to the international arena, arguing instead that states have special obligations to their citizens (1999, 8), including protecting their political culture (39n). Rawls says that governments must take care of their own territory without using emigration as a crutch to maintain illiberal policies (39). He also claims that any nation with a liberal government and sound institutions can be a just society, regardless of resource endowment (1998, 107). This suggests that immigration would cease to be an important issue in a Rawlsian world, allowing individual nations to maintain their own cultures and identities via restriction.

Communitarian. The notion of special obligations and group bonds is a common factor underlying the argument for restriction. Utilitarians recognize that neighbors are better providers than strangers, some libertarians acknowledge that private communities can assert

group interests, and among liberal egalitarians even Rawls himself sees nation-states as having special commitments to their own citizens. All of these positions suggest an underlying justification for regulating immigration—nations have special obligations that compel them to act in their citizens' best interest. If restricting immigration is in the national interest, then it is a defensible policy.

The most prominent defense of national interests, and consequently of the right to immigration restriction, is Michael Walzer's *Spheres of Justice* (1983). Walzer likens nations to neighborhoods, clubs, and families, all of which have the right to regulate their membership in varying ways. He considers the regulation of group membership to be crucial to "complex equality"—the separation of justice into various spheres of life, from work, to school, to kinship. Under this theory, "communities of character—historically stable, ongoing associations of men and women with some special commitment to one another and some special sense of their common life" become primary goods (1983, 62). Thus, for Walzer, regulating membership in every sphere, including at the level of nations, is essential to justice.

A Framework for Immigration Policy

While there is no philosophical consensus on immigration, using immigration to advance national interests can be legitimate under many different assumptions. For purposes of this discussion, it is sufficient to say that philosophers have identified both the welfare of the nation and the welfare of potential immigrants as important considerations. Intuitively, this conforms to how most Americans view immigration policy. They want a policy that helps themselves, helps other Americans, and helps foreigners, each to varying degrees.

I propose a general principle that conforms to that desire. The U.S. should first define exactly what it wants for itself from its immigration policy. Then, design a selection system that meets those goals, while still providing substantial benefit to potential immigrants. In

mathematical terms, the U.S. should maximize the welfare of its immigrants, subject to the constraint that the selection system meets the country's own goals. Literally optimizing this abstract objective function is probably not possible, but it is a worthy ideal to work toward. As a simple example, if the U.S. decides that its only goal is to add more bricklayers to the country's workforce, then it should take some of the world's poorest and most disadvantaged bricklayers.

To further motivate this principle, consider the following simple thought experiment. Imagine a small business looking to hire a new vice president. The owner can hire either Rich or Susan. Based on experience and qualifications, Rich will make a far better vice president than Susan, but Rich is also the privileged son of a Fortune 500 CEO. He has no need for the vice president's salary, as he already receives a substantial allowance from his father. On the other hand, Susan is a single mother who often has trouble paying her rent. Whom should the owner hire? The answer should be obvious. Although he sympathizes with Susan, the owner must do what is best for his company by hiring Rich. After all, business is business. No company that hired out of compassion rather than self-interest could long survive.

But now consider the same scenario with one key difference. Rich is still more privileged than Susan, but this time the owner has determined objectively that both people would perform about equally as vice president. Now whom does he choose? Again, the answer should be obvious. Rich needs the work much less than Susan does, so Susan should be the choice. The owner has maximized the welfare of his potential employees, subject to the constraint that they in fact help his business. My argument for immigration exactly parallels this story. Require that immigrants make a certain positive contribution to one's country, but then choose those applicants who would most value admission. Specifically, if the U.S. wants its immigrants to be rich and prosperous, it should select immigrants who will become rich in the U.S. but who would otherwise be poor in their native countries.

How Should We Choose Immigrants? Among the major immigrant-receiving Western countries today, there are two main methods for immigration selection, but neither satisfies the principle I described above. Some countries, such as the U.S., primarily emphasize family reunification and low-skill employment. Others, like Canada and Australia, have points systems that encourage highly-educated immigrants. None of these countries is exclusively devoted to either system, and many other idiosyncratic factors are present as well, but the low- versus high-skill dichotomy is a useful simplification. Table 7.1 illustrates the differences.

Table 7.1

Percentage of New, Legal Permanent Residents By Immigration Category in 2006

Country	Economic	Family	Refugee	Other
Australia	60.5	29.8	8.7	1.1
Canada	54.9	28.0	12.9	4.1
United Kingdom	23.7	44.5	22.8	9.1
United States	12.6	63.4	17.1	7.0

Source: See note.[34]

Economic considerations prevail in Australia and Canada, while family reunification dominates the American immigration system. The UK falls between these extremes, but closer to the

---

[34] The source for the Australia data is a 2008 "Immigration Update" report by the Australian Department of Immigration and Citizenship, table 1.5.
http://www.immi.gov.au/media/publications/statistics/immigration-update/update_june07.pdf

Canadian data are from this website maintained by Citizenship and Immigration Canada:
http://www.cic.gc.ca/English/resources/statistics/facts2006/permanent/01.asp

UK data are from a 2007 "Control of Immigration" report by the UK Home Office, table 5.4.
http://www.official-documents.gov.uk/document/cm71/7197/7197.pdf

American data are from a 2007 "Yearbook of Immigration Statistics" report by the Department of Homeland Security, table 9.
http://www.dhs.gov/xlibrary/assets/statistics/yearbook/2006/OIS_2006_Yearbook.pdf

Figures for Australia are based on combined 2006 and 2007 data, and they exclude immigrants from New Zealand, which has an open border agreement with Australia. Numbers for the UK also exclude members of the European Economic Area and Switzerland, for the same reason.

American model. In most cases, economic immigrants are educated, high-skill workers. Family reunification in the U.S., while officially unrelated to economic concerns, is a magnet for low-skill workers and their extended families.

Several analysts have proposed that the U.S. increase its emphasis on educated immigrants.[35] Given the high correlation between education and IQ, such a system certainly would begin to reverse the immigrant IQ deficit, without making IQ an explicit policy concern. But one problem with this Canadian- and Australian-style education selection is that it severely limits the pool of available immigrants. Accepted applicants tend to be from other developed countries, or they are a part of a small elite from developing countries. In other words, immigrants admitted under points systems tend to be those who are least likely to be escaping poverty and disadvantage. The Canadian and Australian systems unnecessarily cast aside the welfare of potential immigrants. In terms of the thought experiment, they take Rich without ever even considering Susan.

Now consider the U.S. and Britain, which have the opposite of a skill-based policy. These countries emphasize low-skill employment and family reunification. This type of system is beneficial to impoverished migrants, but it violates the principle described above, which says that immigration should be constrained to always benefit the receiving country. As the previous chapters have shown, current immigrants to the United States are less intelligent on average than white natives, which leads to less economic assimilation, more underclass behavior, and several other negative outcomes. It is clear that, at the very least, there is room for improvement. The United States is hiring Susan even when Rich is much more qualified.

---

[35] See Borjas 1999, ch. 11; Malanga 2007; and the report of the U.S. Commission on Immigration Reform at http://www.utexas.edu/lbj/uscir/exesum95.html

There appears to be an irreconcilable conflict here between economics and deference to the poor. A low-skill immigrant rarely becomes a high-skill immigrant after migrating. Most Western countries have dealt with this problem inefficiently, by creating two classes of immigrants. One class is allowed to immigrate for charitable reasons, and the other class is expected to be high-skill workers. As table 7.1 indicated, most Western countries simply differ on which class of immigrant they prefer more. There is, however, a selection factor that could potentially unite these conflicting goals. That factor is IQ.

IQ AND IMMIGRATION

We have seen that IQ is a reliable and valid operational measure of intelligence, and that it is correlated with economic success. It can also be measured in ways that do not depend on schooling—for example, the highly *g*-loaded Ravens' Matrices require no literacy whatsoever. As an ability measure that is more independent of socioeconomic circumstances than educational attainment, IQ could help us identify immigrants who will make a substantial contribution despite their disadvantaged circumstances. Use of IQ tests could help us to meet the two concerns about immigration policy that were once thought mutually exclusive, and it comes closer to fulfilling the constrained optimization problem described above, where immigrant welfare is maximized while still benefiting the U.S.

Consider again the low-skill immigration policy of the U.S. Selection by IQ would increase immigrant talent without always shutting out those with little education. Mexicans, for example, tend to be among the least educated immigrants. Under Canadian-style education selection, very few Mexicans would be granted entry.[36] Using the IQ criterion, however, the most intelligent Mexicans could still immigrate, despite their disadvantaged background.

---

[36] According to the 2001 Census, just 0.01% of Canadians were of Mexican origin. In contrast, over 3.7% claimed Chinese ancestry. This indicates how a points system can strongly affect the

Therefore, the use of IQ test scores could actually help to level the playing field for potential immigrants all over the world. It is more egalitarian than elitist. Even those without access to good educations or career paths may have an opportunity to show their potential. For example, despite its low average IQ, there are over one million sub-Saharan Africans alone who have IQs greater than 115, which is one standard deviation above American whites. As chapter 3 pointed out, improved material conditions in Africa would make that available number even higher. Intelligent people from higher-IQ regions are even more numerous.[37]

It is important to note that IQ and socioeconomic status are correlated even in generally poor areas. The small group of elites in the third world are likely to be among the smartest in their countries. It is also possible that traditional class structures, such as the caste system in India, developed around IQ differences, so that the Brahmins have genetic as well as social advantages over the Dalits. However, given the lack of economic development and availability of education in many countries, the level of "cognitive stratification"—that is, the tendency for people to be sorted by their raw intellectual ability into appropriate educational and career tracks—must be substantially lower in undeveloped countries compared to developed ones.[38] There should be no shortage of underprivileged, high-IQ applicants for immigration.

Theoretical Difficulties. It is natural to be uncomfortable with immigration selection for IQ. Given the American Dream and the "pull yourself up by your bootstraps" national creed, Americans are not receptive to using a trait that is heritable and unchangeable (by

---

national background mix of immigrants, especially considering the proximity of Latin America to Canada and the porous North American borders. Cited from Statistics Canada: http://www12.statcan.ca/english/census01/products/highlight/ETO/Table1.cfm?Lang=E&T=501&GV=1&GID=0

[37] Sub-Saharan African IQ is about 70 according to LV. About 0.135% of the population has an IQ 115 or higher in a normal IQ distribution with mean of 70 and standard deviation 15. 0.135% multiplied by an estimated population of 770.3 million gives 1.04 million people.

[38] Increasing cognitive stratification in the U.S. is a major theme of *The Bell Curve*.

adolescence) to differentiate people. But superior cognitive ability is not some kind of free ticket to prosperity. If we define the American Dream as success based on ability and hard work rather than social circumstances, then IQ selection merely increases the chances that the Dream will be fulfilled for each immigrant.

The notion that IQ is an unacceptable criterion for selection because it is unchangeable is an especially inconsistent argument from those who support an education-based system. The reality is that a person's educational level while living in an impoverished region is just as unchangeable as his IQ. The chance of getting a college diploma is essentially zero, even for the very intelligent, in many parts of the world. Education selection necessarily ignores people in those circumstances, while IQ selection gives them consideration.

Visceral opposition to IQ selection can sometimes generate sensationalistic claims—for example, that this is an attempt to revive social Darwinism, eugenics, racism, etc. Nothing of that sort is true. Group differences in intelligence do exist, but, as I emphasized throughout the text, that does not mean that any individual should ever be judged on the basis of group membership. An IQ selection system could utilize individual intelligence test scores without any resort to generalizations.

A more substantive concern about IQ selection involves "brain drain"—that is, depriving poor countries of their smartest people. If Microsoft or Google were to offer a scholarship program to the smartest Americans from the poor Appalachian region of the country, fears of "brain drain" from Appalachia would be far outweighed by the enthusiasm for those who were finally getting an opportunity.[39] Brain drain would be more worrisome if poorer

---

[39] Henry Chauncey, first president of the Educational Testing Service, had a similar goal. According to Lemann (1999), Chauncey was driven to uncover the best and brightest regardless of social background. He insisted that the SAT be designed as a test of mental ability, not achievement. The degree to which the SAT meets that goal is a matter of controversy today.

countries did not lack the economic and social infrastructure to develop many of their best and brightest. But if enough immigrants were carefully selected from outside a poor country's elite circles, then the cognitive skills of these high-IQ immigrants would not be especially missed. In contrast, Canadian-style education selection inevitably removes some of the few educated elites that poor countries have.

Practical Difficulties. I believe there is a strong case for IQ selection, since it is theoretically a win-win for the U.S. and for potential immigrants. Practically speaking, however, it is a political non-starter because of opposition that I have already discussed. One way to at least blunt the negative reaction is to drop the use of the word *IQ* and to replace it with *skill*. A new immigration policy could use "skill tests" to find disadvantaged people with "raw skill." The tests would still be ordinary intelligence tests, but the emotional baggage that the term *IQ* sometimes carries with it would be much reduced.

The tests themselves could be administered at embassies and consulates, or even over the internet. As described above, a test like Ravens' Matrices, which requires no knowledge of words or numbers, could be used to ensure cultural fairness. If some degree of bias against certain groups is still discovered, applicants from the affected groups could have their scores bumped up by the necessary amount to compensate.

In terms of test administration, however, there is the problem of cost. Testing is a highly efficient screening process used by many large organizations, but it still carries a price tag. When a government agency administers the tests, the cost will be higher still. Here, education selection has the advantage over IQ selection, because education selection is free. A formal policy analysis of IQ selection would need to consider the cost of testing, possibly by examining how the State Department administers its foreign service exam, or how costly the citizenship tests used by the INS are.

An additional difficulty is how to integrate IQ selection into an immigration policy that has several different facets. Illegal immigration, for example, is a major issue that I cannot address here, except to say it must be controlled in order for *any* policy to work as intended. Additionally, other commentators will offer various 'X' factors as alternative selection criteria. These X's can range from increasing racial diversity, to filling labor shortages, to unifying extended families. Fortunately, considering IQ does not preclude the use of other factors. Highly intelligent people can be found all over the world, with all sorts of physical and cultural characteristics. If X is increasing racial diversity, then we should ensure our racially diverse immigrant class is also very smart. If X is filling the labor shortage in the construction industry, then we should find the most intelligent construction workers. Use of IQ as one selection factor is compatible with most any X.

CONCLUSION

As the previous six chapters have discussed, today's immigrants are not as intelligent on average as white natives. The IQ difference between the two groups is large enough to have substantial negative effects on the economy and on American society. The deficit cannot be dismissed as meaningless or transient. It is transferred across generations—whether via genes, environment, or both—in a manner that we do not yet know how to prevent. Although this is a depressing conclusion, it does help us focus on a new opportunity. In trying to reverse the cognitive decline of immigrants, we could begin to seek out underprivileged people who have the raw mental ability to achieve personal success, while still helping ourselves at the same time.

Appendix A : TABLE OF NATIONAL IQ SCORES

The following table presents technical information used for the national IQ calculations in chapter 2. Lynn and Vanhanen's national IQ scores are given for countries recognized by the CPS. Every country in LV's dataset is listed here for the interested reader, but the only countries used in the analysis are those with corresponding CPS codes.

The table also shows how countries were grouped together. Since they are European-derived nations, Canada, Australia, and New Zealand are grouped with Europe. Also, because of its importance to U.S. immigration and its ethnic and cultural differences with the rest of North America, Mexico is listed in its own separate category. Overall, the groupings were designed to reflect similar peoples rather than just similar geography.

Some immigrants in the CPS reported regions rather than actual countries of birth. Wherever possible, these immigrants were given regional IQ scores that are based on averages of nearby countries. Regions are placed in italics in the table, and the calculation of their IQ scores are described below. In some cases—namely, with "North America," "Asia," "Middle East," "Other Africa," and "Elsewhere"—not enough information was given to create a reasonable IQ score for the individual.

Observations were dropped if they were ambiguous or missing. The dropped data amounted to 993 cases out of 24,492 immigrants in the 2006 CPS. LV had no IQ data for Azores or Palestine even though these territories are listed in the CPS. Their IQ scores are imputed, and they are listed with a double asterisk. The imputation method is described below the table. Note that people born in U.S. territories—American Samoa, Guam, Northern Mariana Islands, Puerto Rico, and U.S. Virgin Islands—are technically not immigrants and are not counted as such here. Immigrants are defined here as people who answered 4 or 5 (non-native) to the question about their citizenship status (variable PRCITSHP).

Region	Country	IQ	Immigrant % in 2006	CPS Code
Europe	Albania	90		
	Andorra	98		
	Australia	98	0.16	501
	Azores**	95	0.04	130
	Austria	100	0.17	102
	Belarus	97		
	Belgium	99	0.06	103
	Bosnia and Herzegovina	90		
	Bulgaria	93		
	Canada	99	1.85	301
	Croatia	90		
	*Czechoslovakia**	97	0.11	105
	Czech Republic	98	0.07	155
	Denmark	98	0.12	106
	*Europe**	96.59	0.34	148
	Estonia	99		
	Finland	99	0.06	108
	France	98	0.32	109
	Germany	99	1.67	110
	Greece	92	0.37	116
	Hungary	98	0.25	117
	Iceland	101		
	Ireland	92	0.35	119
	Italy	102	1.15	120
	Latvia	98	0.02	183
	Lithuania	91	0.10	184
	Luxembourg	100		
	Macedonia	91		
	Malta	97		
	Moldova	96		
	Netherlands	100	0.31	126
	New Zealand	99	0.04	514
	Norway	100	0.09	127
	Poland	99	0.99	128
	Portugal	95	0.48	129
	Romania	94	0.28	132
	Russia	97	1.25	192
	Serbia	89		
	Slovakia	96	0.07	156
	Slovenia	96		
	Spain	98	0.19	134
	Sweden	99	0.08	136
	Switzerland	101	0.13	137
	Ukraine	97	0.61	195
	*USSR**	97	0.41	180
	United Kingdom	100	1.47	138-140, 142
	*Yugoslavia**	91.2	0.43	147

East Asia	Hong Kong	108	0.54	209
	Japan	105	0.85	215
	Mongolia	101		
	North Korea	106	0.00	217
	China	105	3.89	207
	Taiwan	105	0.83	238
	South Korea	106	2.51	218
Southeast Asia	Brunei	91		
	Cambodia	91	0.44	206
	East Timor	87		
	Indonesia	87	0.23	211
	Laos	89	0.28	221
	Malaysia	92	0.12	224
	Philippines	86	4.43	231
	Singapore	108	0.10	234
	Thailand	91	0.59	239
	Vietnam	94	2.46	242
Southcentral Asia	Afghanistan	84	0.23	200
	Bangladesh	82	0.40	202
	Bhutan	80		
	Burma/Myanmar	87	0.16	205
	India	82	4.06	210
	Iran	84	1.15	212
	Maldives	81		
	Nepal	78		
	Pakistan	84	0.33	229
	Sri Lanka	79		

Western Asia	Armenia	94	0.20	185
	Azerbaijan	87		
	Bahrain	83		
	Cyprus	91		
	Georgia	94		
	Iraq	87	0.30	213
	Israel	95	0.23	214
	Jordan	84	0.20	216
	Kazakhstan	94		
	Kuwait	86		
	Kyrgyzstan	90		
	Lebanon	82	0.35	222
	Oman	83		
	Palestine**	84	0.07	253
	Qatar	78		
	Saudi Arabia	84	0.17	233
	Syria	83	0.15	237
	Tajikistan	87		
	Turkey	90	0.23	240
	Turkmenistan	87		
	United Arab Emirates	84		
	Uzbekistan	87		
	Yemen	85		
North Africa	Algeria	83		
	Egypt	81	0.38	415
	Libya	83		
	Morocco	84	0.10	436
	*North Africa**	80.83	0.17	468
	Sudan	71		
	Tunisia	83		
Pacific Islands	Cook Islands	89		
	Federated States of Micronesia	84		
	Fiji	85	0.06	507
	Kiribati	85		
	Marshall Islands	84		
	New Caledonia	85		
	*Pacific Islands**	85.18	0.18	527
	Papua New Guinea	83		
	Samoa (Western)	88		
	Solomon Islands	84		
	Tonga	86		
	Vanuatu	84		

Sub-Saharan Africa	Angola	68		
	Benin	70		
	Botswana	70		
	Burkina Faso	68		
	Burundi	69		
	Cameroon	64		
	Cape Verde	76		
	Central African Republic	64		
	Chad	68		
	Comoros	77		
	Democratic Republic of the Congo	64		
	Djibouti	68		
	Equatorial Guinea	59		
	Eritrea	68		
	Ethiopia	64	0.24	417
	Gabon	64		
	Gambia	66		
	Ghana	71	0.35	421
	Guinea	67		
	Guinea-Bissau	67		
	Ivory Coast	69		
	Kenya	72	0.21	427
	Lesotho	67		
	Liberia	67		
	Madagascar	82		
	Malawi	69		
	Mali	69		
	Mauritania	76		
	Mauritius	89		
	Mozambique	64		
	Namibia	70		
	Niger	69		
	Nigeria	69	0.42	440
	Republic of the Congo	65		
	Rwanda	70		
	São Tomé and Príncipe	67		
	Senegal	66		
	Seychelles	86		
	Sierra Leone	64		
	Somalia	68		
	South Africa	72	0.32	449
	Swaziland	68		
	Tanzania	72		
	Togo	70		
	Uganda	73		
	Zambia	71		
	Zimbabwe	66		

Mexico	Mexico	88	30.56	315
Central America / Caribbean	Bahamas	84	0.08	333
	*Caribbean**	75.14	0.18	353
	*Central America**	82.57	0.64	318
	Antigua and Barbuda	70		
	Barbados	80	0.21	334
	Belize	84	0.21	310
	Bermuda	90	0.00	300
	Costa Rica	89	0.25	311
	Cuba	85	2.75	337
	Dominica	67	0.05	338
	Dominican Republic	82	2.27	339
	El Salvador	80	3.06	312
	Grenada	71	0.13	340
	Guatemala	79	1.57	313
	Haiti	67	1.13	342
	Honduras	81	1.38	314
	Jamaica	71	1.62	343
	Nicaragua	81	0.49	316
	Panama	84	0.26	317
	Saint Kitts and Nevis	67		
	Saint Lucia	62		
	Saint Vincent and the Grenadines	71		
	Trinidad and Tobago	85	0.47	351
South America	Argentina	93	0.39	375
	Bolivia	87	0.20	376
	Brazil	87	0.83	377
	Chile	90	0.25	378
	Colombia	84	1.76	379
	Ecuador	88	1.06	380
	Guyana	87	0.58	383
	Paraguay	84		
	Peru	85	0.99	385
	*South America**	87.83	0.16	389
	Suriname	89		
	Uruguay	96	0.13	387
	Venezuela	84	0.38	388
Dropped Due To Ambiguity	North America		0.10	304
	Asia		0.50	245
	Middle East		0.12	252
	Other Africa		0.91	462
	Elsewhere		2.36	555

* These are regions that are used when an immigrant's actual country of birth is unknown. Regional IQ scores are calculated as follows:

Czechoslovakia = average of Czech Republic and Slovakia
Europe = average of countries of Europe (regions, territories, Canada, Australia, and New Zealand excluded)
USSR = Russia
Yugoslavia = average of Bosnia and Herzegovina, Croatia, Macedonia, Serbia, Slovenia
North Africa = average of countries of North Africa
Central America = average of countries of Central America
South America = average of countries of South America
Caribbean = average of countries of Caribbean
Pacific Islands = average of countries of the Pacific Islands

** These territories are listed in the CPS but have no IQ scores from LV. They are imputed as follows:

Azores = Portugal
Palestine = Jordan

Appendix B: DETAILS OF IQ CALCULATIONS

The ASVAB section of chapter 2 tested Spearman's hypothesis using the method of correlated vectors (MCV). The technical details of MCV are discussed in Jensen (1998), where all the individual page citations in this section refer.

The formula for the congruence coefficient is $\sum XY / \sqrt{\sum X^2 \sum Y^2}$ (99n8).

The *g*-loadings used to calculate the correlations are an average of the loadings for white natives and the immigrant group being compared. The formula for the average is $\sqrt{(a^2 + b^2)/2}$, where *a* is the vector of *g*-loadings for natives and *b* is the vector for the immigrant comparison group (406).

Both the *g*-loadings and the group differences are adjusted by dividing by the square root of the subtest reliabilities, given in Bock and Moore (1986, 197), to correct for attenuation. The only paper to perform a similar MCV analysis with the ASVAB is Hartmann et al. (2007), which tested Spearman's hypothesis on the white-Hispanic difference, without considering immigrant generation at all. The result was that the correlation in question, although initially quite high, was reduced to insignificance when the reliabilities were accounted for. The authors reach this result probably because they do not use the actual reliabilities; rather, they use the communalities, which are a lower bound on the reliabilities. Unaware of Bock and Moore (1986), they say the reliabilities are unavailable.

I used the DIFPACK software, version 1.7, to implement SIBTEST on the PIAT-R Math in chapter 2. DIFPACK is produced by the Roussos-Stout Software Development Group. It is available for purchase at: http://www.assess.com/xcart/product.php?productid=224. This version of the software includes the Jiang and Stout (1998) regression correction to better control Type I error.

SIBTEST was run using a minimum cell size of 2, but higher minimums made little difference in the results. The one-tailed p-value was 0.5.

Respondents do not answer every item on the PIAT-R. Instead, they answer items that come between a basal (lowest item answered correctly) and a ceiling (highest item answered correctly). The basal and ceiling are determined dynamically by how well the respondent performs. All items coming before the basal are assumed to be correct, and all items after the ceiling are assumed incorrect. This procedure may have indirectly reduced the bias of the overall test, since a biased early or late item would not often be encountered by the respondents.

I performed two other internal validity tests that corroborate the SIBTEST results, but I did not include them in the text because they may have methodological problems. The first was the item rank-order correlation between natives and immigrants, which was over 0.99, indicating no bias. According to Wicherts (2007, 134), this method is antiquated. The second is the Mantel-Haenszel procedure, which identified a handful of biased items that, as with SIBTEST, had little impact on the overall scores. According to Roussos et al. (1999), Mantel-Haenszel can produce misleading results in certain cases.

On the digit span tests, older norms were used, which suggests a problem with the Flynn effect. Due to the Flynn effect, which is discussed in chapter 1, a 2003 sample given a full-scale IQ test normed to 100 in 1991 may be expected to show a mean of 103 (Flynn 1998). Since they were compared against norms that are too low for today's standards, the *d* of 0.16 for immigrants may actually be too small in magnitude, by about 3/15 = 0.2 standard deviations.

However, IQ inflation varies considerably on subtests. In the case of the digit span, the degree of score inflation appears to be small relative to full-scale gains. One paper (Wicherts et. al. 2004) found large Flynn effects between 1968 and 1999 on each subtest of the adult version

of the Wechsler. Digit span increased by about half a standard deviation over 31 years, right in line with Flynn's estimate of 0.25 IQ points per year, but this was actually the smallest increase of any test in the battery. Since participants in the Wicherts et al. study had taken another version of the Wechsler less than three months prior, a retest effect probably caused overestimation of the Flynn effect on each subtest.

Two other studies (Rodgers and Wanstrom 2006; Murray 2006) found no Flynn effect at all on the digit span given to the children of NLSY participants. Since the data are not clear on the subject, and any actual Flynn effect on the digit span appears to be small, I do not make any Flynn adjustment in the text. Therefore, the native-immigrant *d* of 0.16 is, if anything, biased in favor of immigrants rather than against them.

Somewhat confusingly, the age variable provided by the NIS is the child's age when first sampled for the survey. The actual digit span test was conducted up to a year after the original sampling. To calculate each child's true age at the time of the test, I subtracted birth year and month from the year and month that the test was administered. The children's birth years and months could be found only in the adult sample, where each adult had information about his or her children.

In calculating the digit span *d*, I was careful to exclude the children of immigrants from the NIS who were born in the United States, as they are not technically immigrants at all. There was also an issue of test conditions. From the tester comments appended to some of the children's digit span scores, one can see they were not ideal. Parents and siblings were often in the room when the test was being conducted. If the tester reported that the child was at all distracted during administration, the child's case was dropped from the analysis. (If the variable ds1a2=2 or was missing, then the child was considered distracted.)

Appendix C: LIST OF COUNTRIES BY 1970 EDUCATION LEVEL

Country	IQ[1]	1970 Education[2]	CPS code	1970 Census Code
Afghanistan	**84**		200	
Argentina	**93**	150.3	375	30005
Armenia	**94**		185	
Australia	**98**	158.4	501	70010
Austria	**100**	143.2	102	45000
Bangladesh	**82**		202	
Barbados	**80**		334	
Belgium	**99**	150.6	103	42000
Belize	**84**	136.0	310	21010
Bermuda	**90**	127.3	300	16000
Bolivia	**87**	159.7	376	30010
Brazil	**87**	148.9	377	30015
Cambodia	**91**		206	
Canada	**99**	143.4	301	15000
Chile	**90**	155.9	378	30020
Colombia	**84**	136.0	379	30025
Costa Rica	**89**	132.9	311	21020
Cuba	**85**	132.7	337	25000
Czech Republic	**98**	138.2	155	45200
Denmark	**98**	147.9	106	40000
Dominican Republic	**82**	113.8	339	26010
Ecuador	**88**	135.9	380	30030
Egypt	**81**	167.9	415	60012
El Salvador	**80**	134.6	312	21030
Ethiopia	**64**		417	
Fiji	**85**		507	
Finland	**99**	138.2	108	40100
France	**98**	152.4	109	42100
Germany	**99**	145.5	110	45300
Ghana	**71**		421	
Greece	**92**	120.3	116	43300
Grenada	**71**		340	
Guatemala	**79**	137.2	313	21040
Guyana	**87**		383	
Haiti	**67**	143.1	342	26020
Honduras	**81**	131.9	314	21050
Hong Kong	**108**		209	
Hungary	**98**	138.7	117	45400
India	**82**	184.8	210	52100
Indonesia	**87**		211	
Iran	**84**	163.8	212	52200
Iraq	**87**		213	
Ireland	**92**	133.1	119	41400
Israel	**95**	156.6	214	53400
Italy	**102**	109.7	120	43400
Jamaica	**71**	137.9	343	26030
Japan	**105**	151.4	215	50100
Jordan	**84**	131.7	216	53500
Kenya	**72**		427	
Laos	**89**		221	
Latvia	**98**	156.7	183	46100
Lebanon	**82**	145.1	222	53700
Lithuania	**91**	139.4	184	46200

Malaysia	**92**		224	
Mexico	**88**	93.8	315	20000
Morocco	**84**		436	
Myanmar	**87**		205	
Netherlands	**100**	147.5	126	42500
New Zealand	**99**	165.1	514	70020
Nicaragua	**81**	130.8	316	21060
Nigeria	**69**		440	
North Korea	**106**	160.6	217-218	50200
Norway	**100**	140.2	127	40400
Pakistan	**84**	168.7	229	52140
Panama	**84**	146.7	317	21070
People's Republic of China	**105**	138.2	207	50000
Peru	**85**	150.5	385	30050
Philippines	**86**	147.4	231	51500
Poland	**99**	125.5	128	45500
Portugal	**95**	87.4	129	43600
Puerto Rico	**84**		72	
Republic of China	**105**		238	
Romania	**94**	133.2	132	45600
Russia	**97**		192	
Saudi Arabia	**84**		233	
Singapore	**108**		234	
Slovakia	**96**	138.2	156	45200
South Africa	**72**	164.3	449	60094
South Korea	**106**	160.6	217-218	50200
Spain	**98**	127.9	134	43800
Sweden	**99**	141.4	136	40500
Switzerland	**101**	155.4	137	42600
Syria	**83**	132.4	237	54100
Thailand	**91**		239	
Trinidad and Tobago	**85**	144.7	351	26060
Turkey	**90**	140.3	240	54200
Ukraine	**97**	124.4	195	46530
United Kingdom	**100**	151.3	138-140, 142	41000
Uruguay	**96**	146.1	387	30060
Venezuela	**84**	154.4	388	30065
Vietnam	**94**	150.4	242	51800

Table Notes

[1] Chapter 6 was written a year before the rest of the dissertation, so the national IQ scores used in it do not include some of the minor revisions used in chapter 2 and shown in Appendix A.

[2] These are raw education scores averaged directly from the 1970 census codes. A score of 80 corresponds to completion of 5th grade, and then an increment of 10 on the raw score corresponds to one additional grade level: 90 = 6th grade, 100 = 7th grade, …, 150 = 12th grade, …, 190 = "16th grade" or college completion.

REFERENCES

Abedi, Jamal and Carol Lord. 2001. The language factor in mathematics tests. *Applied Measurement in Education* 14: 219-234.

Alba, Richard, Amy Lutz, and Elena Vesselinov. 2001. How enduring were the inequalities among European immigrant groups in the United States? *Demography* 38: 349-356.

Andreasen, Nancy C., Michael Flaum, Victor Swayze II, Daniel S. O'Leary, Randall Alliger, Gregg Cohen, James Ehrhardt, and William T.C. Yuh. 1993. Intelligence and brain structure in normal individuals. *American Journal of Psychiatry* 150: 130-134.

Barber, Nigel. 2005. Educational and ecological correlates of IQ: A cross-national investigation. *Intelligence* 33: 273-284.

Barnett, Susan M. and Wendy Williams. 2004. National intelligence and the emperor's new clothes: A review of *IQ and the wealth of nations*. *Contemporary Psychology* 49: 389-396.

Bayley, Nancy. 1949. Consistency and variability in the growth of intelligence from birth to eighteen years. *Journal of Genetic Psychology* 75: 165-196.

Bock, R. Darrell and Elsie G.J. Moore. 1986. Advantage and disadvantage: A profile of American youth. Hillsdale, NJ: Lawrence Erlbaum Associates.

Borjas, George J. 1987. Self-selection and the earnings of immigrants. *American Economic Review* 77: 531-553.

------. 1994a. The economics of immigration. *Journal of Economic Literature* 32: 1667-1717.

------. 1994b. Long-run convergence of ethnic skill differentials: The children and grandchildren of the great migration. *Industrial and Labor Relations Review* 47: 553-573.

------. 1995. The economic benefits from immigration. *Journal of Economic Perspectives* 9: 3-22.

------. 1998. Do blacks gain or lose from immigration? *Help or hindrance: The economic implications of immigration for African Americans*. Ed. Daniel S. Hamermesh and Frank D. Bean. New York: Russell Sage Foundation. 51-74.

------. 1999. Heaven's door. Princeton: Princeton University Press.

------. 2001. Long-run Convergence of ethnic skill differentials, revisited. *Demography* 38: 357-361.

------. 2003. The labor demand curve *is* downward sloping: Reexamining the impact of immigration on the labor market. *Quarterly Journal of Economics* 113: 1335-1374.

Bowles, Samuel, Herbert Gintis, and Melissa Osborne. 2001. The determinants of earnings: A behavioral approach. *Journal of Economic Literature* 39: 1137-1176.

Brand, Chris. 1987a. The importance of general intelligence. *Arthur Jensen: Consensus and Controversy*. Ed. Sohan Mogdil and Celia Mogdil. New York: Routledge**.**

------. 1987b. Bryter still and bryter? *Nature* 328: 110.

Brigham, Carl C. 1923. A study of American intelligence. Princeton: Princeton University Press.

------. 1930. Intelligence tests of immigrant groups. *Pyschological Review* 37: 158-165.

Brooks-Gunn, Jeanne, Cecilia M. McCarton, Patrick H. Casey, Marie C. McCormick, Charles R. Bauer, Judy C. Bernbaum, Jon Tyson, Mark Swanson, Forrest C. Bennett, David T. Scott, James Tonascia, and Curtis L. Meinert. 1994. Early intervention in low-birth-weight premature infants: Results through age 5 years from the Infant Health and Development Program. *Journal of the American Medical Association* 272: 1257-1262.

Butcher, Kristin F. and Anne Morrison Piehl. 2008. Crime, corrections, and California: What does immigration have to do with it? *California Counts* 9: 1-22.

Campbell, Francis A., Craig T. Ramey, Elizabeth Pungello, Joseph Sparling, and Shari Miller-Johnson. 2002. Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science* 6: 42-57.

Carens, Joseph H. 1987. Aliens and citizens: The case for open borders. *The Review of Politics* 49: 251-273.

Carroll, John B. 1993. Human cognitive abilities: A survey of factor-analytic studies. New York: Cambridge University Press.

------. 1995. Reflections on Stephen Jay Gould's *The mismeasure of man* (1981): A retrospective review. *Intelligence* 21: 121-134.

Chiquiar, Daniel and Gordon H. Hanson. 2005. International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States. *Journal of Political Economy* 105: 239-281.

Colom, Roberto, Rex E. Jung, and Richard J. Haier. 2006. Distributed brain sites for the *g*-factor of intelligence. *NeuroImage* 31: 1359-1365.

da Costa Pinto, Amancio. 1991. Reading rates and digit span in bilinguals: The superiority of mother tongue. *International Journal of Psychology* 26: 471-483.

Daniels, Michael, Bernie Devlin, and Katherine Roeder. 1997. Of genes and IQ. *Intelligence, genes, and success: Scientists respond to the bell curve*. Ed. Bernie Devlin, Stephen E. Fienberg, Daniel P. Resnick, and Kathryn Roeder. New York: Springer-verlag. 45-70.

Davis, Bernard. 1983. Neo-Lysenkoism, IQ, and the press. *The Public Interest* 74: 41-59.

de Mel, Suresh, David McKenzie, and Christopher Woodruff. 2007. Returns to capital in microenterprises: Evidence from a field experiment. World Bank Policy Research Working Paper 4230.

de Wit, Harriet, Janine D. Flory, Ashley Acheson, Michael McCloskey, and Stephen B. Manuck. 2007. IQ and non-planning impulsivity are independently associated with delay discounting in middle-aged adults. *Personality and Individual Differences* 42: 111-121.

Dickens, William T. and James R. Flynn. 2001. Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review* 108: 346-369.

------. 2006. Black Americans reduce the racial IQ gap: Evidence from standardization samples. *Psychological Science* 17: 913-920.

Dickerson, Richard E. 2006. Exponential correlation of IQ and the wealth of nations. *Intelligence* 34: 291-295.

Djankov, Simeon, Gérard Roland, Edward Miguel, Ekaterina Zhuravskaya, and Yingyi Qian. 2005. Who are Russia's entrepreneurs? *Journal of the European Economic Association* 3: 1-11.

Donoghue, John R. and Nancy L. Allen. 1993. Thin versus thick matching in the Mantel-Haenszel procedure for detecting DIF. *Journal of Educational Statistics* 18: 131-154.

Douglas, Jeffrey A., Louis A. Roussos, and William Stout. 1996. Item-bundle DIF hypothesis testing: Identifying suspect bundles and assessing their differential functioning. *Journal of Educational Measurement* 33: 465-484.

Duncan, Greg J., Jens Ludwig, and Katherine A. Magnuson. 2007. Reducing poverty through pre-school interventions. *The Future of Children* 17: 143-160.

Ellis, Nick C. and R.A. Hennelly. 1980. A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology* 71: 43-51.

Entine, Jon. 2007. Abraham's children: Race, identity, and the DNA of the chosen people. New York: Grand Central Publishing.

Feldman, Marcus W., Sarah P. Otto, and Freddy B. Christiansen. 2000. Genes, culture, and inequality. *Meritocracy and Economic Inequality*. Ed. Kenneth Arrow, Samuel Bowles, and Steven Durlauf. Princeton: Princeton University Press.

Ferguson, Ronald F. 2001. A diagnostic analysis of Black-White GPA disparities in Shaker Heights, Ohio. *Brookings papers on education policy: 2001*. Ed. Diane Ravitch. Washington: Brookings Institution Press.

Fischer, Claude S., Michael Hout, Martin Sanchez Jankowski, Samuel R. Lucas, Ann Swidler, and Kim Voss. 1996. Inequality by design. Princeton: Princeton University Press.

Flynn, James R. 1984. The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin* 95: 29-51.

------. 1987. Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin* 101: 171-191.

------. 1998. WAIS-III and WISC-III: IQ gains in the United States from 1972 to 1995; how to compensate for obsolete norms. *Perceptual and Motor Skills* 86: 1231-1239.

------. in press. Requiem for nutrition as the cause of IQ gains: Raven's gains in Britain 1938 to 2008. *Economics and Human Biology*.

Frey, Meredith C. and Douglas K. Detterman. 2004. Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science* 15: 373-378.

Gardner, Howard. 1983. Frames of mind: The theory of multiple intelligences. New York: Basic Books.

------. 2006. Replies to my critics. *Howard Gardner under fire: The rebel psychologist faces his critics*. Ed. Jeffrey A. Schaler. Chicago: Open Court**.** 277-344.

Gordon, Robert A. 1997. Everyday life as an intelligence test: Effects of intelligence and intelligence context. *Intelligence* 24: 203-320.

Gottfredson, Linda S. 1986. Occupational aptitude patterns map: Development and implications for a theory of job aptitude requirements. *Journal of Vocational Behavior* 29: 254-291.

------. 1988. Reconsidering fairness: A matter of social and ethical priorities. *Journal of Vocational Behavior* 33: 293-319.

------. 1994. Mainstream science on intelligence. *Wall Street Journal*, December 13.

------. 1997. Why g matters: The complexity of everyday life. *Intelligence* 24: 79-132.

------. 2005. What if the hereditarian hypothesis is true? *Psychology, Public Policy, and Law* 11: 311-319.

Gould, Stephen Jay. 1981. The mismeasure of man. New York: W.W. Norton.

Grant, Madison. 1916. The passing of the great race. New York: C. Scribner.

Haier, Richard J., Rex E. Jung, Ronald A. Yeo, Kevin Head, and Michael T. Alkire. 2004. Structural brain variation and general intelligence. *NeuroImage* 23: 425-433.

Hamermesh, Daniel S. 1993. Labor demand. Princeton: Princeton University Press.

Hartmann, Peter, Nanna Hye Sun Kruusea, and Helmuth Nyborg. 2007. Testing the cross-racial generality of Spearman's hypothesis in two samples. *Intelligence* 35: 47-57.

Heckman, James J. 1995. Lessons from the bell curve. *Journal of Political Economy* 103: 239-281.

Herrnstein, Richard J. and Charles Murray. 1994. The bell curve: Intelligence and class structure in American life. New York: Free Press Paperback.

Hoosain, Rumjahn. 1979. Forward and backward digit span in the languages of the bilingual. *The Journal of Genetic Psychology* 135: 263-268.

Hoppe, Hans-Hermann. 2001. Democracy—The god that failed: The economics and politics of monarchy, democracy, and natural order. New Brunswick: Transaction Publishers.

Huntington, Samuel P. 2004. Who are we? The challenges to America's national identity. New York: Simon and Schuster.

Hymowitz, Kay S. 2006. Marriage and caste in America : Separate and unequal families in a post-marital age. Chicago: Ivan R. Dee.

Jencks, Christopher. 1980. Heredity, environment, and public policy reconsidered. *American Sociological Review* 45: 723-736.

------. 1992. Rethinking social policy. New York: Harper Collins.

------. 2001. Who should get in? *The New York Review of Books* 48: 57-66.

Jensen, Arthur R. 1980. Bias in mental testing. New York: The Free Press.

------. 1985. The nature of the black-white difference on various psychometric tests: Spearman's hypothesis. *Behavioral and Brain Science* 8: 193-219.

------. 1998. The g factor: The science of mental ability. Westport: Praeger.

Jiang, Hai and William Stout. 1998. Improved Type I error control and reduced estimation bias for DIF detection using SIBTEST. *Journal of Educational and Behavioral Statistics* 23: 291-322.

Jones, Garett. 2008. Are smarter groups more cooperative? Evidence from prisoner's dilemma experiments, 1959–2003. *Journal of Economic Behavior and Organization* 68: 489-497.

Jones, Garret and W. Joel Schneider. 2006. Intelligence, human capital, and economic growth: A Bayesian averaging of classical estimates (BACE) approach. *Journal of Economic Growth* 11: 71-93.

Kamin, Leon. 1974. The science and politics of IQ. Potomac, MD: Lawrence Erlbaum Associates.

Kaufman, Alan S. 1979. Intelligence testing with the WISC-R. New York: Wiley and Sons.

Kirsch, Irwin S., Ann Jungeblut, Lynn Jenkins, and Andrew Kolstad. 1993. Adult literacy in America: A first look at the findings of the National Adult Literacy Survey. Princeton: Educational Testing Service.

Korenman, Sanders and Christopher Winship. 2000. A reanalysis of *The bell curve*: Intelligence, Family Background, and Schooling. *Meritocracy and Economic Inequality*. Ed. Kenneth Arrow, Samuel Bowles, and Steven Durlauf. Princeton: Princeton University Press. 137-178.

Krikorian, Mark. 2008. The new case against immigration: Both legal and illegal. New York: Sentinel.

Larsen, Lars, Peter Hartmann, and Helmuth Nyborg. 2008. The stability of general intelligence from early adulthood to middle-age. *Intelligence* 36: 29-34.

Lemann, Nicholas. 1999. The big test: The secret history of American meritocracy. New York: Farrar, Strauss, and Giroux.

Levin, Michael. 1994. Comment on the Minnesota transracial adoption study. *Intelligence* 19: 13-20.

Loehlin, John C., Stephen G. Vandenberg, and R. Travis Osborne. 1973. Blood-group genes and Negro-White ability differences. *Behavior Genetics* 3: 263-270.

Loehlin, John C. 2002. The IQ paradox: Resolved? Still an open question. *Psychological Review* 109: 754-758.

Lynch, James P. and Rita J. Simon. 2003. Immigration the world over: Statutes, policies, and practices. New York: Rowman and Littlefield.

Lynn, Richard. 1990. The role of nutrition in secular increases in intelligence. *Personality and Individual Differences* 11: 273-285.

------. 2008. The global bell curve. Augusta, Georgia: Washington Summit Publishers.

Lynn, Richard and Tatu Vanhanen. 2002. IQ and the wealth of nations. Westport: Praeger.

------. 2006. IQ and global inequality. Augusta: Washington Summit Publishers.

Mahler, Sarah J. 1995. American dreaming. Princeton: Princeton University Press.

Malanga, Steven. 2007. The right immigration policy. *The immigration solution*. Ed. Heather MacDonald, Victor Davis Hanson, and Steven Malanga. Chicago: Ivan R. Dee. 170-184.

Markwardt, Frederick C. 1998. PIAT-R manual: Normative update. Circle Pines: American Guidance Service, Inc.

Mas-Colell, Andreu, Michael D. Winston, and Jerry R. Green. 1995. Microeconomic theory. New York: Oxford University Press.

Millet, Kobe and Siegfried Dewitte. 2006. Altruistic behavior as a costly signal of general intelligence. *Journal of Research in Personality* 41: 316-326.

Munford, Paul R. and Alejandra Munoz. 1980. A comparison of the WISC and WISC-R on Hispanic children. *Journal of Clinical Psychology* 36: 452-457.

Murray, Charles. 1984. Losing ground: American social policy, 1950-1980. New York: Basic Books.

------. 1995. The bell curve and its critics. *Commentary*, May: 23-30.

------. 1999. The underclass revisited. Washington, DC: AEI Press.

------. 2005. The inequality taboo. *Commentary*, September: 13-22.

------. 2006. Changes over time in the black–white difference on mental tests: Evidence from the children of the 1979 cohort of the National Longitudinal Survey of Youth. *Intelligence* 34: 527-540.

------. 2007a. Jewish genius. *Commentary*, April: 29-35.

------. 2007b. The magnitude and components of change in the black–white IQ difference from 1920 to 1991: A birth cohort analysis of the Woodcock–Johnson standardizations. *Intelligence* 35: 305-318.

------. 2008. Real education. New York: Crown Forum.

Neal, Derek. 2006. Why has black-white skill convergence stopped? *Handbook of the economics of education*. Ed. Eric A. Hanushek and Finis Welch. New York: Elsevier. 511-576.

Neal, Derek and William R. Johnson. 1996. Premarket factors in black-white wage differences. *Journal of Political Economy* 104: 869-895.

Nechyba, Thomas J. 2004. Review of *IQ and the wealth of nations*. *Journal of Economic Literature* 42: 220-221.

Neisser, Ulric, Gwyneth Boodoo, Thomas J. Bouchard, A. Wade Boykin, Nathan Brody, Stephen J. Ceci, Diane E Halpern, John C. Loehlin, Robert Perloff, Robert J. Sternberg, and Susana Urbina. 1996. Intelligence: Knowns and unknowns. *American Psychologist* 52: 77-101.

Nielsen, Francois. 1997. Happy days are here again! *Social Forces* 76: 701-704.

Nisbett, Richard E. 2005. Heredity, environment, and race differences in IQ: A commentary on Rushton and Jensen. *Psychology, Public Policy, and Law* 11: 302-310.

Nozick, Robert. 1974. Anarchy, state, and utopia. New York: Basic Books.

Ogbu, John U. 2003. Black American students in an affluent suburb: A study of academic disengagement. Mahwah, NJ: Lawrence Erlbaum Associates.

Olazaran, Javier., Diane M. Jacobs, and Yaakov Stern. 1996. Comparative study of visual and verbal short-term memory in English and Spanish speakers: Testing a linguistic hypothesis. *Journal of the International Neuropsychological Society* 2: 105-110.

Pennock-Roman, Maria. 1992. Interpreting test performance in selective admissions for Hispanic students. *Psychological testing of Hispanics*. Ed. Kurt F. Geisinger. Washington, DC: American Psychological Association**.** 99-135.

Portes, Alejandro and Min Zhou. 1993. The new second generation: Segmented assimilation and its variants. *Annals of the American Academy of Political and Social Science* 530: 74-96.

Prokosch, Mark D., Ronald A. Yeo, and Geoffrey F. Miller. 2005. Intelligence tests with higher g-loadings show higher correlations With body symmetry: Evidence for a general fitness factor mediated by developmental stability. *Intelligence* 33: 203-213.

Putnam, Robert D. 1995. Bowling alone: America's declining social capital. *Journal of Democracy* 6: 65-78.

------. 2007. E pluribus unum: Diversity and community in the twenty-first century. *Scandinavian Political Studies* 30: 137-174.

Rawls, John. 1971. A theory of justice. Cambridge: Harvard University Press.

------. 2001. The law of peoples: With the idea of public revision revisited. Cambridge: Harvard University Press.

Reeve, Charlie L. and Jennifer E. Charles. 2008. Survey of opinions on the primacy of *g* and social consequences of ability testing: A comparison of expert and non-expert views. *Intelligence* 36: 681-688.

Rice, Tom W. and Jan L. Feldman. 1997. Civic culture and democracy from Europe to America. *The Journal of Politics* 59: 1143-1172.

Rodgers, Joseph Lee and Linda Wanstrom. 2006. Identification of a Flynn effect in the NLSY: Moving from the center to the boundaries. *Intelligence* 35: 187-196.

Roth, Philip L, Craig A. Bevier, Philip Bobko, Fred S. Switzer, and Peggy Tyler. 2001. Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology* 54: 297-330.

Roussos, Louis and William Stout. 1996. A multidimensionality-based DIF analysis paradigm. *Applied Psychological Measurement* 20: 355-371.

Roussos, Louis A., Deborah L. Schnipke, and Peter J. Pashley. 1999. A generalized formula for the Mantel-Haenszel differential item functioning parameter. *Journal of Educational and Behavioral Statistics* 24: 293-322.

Rumbaut, Ruben G. and Walter A. Ewing. 2007. The myth of immigrant criminality and the paradox of assimilation: Incarceration rates among native and foreign-born men. Washington, DC: American Immigration Law Foundation.

Rushton, J. Philippe. 2000. Race, evolution, and behavior. Port Huron: Charles Darwin Research Institute.

Rushton, J. Philippe and Arthur R. Jensen. 2005a. Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law* 11: 235-294.

------. 2005b. Wanted: More race realism, less moralistic fallacy. *Psychology, Public Policy, and Law* 11: 328-336.

------. 2006. The totality of available evidence shows the race IQ gap still remains. *Psychological Science* 17: 921-922.

Rushton, J. Philippe and Richard M. Sorrentino. 1981. Altruism and helping behavior: Current perspectives and future possibilities. *Altruism and helping behavior: social, personality, and developmental perspectives*. Ed. J. Philippe Rushton and Richard M. Sorrentino. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rushton, J. Phillippe and J. Weiner. 1975. Altruism and cognitive development in children. *British Journal of Social and Clinical Psychology* 14: 341-349.

Sackett, Paul R., Chaitra M. Hardison, and Michael J. Cullen. 2004. On interpreting stereotype threat as accounting for African American–White differences on cognitive tests. *American Psychologist* 59: 7-13.

Sackett, Paul R. and Stephanie L. Wilk. 1994. Within group norming and other forms of score adjustments in preemployment testing. *American Psychologist* 49: 929-954.

Sala-I-Martin, Xavier X. 1997. I just ran two million regressions. *American Economic Review* 87: 178-183.

Sampson, Robert J, Jeffrey D. Morenoff, and Stephen Raudenbush. 2005. Social anatomy of racial and ethnic disparities in violence. *American Journal of Public Health* 95: 224-231.

Scarr, Sandra. 1997. Behavior– genetic and socialization theories of intelligence: Truce and reconciliation. *Intelligence, Heredity, and Environment*. Ed. Robert J. Sternberg and Elena. L. Grigorenko. New York: Cambridge University Press. 3–41.

Scarr, Sandra, Andrew J. Pakstis, Solomon H. Katz, and William B. Barker. 1977. Absence of a relationship between degree of White ancestry and intellectual skills within a Black population. *Human Genetics* 39: 69-86.

Schmitt, Alicia P. 1988. Language and cultural characteristics that explain differential item functioning for Hispanic examinees on the Scholastic Aptitude Test. *Journal of Educational Measurement* 25: 1-13.

Shamosh, Noah A. and Jeremy R. Gray. 2008. Delay discounting and intelligence: A meta-analysis. *Intelligence* 36: 289-305.

Shealy, Robin and William Stout. 1993. A model-based standardization approach that separates true bias/DIF from group ability differences and detects test bias/DIF as well as item bias/DIF. *Psychometrika* 58: 159-194.

Singer, Peter. 1993. Practical ethics. New York: Cambridge University Press.

Snyderman, Mark and Stanley Rothman. 1988. The IQ controversy: The media and public policy. New Brunswick: Transaction Press.

Snyderman, Mark and Richard J. Hernnstein. 1983. Intelligence tests and the immigration act of 1924. *American Psychologist* 38: 986-995.

Sowell, Thomas. 1978. Race and IQ reconsidered. *Essays and data on American ethnic groups*. Ed. Thomas Sowell. Washington, DC: Urban Institute**.**

------. 2005. Black rednecks and white liberals. San Francisco: Encounter Books.

Spitz, Herman H. 1992. Does the Carolina Abecedarian Early Intervention Project prevent sociocultural mental retardation? *Intelligence* 16: 225-237.

Steele, Claude M. and Joshua Aronson. 1995. Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology* 69: 797-811.

Sternberg, Robert J. 1985. Beyond IQ: A triarchic theory of human intelligence. New York: Cambridge University Press.

------. 1996. Myths, countermyths, and truths about intelligence. *Educational Researcher* 25: 11-16.

Stigler, James W., Shin-Ying Lee, and Harold W. Stevenson. 1986. Digit memory in Chinese and English: Evidence for a temporally limited store. *Cognition* 23: 1-20.

Strenze, Tarmo. 2007. Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence* 35: 401-426.

Taylor, Jared. 1992. Paved with good intentions: The failure of race relations in contemporary America. New York: Carroll and Graf.

te Nijenhuis, Jan and Henk van der Flier. 1999. Bias research in the Netherlands: Review and implications. *European Journal of Psychological Assessment* 15: 165-175.

------. 2003. Immigrant-majority group differences in cognitive performance: Jensen effects, cultural effects, or both? *Intelligence* 31: 443-459.

te Nijenhuis, Jan, Mart-jan de Jong, Arne Evers, and Henk Van Der Flier. 2004. Are cognitive differences between immigrant and majority groups diminishing? *European Journal of Personality* 18: 405-434.

Teasdale, Thomas W. and David R. Owen. 2008. Secular declines in cognitive test scores: A reversal of the Flynn Effect. *Intelligence* 36: 121-126.

Telles, Edward E. and Vilma Ortiz. 2008. Generations of exclusion: Mexican Americans, assimilation, and race. New York: Russell Sage Foundation.

Thorndike, Robert M. and David F. Lohman. 1990. A century of ability testing. Chicago: Riverside Publishers.

Valencia, Richard R. and Richard J. Rankin. 1985. Evidence of content bias on the McCarthy Scales with Mexican-American children: Implications for test translation and non-biased assessment. *Journal of Educational Psychology* 77: 197-207.

Volken, Thomas. 2003. IQ and the wealth of nations: A critique of Richard Lynn and Tatu Vanhanen's recent book. *European Sociological Review* 19: 411–412.

Voracek, Martin. 2004. National intelligence and suicide rate: An ecological study of 85 countries. *Personality and Individual Differences* 37: 543-553.

Walzer, Michael. 1983. Spheres of justice: A defense of pluralism and equality. New York: Basic Books.

Wechsler, David. 1991. WISC-III manual. New York: The Psychological Corporation.

Weede, Erich and Sebastian Kampf. 2002. The impact of intelligence and institutional improvements on economic growth. *Kyklos* 55: 361–380.

Weinberg, Richard A., Sandra Scarr, and Irwin D. Waldman. 1992. The Minnesota transracial adoption study: A follow-up of IQ test performance at adolescence. *Intelligence* 16: 117-135.

Weissbrodt, David and Laura Danielson. 2005. Immigration law and procedure in a nutshell. St. Paul, MN: West Publishing Company.

Whetzel, Deborah L. and Michael A. McDaniel. 2006. Prediction of national wealth. *Intelligence* 34: 449-458.

Whitworth, Randolph H. and Sabine M. Chrisman. 1987. Validation of the Kaufman Assessment Battery for children comparing Anglo and Mexican-American preschoolers. *Educational and Psychological Measurement* 47: 695-702.

Wicherts, Jelte M. 2007. Group differences in intelligence test performance. Amsterdam: Printpartners Ipskamp.

Wicherts, Jelte M., Conor V. Dolan, David J. Hessen, Paul Oosterveld, G. Caroline M. van Baal, Dorret I. Boomsma, and Mark M. Span. 2004. Are intelligence tests measurement invariant over time? Investigating the nature of the Flynn effect. *Intelligence* 32: 509-537.

Wierzbicki, Susan. 2004. Beyond the immigrant enclave: Network change and assimilation. El Paso, TX: LFB Scholarly Publishing.

Wilson, James Q. and Richard Herrnstein. 1985. Crime and human nature. New York: Simon and Schuster.

Wilson, William Julius. 1987. The truly disadvantaged: The inner city, the underclass, and public policy. Chicago: University of Chicago Press.

------. 1996. When work disappears. New York: Alfred A. Knoph.