Human Intelligence

EARL HUNT
University of Washington

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo, Mexico City

Cambridge University Press
32 Avenue of the Americas, New York, NY 10013-2473, USA

www.cambridge.org
Information on this title: www.cambridge.org/9780521707817

First published 2011

Printed in the United States of America

A catalog record for this publication is available from the British Library.

Library of Congress Cataloging in Publication data

Hunt, Earl B.
Human intelligence / Earl Hunt.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-521-88162-3 (hardback) – ISBN 978-0-521-70781-7 (pbk.)
1. Intellect. I. Title.
BF431.H814 2010
153.9–dc22 2010038696

ISBN 978-0-521-88162-3 Hardback
ISBN 978-0-521-70781-7 Paperback

CHAPTER 1

The Issue of Intelligence

So it is that gods do not give all men the gifts of grace . . . neither good looks nor intelligence nor eloquence.

Homer, *The Odyssey*

There's many a man has more hair than wit.

Shakespeare, *Comedy of Errors*, act 2, scene 3

1.1. The Idea of Intelligence

Homer and Shakespeare lived in very different times, more than two thousand years apart, but they both captured the same idea; we are not all equally intelligent. I suspect that anyone who has failed to notice this is somewhat out of touch with the species. However, we cannot simply sort people into the "intelligent" and the "not-so-intelligent." Homer observed that few people have great gifts. Shakespeare, more pithily, observed that all too many of us do not do terribly well at problem solving. Most of us, though, fall in between Homer's desire for eloquence and Shakespeare's worry about lack of wit.

In this book I will talk about the nature of intelligence, its causes, who has it, and how it is used. I will do so without the eloquence of Homer and Shakespeare. I will take a scientific view. Modern psychology has a great deal to say about intelligence, and somehow a great deal that has been said has been seriously misunderstood. The popular media sometimes report that the psychologists who study intelligence say almost the opposite of what the psychologists actually said.[1]

There is a reason for this. The study of intelligence is not an isolated academic topic; our intelligence has social consequences. We want our leaders to be intelligent, and exhibit concern if we think they are not. There were politically motivated attacks on the intelligence of Presidents Lincoln, Truman, Harding, and Ford. Serious concerns about mental competence

1 See Tannenbaum (1996) for a discussion of this issue and references to earlier discussions of the topic. Gottfredson (2005) provides a spirited discussion of how failing to consider the implications of psychological research on intelligence can be costly to society.

were raised about Wilson and Eisenhower, following strokes, and Reagan, due to early symptoms of Alzheimer's disease. Lincoln and Truman, who received the most vicious attacks, are now considered two of our finest presidents. Eisenhower recovered to function well; Wilson did not. The effect of Reagan's illness upon his second term is still a matter of debate.

Concerns about intelligence are not confined to concerns about our leaders. Our school systems use cognitive tests to stream high school students into different programs. Colleges use cognitive tests to screen applicants for higher education programs. These tests are never called "intelligence tests," but they correlate highly with them.

Testing is not confined to the educational system. Volunteers for the United States military services must obtain passing scores on a test of general mental competence, the Armed Forces Qualifying Test (AFQT). Similar tests are used in many other countries. Toward the bottom end of the scale there are a variety of special assistance programs for people who simply do not have the cognitive competence to cope with the complexities of the modern world. Low intelligence test scores can be offered as evidence of diminished mental capacity during the penalty phase of a criminal trial.

While there is broad agreement that some people are smarter than others, things become more complex when we try to be precise. I think that every knowledgeable person would agree that Albert Einstein and Thomas Jefferson were both highly intelligent. Who was the more intelligent? That is hard to say; they were brilliant in different ways at different times. It would be easy to find other examples of the same point. There are clearly varieties of cognitive skill, especially at the top. As a result some modern observers have concluded that there is no single dimension of intelligence.

This idea is not new. In the sixteenth century the Spanish physician/philosopher Juan Huarte de San Juan[2] drew a remarkably cogent picture of individual differences in human thought. Huarte believed that when people attack problems some will use their imaginations to envisage how a solution might work out, while others will rely on their memories of solutions that have worked in the past. Huarte also defined "understanding" (*entendimiento*) as a separate capacity, implying that one can be bright without having a good understanding of a situation. Huarte's distinction between problem solving by imagination or by memory is mirrored in contemporary theories that distinguish between the ability to do abstract reasoning and the ability to apply previously learned solution methods.[3] Robert Sternberg, a prolific modern writer on intelligence, has emphasized the distinction between analytic intelligence and the ability to understand complex social situations.[4]

Huarte anticipated another modern idea, the need to have a biological explanation for intelligence. Huarte offered a theory based on the sixteenth-century notion that the body is governed by four "humors" – blood, bile, black bile, and phlegm. This theory of biology has long since been discarded. The idea that there should be a biological explanation for individual variations in cognition has been retained. One of the most active areas of modern intelligence research deals with the relation between intelligent behavior and the brain.

Let us leap from the sixteenth century to the nineteenth, and to one of the most colorful characters in the history of science, the Victorian physician, mathematician, and explorer Sir Francis Galton. Galton explored in Africa, made major contributions to the development of statistics, and conducted research in psychology, most noticeably on intelligence. He wholeheartedly endorsed the theory of evolution proposed by his cousin Charles Darwin. Galton believed that human intelligence was largely inherited. He also maintained that intelligence was one manifestation of a person's overall constitutional fitness. Therefore,

2 Huarte de San Juan, 1575/1991.

3 Horn & Noll, 1994.

4 Sternberg, 2003.

it should be possible to learn something about a person's intelligence by examining his or her physique, including brain size, and by determining the efficiency of the person's nervous system by doing such things as recording the speed with which he or she reacted to a signal to strike a bag. These ideas are alive, in much expanded form, today.

The next step was taken at the start of the twentieth century, when the Frenchman Alfred Binet developed the first intelligence tests to be used in schools. Testing has dominated the study of intelligence since then, so we need to look more closely at the idea.

1.2. Testing

If you want to go beyond saying "people are different" you have to offer some way of measuring those differences. There is an imperative to develop such measures if a society wants to assign different roles to different people, based on their personal characteristics. Not all societies do this, all the time. There was no intelligence test for the ruler in the hereditary kingships of medieval Europe. The Hindu caste system pre-assigned people to social roles, based upon their birth. It is notable, though, that both these societies experienced a good deal of conflict due to their restricting people's social roles.[5]

In village-based societies personal knowledge of individuals plays a major role in assigning people to jobs. When American pioneers began to move into the northern plains states Sitting Bull, the paramount chief of the Lakota (Sioux) Indians, selected Crazy Horse to be war leader from among people whom he knew personally.[6] That technique does not work in today's large societies, where there are many positions to be filled in both government and industries. Leaders cannot possibly know all their subordinates, let alone the subordinates' subordinates. Our society requires formal machinery for selecting candidates either into employment, directly, or into educational systems that serve as channels to future employment.

Many societies solve this problem by an elaborate form of recommendations. A boy or (historically less often) a girl who is thought to be talented is sent off for training and/or apprenticeships. While the details have been lost, this appears to be the way the ancient Egyptians selected children to be trained as scribes. It was also the way in which second, third, and fourth sons were recruited for the priesthood (or the army) in medieval Europe. The person was not needed at home, and somebody had connections enough to start them on a career. The use of connections is certainly not unknown in modern times. But we do rely on another method of personnel selection: testing.

There have been many objections to testing. In evaluating them it is well to keep one thing in mind. Society needs a mechanism for personnel selection. Not everyone can have whatever they want. Students have to be selected, jobs have to be filled, and when behavioral problems arise, mental competence must be assessed. If you do not like testing, what is your alternative?

1.2.1. *Testing Before Psychology*

Modern psychologists did not invent testing. In the days when the Chinese emperor claimed to rule "the Earth, the Moon, and three quarters of the Sun" an elaborate series of local, regional, and nationwide tests was used to select officers for the imperial bureaucracy. Candidates had to write traditional poetry and to explain the importance of fearing the will of heaven and knowing the words of the sages. Evidently it was assumed that a person who could do these things could collect the imperial taxes or be an ambassador to the Mongols.

5 During the Hundred Years' War between England and France (roughly 1350–1450) the French court was disrupted when Charles the Foolish inherited the throne. He probably suffered from a bipolar psychosis. In India the Sikhs were formed largely as a protest against the rigid social structure enforced by the Hindu caste system.

6 He did well. Crazy Horse defeated General Custer at the Battle of the Little Big Horn. This was the most stunning Native American victory in the history of the western expansion.

Some centuries later the British Empire emulated the Chinese Empire. Until after World War II career positions in both the Indian empire and the British upper bureaucracy were filled largely from the ranks of people who had read history, classics, and occasionally economics at the elite universities, especially Oxford and Cambridge. The British assumed that someone who could do well on oral and written examinations of the writings of Horace and the wars of Caesar would have the ability to administer India or to ferret out secrets while on Her Majesty's Service.[7]

Such techniques of personnel selection may sound quaint to us, but, on the whole, they worked. The Chinese bureaucracy held the empire together in a way that no succession of able emperors could ever have done alone. Properly educated British gentlemen administered India reasonably well for two hundred years. The classics program at Cambridge University produced at least four remarkably effective twentieth-century spies. Unfortunately, they spied against Britain for the Soviet Union, but that is a motivational rather than a cognitive aberration.[8] They were good at what they did.

Why did these exercises in testing apparently irrelevant knowledge do a reasonably good job of selecting people able to run very large, complex empires? Or for that matter, able to fool a modern counterintelligence agency?

What the British and the Chinese had stumbled on, and what we today attempt to evaluate, was a collection of mental traits that, collectively, we call *intelligence*. These traits define individual differences in skills that have broad application in many settings. One of the most important aspects of intelligence is an ability to learn. You demonstrate this by showing that after exposure to knowledge you have learned something. The skills needed to learn the wisdom of Confucius or the philosophical ideas of Socrates are not exactly the skills you need to run an empire, *but there is an overlap*. For that matter, the skills needed to do well on a college entrance test are not exactly the skills you need to acquire a bachelor's degree, *but there is an overlap*. That is why both the classic and the modern testing systems work. It is also why they work imperfectly.

1.2.2. *Alfred Binet Invents Modern Intelligence Testing*

At the start of the twentieth century the French Ministry of Education had a problem. The idea of universal public education had been accepted, but the schools did not seem to work for some students. How did this problem arise?

France, like all modern democracies, was (and is) committed to providing public education for all its citizens, so that all children have an opportunity to compete for desirable positions in society. This goal is not easy to achieve.

Modern schooling is an historically unusual form of education. Before 1800 most humans were educated "on the job" – observing and then helping adults, and serving as apprentices. Universal education, the requirement that every child learn by practicing seemingly esoteric exercises in a setting divorced from everyday life, is a late nineteenth-/early twentieth-century idea. By 1900 it was apparent to educators that some children have a great deal of trouble learning in this manner. The French educational administration needed to have a way of identifying such children, so that they could either be dropped from the system or channeled into an educational program more suited to their capabilities.

The fact that France was a democracy imposed an added constraint. French educators needed an objective method, in preference to the subjective impressions of "persons in authority," the teachers and principals. In an authoritarian regime there is no

7 Gardner, Kornhaber, & Wake, 1996, pp. 12–16.
8 Anthony Blunt, George Burgess, David MacLean, and Kim Philby. Burgess, MacLean, and Philby fled to the Soviet Union when they were about to be exposed. Blunt stayed in England, undiscovered, and became a distinguished art historian. His espionage role, which seems to have lasted through the 1940s and 50s, was not publicly revealed until 1979.

need for such a method; if the authorities don't like you, you're out.

So from the very first, testing was embedded in the society that required it.[9]

In order to meet this challenge the Education Ministry hired Alfred Binet, who had worked for a while with Galton. Binet began by making two important assumptions. The first was that mental competence increases over the childhood years. The typical six-year-old can solve problems a four-year-old cannot; a four-year-old can solve problems a two-year-old cannot, and so on, at least from birth to the late teenage years. Therefore, it makes sense to talk about *mental age* (MA) – the level of mental competence at which a child is operating.

Binet took a pragmatic approach to the measurement of mental age. He asked experienced teachers what sorts of problems children could solve at different ages. Once he had a set of problems typical of the ones children could solve at age six, seven, eight, and so on, he could determine a person's mental age by finding the most difficult problems that a child could solve. Mental age could then be compared to chronological age (CA), to determine whether a child has been performing below, at, or above the cognitive level that would be expected.

Binet then made his second, more arguable, assumption. He assumed that a child's relative standing in mental development, compared to his age group, will remain fairly constant as the child grows up. If Sammy and Tommy are both six, but Sammy has a mental age of eight and Tommy one of five, Binet assumed that four years later, when they are both ten, Sammy would have a mental age higher than ten, and Tommy a mental age lower than ten. Therefore, it follows that if you test children on entrance to school (age six or seven), and you find that some are markedly behind (i.e., have mental ages in the three-to-four range), those children are likely to be behind their classmates at all ages, and therefore are candidates for removal from the normal school program. That is what the French education system wanted to know.[10]

The Education Ministry accepted Binet's argument. The modern era of intelligence testing had been born.

1.2.3. *The Intelligence Quotient (IQ)*

Binet did not use the term *Intelligence Quotient* (IQ) because the concept of mental age was sufficient for classifying children who were entering school. As mental testing expanded to the evaluation of adolescents and adults, however, there was a need for a measure of intelligence that did not depend upon mental age. Accordingly the intelligence quotient (IQ) was developed. There have been changes in the definition and use of the term since its introduction. The details are provided in panel 1.1. Here we proceed directly to the modern use of the term.

The term IQ is used in two ways, which I will call the *narrow* and *broad* uses of the term.

The narrow definition of IQ is a score on an intelligence test, developed according to a scoring protocol where "average" intelligence, that is, the median level of performance on an intelligence test, receives a score of 100, and other scores are assigned so that the scores are distributed normally about 100, with a standard deviation of 15. Some of the implications are that:

1. Approximately two-thirds of all scores lie between 85 and 115.
2. Five percent (1/20) of all scores are above 125, and one percent (1/100) are above 135. Similarly, five percent are below 75 and one percent below 65.

9 The contrast between the French and the Chinese and British imperial systems is informative. The Chinese and British systems were designed to select a sufficient number of qualified candidates for government functions. So long as the supply of young officers and bureaucrats was adequate, there was little concern that the system might have shut out potential candidates. The French testing program was designed to staff society, without favoring some citizens over others. Any government has to solve its staffing problems. Only democracies have to justify the staffing system to the citizens.

10 Binet & Simon, 1905.

Panel 1.1. The Intelligence Quotient (IQ)

Mental age is inadequate as a means of comparing the intelligence of two children of different chronological ages (CA). Suppose a six-year-old and a ten-year-old both have a mental age (MA) of eight. Cognitively, they are likely to be very different individuals, for one is developing rapidly and the other is developing slowly. The need for a measure of mental development that is independent of chronological age led to the concept of the *Intelligence Quotient*, IQ, which was originally defined as the ratio of mental age to chronological age, multiplied by 100.*

$$IQ = 100 \cdot \frac{MA}{CA}. \tag{1.1}$$

To illustrate, a ten-year-old who can solve problems at the level of difficulty expected of a twelve-year-old would have $MA = 12$, $CA = 10$, $IQ = 120$. An IQ of 100 indicates that the child's cognitive development is proceeding on schedule, an IQ above 100 indicates acceleration, and an IQ below 100 indicates slowed development. In the case of our hypothetical six-year-old and ten-year-old, both of whom have a mental age of eight, the first child would have an IQ of 133 and the second an IQ of 80. In modern educational terms the first child might be considered for an accelerated program, while if the second's IQ score were accompanied by difficulties with schoolwork he or she would be a candidate for a special education program.

This method of calculating IQ will not work with adults, because intelligence does not increase linearly with age past childhood. A man of sixty whose mental powers are equal to those expected of a forty-year-old would not be considered a case of retarded development! Therefore, the IQ ratio just described has been replaced by a measure based on the notion that IQ should reflect a person's relative standing within his or her own age group.

Intelligence tests are *standardized* by giving them to a relatively large sample of people chosen to be representative of the population for whom the test is intended. In the case of a test intended for broad use, such as establishing the mental competency of adults, this is essentially the entire population of the country, so attempts are made to obtain a large representative sample of the United States, the United Kingdom (for the British version), Spain (for the Spanish version), and so forth.† The sample should be sufficiently large that a distribution of scores can be obtained for different age groups.

The *raw* test score is based on the number of items correctly answered and/or the difficulty level of the item. (The scoring algorithm varies somewhat with the test, as discussed in Chapter 2). The IQ score is derived from the raw score in the following way.

Let Y be a raw score for a person in a particular age group, and let M be the mean and S the standard deviation for all scores in the reference group.‡ The corresponding *standard score* is

$$z = \frac{Y - M}{S}. \tag{1.2}$$

The IQ score is derived from this by the conversion

$$IQ = 15 \cdot z + 100. \tag{1.3}$$

If the raw scores were normally distributed, then the resulting IQ scores will be normally distributed with a mean of 100 and a standard deviation of 15. The graphic depiction of this distribution is the famous "Bell Curve." The Bell Curve for IQ distributions is shown in Figure 1.1.

Why not record scores in the standard score format? Statisticians, psychometricians, and research workers would probably prefer to do this. The advantages have to be weighed against two countervailing trends: tradition and public relations. IQ scores were introduced almost a century ago; people are used to them. Standard scoring is a bit esoteric for the nonstatistician. For a statistician, having a scale with a mean of zero and a standard deviation of 1 is a convenience, nothing more. If a (nonstatistical) parent were to be told that their child had scored a zero on an intelligence test, they might interpret this as a claim that their pride and joy had the intelligence of a rock! Half the people who took the test would receive negative scores, which could lead to all sorts of misunderstanding. Appearances can be important.

* The concept of IQ was developed by the German psychologist William Stern, not by Binet.

† This procedure contains the implicit assumption that the distributions of intelligence are the same across populations. Thus if you consider a score of 100 on the Spanish form of the test to indicate the same thing as a score of 100 on the American version of the test, you are implicitly assuming that Americans and Spaniards have equivalent intelligences, on a population basis. Such assumptions have been vigorously debated. The controversies are discussed in detail in Chapter 11.

‡ The standard deviation is a measure of the amount of variation in a population. More details are given in Chapter 2.

Thus IQ, in the narrow sense, is a score indicating a person's relative performance on an intelligence test, compared to the performance of people in an appropriately chosen comparison group. This does not completely clarify the matter, because there can be debate about what counts as an intelligence test. This matter is also discussed on page 8. I will attempt to be clear about how the term is being used in various contexts.

In the broad sense the term IQ is used as a synonym for intelligence, that is, as a shorthand term for individual differences in cognition. This can lead to confusion, so I shall attempt to use IQ only in the sense of a test score. The term *intelligence* will be used to refer to the broader concept of individual differences in mental ability. In my usage, a person who has high intelligence will probably have a high IQ score, but the distinction between the two is important.

In interpreting IQ scores it is often useful to think of percentiles, which indicate the percentage of people in the referent group whose scores are below a certain level. What that level is will be determined by the IQ score and by the properties of the Bell (normal) Curve itself. Table 1.1 gives some important reference scores. The properties of these scores follow from the assumption that IQ scores will follow the normal distribution, which is illustrated in Figure 1.1.

As a result, in terms of the modern scoring, if someone says that their child has an IQ of, say, 120, this does not mean that the child's mental age is 20% higher than his chronological age. It means that the child has a test score in the top 9% of test scores at the child's age.

Why can we assume that IQ scores are distributed normally? The answer is simple. IQ tests (and many other tests) are constructed by choosing appropriate numbers of easy, intermediate, and hard items, so that the total scores will be normally distributed in the population for which the test was intended. The Bell Curve is an artifact of the way the test is constructed! There is no definition of IQ independent of the tests themselves. This contrasts with a variable like height, which is defined independently of yard sticks or meter sticks. Height happens to be distributed approximately normally, within the populations of adult males and females. The distribution of height is a fact of nature. The fact that IQ test scores are normally distributed is an outcome of the test construction procedure. Nevertheless, it is a reasonable thing to do. Why?

IQ scores are used to describe people *relative to each other*. They are also used to

Table 1.1. The distributions of standard and IQ scores in terms of the percentage of people above or below selected scores. An IQ of 65 would, if accompanied by other indications of mental incompetence, be cause for considering a person mentally disabled. If IQ is distributed normally, about one percent of all people have IQ scores this low. Average IQ is, by definition, 100. Approximately half of all scores lie between 90 and 110. About 16% of the scores lie above 115, 2% above 130, and 1% above 135. *MENSA*, an organization whose members have high IQ scores, defines the *4 sigma* group, people with IQs over 160. (*Sigma* is a term frequently used to refer to the standard deviation.) This level of score would be expected three times in every 100,000 observations.

Standard Score (z)	*IQ Score*	*% Below*	*% Above*
−2.33	65	0.982	99.018
−2.00	70	2.275	97.725
−1.00	85	15.866	84.134
−.67	90	25.249	74.751
0.00	100	50.000	50.000
.67	110	74.751	25.249
1.00	115	84.134	15.866
2.00	130	97.725	2.275
2.33	135	99.018	0.982
3.00	145	99.865	0.135
4.00	160	99.997	0.003

make predictions and to indicate associations, as in predicting a student's likely academic progress or investigating the association between intelligence and income. There are technical reasons for wanting to deal with normally distributed scores when we apply the statistical methods used for making predictions and analyzing associations.

There is another, less technical reason for requiring that IQ scores be normally distributed. Many other human qualities that can be measured on scales with physical interpretations, like height and weight, are distributed normally. It seemed to many of the early researchers that if we could measure intelligence in some physical manner, such as measuring the efficiency of the nervous system, these measures would probably turn out to be normally distributed. Therefore, it seemed appropriate to require that IQ scores be normally distributed.

In the late nineteenth and early twentieth century this reasoning seemed compelling, because the normal distribution itself was regarded (almost mystically) as a Law of Nature. Today we are a bit more skeptical, but there is still a good argument for assuming normality. If a person's intelligence is due to a large number of independent causes, each of which has a small effect, intelligence would be distributed normally across the population.

A certain amount of the confusion between the broad and narrow senses of IQ is due to the way in which cognitive tests are described. Some tests are explicitly marketed as intelligence tests. But because the term IQ, and sometimes even the term *intelligence*, have acquired a bad taste in certain circles, many tests of cognitive skills are not marketed as intelligence tests, even though these tests are highly correlated with tests that *are* marketed as intelligence tests! For instance, in a widely read and highly controversial report, Richard Herrnstein and Charles Murray used the Armed

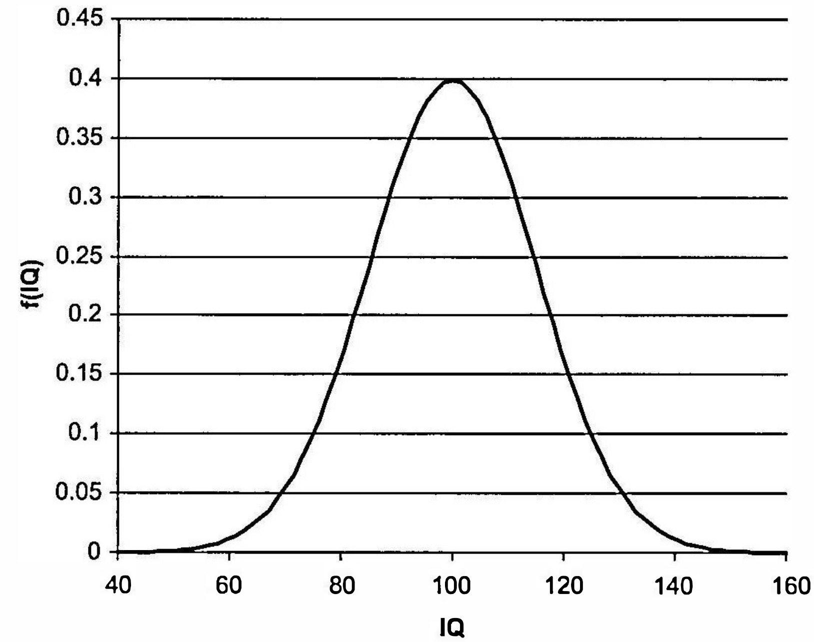

Figure 1.1. The "Bell Curve" for IQ. The area under this curve represents 100% of the population. The area under the curve and to the left of a given IQ value (on the abcissa) represents the fraction of people in a population who have IQs lower than the indicated IQ. Conversely, the area to the right indicates the fraction of people who have this IQ or a higher one. For example, 50% of the area under the curve lies to the left of IQ = 100, indicating that half the population has an IQ of less than 100. Nine percent of the area under the curve lies to the right of 120, indicating that only nine percent of all people have IQs of 120 or higher. The Bell Curve for IQ scores is a special example of the normal, or Gaussian, distribution. At the extremes, the curve never quite touches the abscissa ("x axis"), but this cannot be shown on the graph.

Services Qualifying Test (AFQT) as a measure of intelligence, and treated AFQT and IQ scores as being virtually synonymous.[11] The US Department of Defense never refers to the AFQT as an intelligence test. Similar confusions arise with the SAT. Many research projects have used SAT scores as a measure of intelligence, although the test's publisher, the Educational Testing Service, does not describe it as an intelligence test.

There is great controversy over whether or not IQ scores should be treated as real indicators of mental ability. Panel 1.2 presents a historical debate that took place in the 1920s, but in many ways foreshadowed contemporary arguments. I shall come down squarely in the middle of the controversy. I will argue that the scores certainly do mean something, but they may not mean as much as some enthusiasts claim.

11 Herrnstein & Murray, 1994.

1.2.4. *What Binet Discovered: "Drop in from the Sky" Testing Works*

Let us take a closer look at what Binet assumed and what he found.

Binet's assumption that mental competence increases as children grow older is certainly correct. Mental competence may decrease in old age, but that is another story, and was of no concern to Binet. He was also correct that there are marked individual differences in the rate at which mental competence increases.

His second assumption, that relative standings remain constant as children age, is true on the whole, but there are exceptions. As a toddler, Albert Einstein was a relatively

Panel 1.2. Defining Intelligence: The Debate between Mr. Lippmann and Professor Boring

In science clarity of definition is essential, for good definitions make clear what the important questions are. The study of intelligence has been plagued by a lack of precise definitions. The debate between Lippmann and Boring, early in the twentieth century, shows how the failure to define terms introduced confusions that continue to this day.

Following the use of tests in World War I, intelligence testing became a growth industry. So, inevitably, it attracted the attention of learned commentators – people who, if there had been TV in those days, would have appeared as talking heads on the Sunday morning pundit shows. One of the most respected of these commentators, Walter Lippmann, did not at all like the new technology. He was particularly incensed by a claim, based on analyses of the Army Alpha data, that the average American had a mental age of fourteen. In Lippmann's own words:

> *The intelligence test, then, is an instrument for classifying a group of people, rather than "a measure of intelligence." People are classified within a group according to their success in solving problems which may or may not be tests of intelligence. They are classified according to the performance of some Californians in the years 1910 to about 1916 with Mr. Terman's notion of the problems that reveal intelligence. They are not classified according to their ability in dealing with the problems of real life that call for intelligence.*
>
> *(Lippmann, 1922a)*

Lippmann argued that the test developers had produced a barrage of statistics that had the trappings of science, but were not scientific. The tests themselves were based on hunches by people such as Terman that this or that behavior indicated intelligence, rather than on any scientific theory of what constituted intelligence. Lippmann also doubted that a classification of people on the basis of test scores would map onto a classification according to their ability to "deal with problems of real life that call for intelligence."

Academia responded. The Harvard professor E. G. Boring* clarified the matter by asserting:

> *Intelligence is what the intelligence tests test.*
>
> *(Boring, 1923)*

So there!

The exchange between Lippmann and Boring foreshadowed a debate that is active today. Should a test be developed inductively, by the pragmatic procedure of identifying people who are believed to vary in intelligence, seeing what behaviors distinguish those who are intelligent from those who are not, and then incorporating these behaviors into a test? Or should a test be based on a theory of how individual differences in cognitive power come to arise?

The answer is not a simple one. To see why, consider the following analogy. I think that everyone would agree that people differ in their physical fitness. But what is physical fitness? You could take the approach of an athletic coach, and ask people to run, jump, throw weights, and so forth in order to determine physical fitness. Alternatively, you could take a medical approach. Physical fitness depends on muscular adequacy, reaction, and the ability of the heart and lungs to provide fuel to the muscles. So let us take measures of cardio-pulmonary capacity and construct tests of the strength of isolated muscle groups and the speed of neural impulses.

Binet, Terman, and their many successors took the coach's approach. Lippmann seems to have wanted a more theoretically justified approach, although he did not offer one. ~~(After all~~, he was a journalist, not a scientist.) Lippmann did not deny that something called intelligence exists, for he spoke of dealing with problems that call for intelligence. His objection was to the tests offered to evaluate it.

Boring's response, identifying intelligence with a test score, strikes many (including me) as somewhat arrogant, for it confounds the concept of intelligence with a score on an imperfect indicator of intelligence. However, it does lead to a research agenda.

The first item on the agenda is to ask whether or not the tests identify people who would otherwise be considered intelligent. Lippmann said they would not, but that was his opinion, not an observation of fact.

To some extent the way in which the tests were developed ensured that they did identify "the intelligent." Binet and Terman began with groups who had known variations in cognitive competence (e.g., younger and older schoolchildren), and then developed tests whose scores reflected the variation. Lippmann pointed out that this procedure makes the definition of intelligence dependent upon an appropriate choice of examinees for the initial (*standardization*) testing. This was an important observation. Today there is a lively debate over the appropriateness of using intelligence tests that have been standardized in the post-industrial world to evaluate intelligence in the developing nations of Africa, Asia, and Latin America.

But Boring had a point, too. If test scores are accepted as good indicators of intelligence, the tests become a powerful tool for investigating individual differences in cognition. You can determine how test scores are related to variations in brain structure and/or process, genetic makeup, schooling, family support, and a variety of other causal factors. Test scores can also be used as predictors of success or failure in all walks of life. We can then consider different explanations for observed relationships. Science advances by accounting for relationships between observables, not by debating beliefs about what ought to be the case.

Boring's view has lead to the *psychometric* approach to intelligence. It has resulted in a viable research agenda. But it contains a weakness.

If intelligence is equated with an IQ score, the study of intelligence is reduced to an analysis of the causes and consequences of test scores. So if the original process of test construction missed something important about intelligence, the assertion that intelligence is what the tests test makes it hard to incorporate new perspectives. Lippmann was right to be concerned over the conceptual limitations of a fascination with test scores.

The argument between Lippmann and Boring foreshadowed future developments. There are many advocates of the psychometric approach today. It is not entirely devoid of theory. In fact, the analysis of pragmatically defined test scores has suggested theoretical positions about intelligence that can then be expanded to include new measures. Some of these efforts will be presented in later chapters.

There are also contemporary psychologists who accept Lippmann's concerns (although he is seldom cited!) and then go beyond criticizing to try to develop a theoretical basis for intelligence and intelligence testing. Such a theory will be useful if it leads to a coherent picture of how individual differences in cognitive power arise and influence human experience from birth to death.

In Section 1.4 I present my own views about what such a theory has to deal with, and explain the difficulties that it encounters.

* Boring, 1923.

late talker! Most of us can think of someone who was not terribly impressive in grade school or high school, but who had a great college career. Things go the other way, too. The smartest kid in grade school does not always become a Phi Beta Kappa in college, let alone becoming a prize-winning physicist or a fabulously wealthy entrepreneur.

Such examples are striking, but they are exceptions to a general rule. Past the age of about ten, indicators of relative cognitive competence are fairly stable. A Scottish study found a substantial correlation between intelligence test scores taken at age eleven and subsequent measures taken when the examinees were in their sixties and seventies.[12] In psychological jargon intelligence is a *trait*, a characteristic of the individual that is reasonably stable over time and that is revealed in many situations.

Binet's third assumption was the unnoticed 600-pound gorilla in the room. Let us accept that there are individual differences in the mental competencies required to be successful in modern society.[13] These competencies, collectively, are what we mean when we say "intelligence." Binet assumed, and then showed, that it is possible to make reliable measurements of a substantial number of the traits that constitute intelligence, within the space of three or four hours.

I suggest that the reader go back and weigh each clause of that sentence. Binet showed that *some* of the cognitive traits important to success can be evaluated within a single testing session, conducted outside of the context of everyday activities. I like a term originally coined by Robert Mislevy,[14] *Drop in from the Sky* testing, because it captures the examinee's view of the out-of-context nature of the assessment. Much of the research in the century following Binet can be looked upon as an attempt to refine and expand the measurements that can fit into the "Drop in from the Sky" paradigm.

Throughout this book I shall argue that by accepting the conventional testing paradigm as a given, intelligence researchers have too often ignored traits that cannot be measured within the paradigm but that, by any reasonable definition, are part of intelligence. Nonetheless, a nontrivial subset of human cognition does fit into the "Drop in from the Sky" paradigm, and hence can be evaluated using conventional testing methods.

1.2.5. *Expansions of Binet's Work: The Stanford-Binet and Wechsler Tests*

Binet's tests excited a great deal of interest. Educationally, the need for classifying students was a problem for the burgeoning public educational systems of the early twentieth century. Intellectually, the tests provided a foothold for applying the new, but rapidly developing, science of psychology. Lewis Terman, a professor at Stanford University, translated and modified Binet's tests for use in the United States. Due to the interruptions caused by World War I, Terman did not complete his work until the 1920s. The resulting test, the Stanford-Binet Intelligence Test, is still used in regularly updated form today.

The Binet and Stanford-Binet tests were intended for use with schoolchildren, roughly age fifteen and younger. In the late 1930s David Wechsler, a clinical psychologist working at New York City's Bellevue Hospital, created a similar test for adults, the *Wechsler-Bellevue* test. It has subsequently been modified into the *Wechsler Adult Intelligence Scale* (WAIS). It and a companion test for children, the *Wechlser Intelligence Scale for Children* (WISC), are the most widely used individually administered intelligence tests today. The Wechsler tests are often referred to as the "gold standard" for intelligence tests.[15] As of 2010 the fourth

12 Deary et al., 2000.

13 Individual differences in mental competence probably occur in every human society, although the relative importance of specific competencies may vary from society to society. The ability to locate one's self in space is less important in a society that has street signs and global positioning systems than in a society of hunter-gatherers.

14 Mislevy, formerly at ETS and now a professor at the University of Maryland, is a highly respected specialist in the field of psychometrics, the mathematical analysis of test scores.

15 Matarazzo, 1972.

revisions of the tests, WAIS-IV and WISC-IV, were in use.

Both the Wechsler and the Stanford-Binet tests are individually administered. The examinee sits down with a trained examiner and attempts to solve a series of problems, divided pragmatically into problems that do or do not involve language, and that vary in the demands that they place on memory. Wechsler has described this as an opportunity for the examinee to display his or her cognitive skills during a standardized interview with an experienced observer.[16]

The resulting IQ scores have proven to be highly useful, both in the educational system and in other settings. For instance, the WAIS is widely used to evaluate a person who, for whatever reason, is of suspected mental disability. Examples of such use are the adjudication of legal competence and the analysis of status following brain injury. Other applications of these tests, and there are many, are extensions of these ideas. The focus is always on the individual. The cost of testing is evaluated relative to the potential benefits of the results in making judgments about an individual case. Are the decisions made about this person improved by knowing test scores, and if they are, is the value of a typical decision enough to justify the costs of the test?[17]

The Wechsler and Stanford-Binet tests are not the only individually administered intelligence tests. However, they have played a very important role in the development of testing. Some further tests will be discussed in Chapter 2, where the tests are described in more detail.

16 Wechsler, 1975.

17 The law provides a dramatic example. The United States Constitution prohibits "cruel and unusual punishment" but does not define it. Several of the states and the current federal code include capital punishment for treason and for particularly vicious cases of murder. The Supreme Court has held that it would be cruel and unusual punishment to execute a mentally retarded person for committing a capital crime, on the grounds that the person, though guilty, could not understand the crime or its consequences. As of 2009 a defendant's attorney could offer an IQ test score as one piece of evidence of mental retardation and, hence, as a reason for not assessing the death penalty.

1.2.6. *The Development of Group Testing*

The next major step in intelligence testing was a spin-off of technology from a military application. When America entered World War I the army had to make rapid mental evaluations of large numbers of incoming soldiers. The War Department[18] sponsored development of a test that could be administered to large groups of recruits. Psychologists responded with the *Army Alpha* Test, a written test suitable for group administration, and the *Beta* test, which could be given to nonliterates.[19]

The program was considered a success. Today the militaries in virtually all developed nations routinely use cognitive tests to screen recruits. The tests that the U.S. military uses will be described in Chapter 2.

The military tests are examples of personnel classification tests. Cognitive tests for personnel classification are widely used in the civilian sector as well as in the military. The costs and benefits of testing within a personnel classification system are not the same as the costs and benefits of testing intended for individual counseling and/or placement. In a personnel classification system correct classifications have a value and incorrect classifications have a cost, as seen from the perspective of the institution setting the test, rather than as seen from the perspective of the examinee. A classification test is economical if, on the average, the cost of administering the test is less than the value of improved decision making. This view shifts the focus from decisions about an individual to the average value of a decision, calculated over the entire population. The shift greatly affects the economics of testing. A cheap test, which makes only a moderate improvement in the accuracy of the selection decisions, administered to thousands of or even millions of people, can be a valuable classification instrument.

18 Now renamed the Department of the Army and absorbed into the larger Department of Defense.

19 In the United States today illiteracy is almost a signal that the individual is of suspect intelligence. That was not true in 1917–18, for universal public education was not the case. Many people of normal intelligence were illiterate, especially in rural areas.

Well over a hundred group-administered classification tests have been developed. They include the Scholastic Ability Test (now renamed by its acronym, the SAT) used in the college admissions process, and the General Aptitude Test Battery (GATB), which for years was used by the United States Department of Labor to provide a test score to guide in industrial hiring, and many more.

Describing all the tests in use today literally takes a volume, and the volume has to be updated annually. Some of the more prominent tests are described in detail in Chapter 2. The important point is that since the 1930s intelligence testing has been widely used to make important decisions about people's academic and vocational careers. Testing is also used as a guide in medical rehabilitation, such as evaluating the course of treatment following insults to the brain. The tests are also widely used in research on the description, causes, and consequences of being intelligent.

1.3. Do the Tests Work? Efficacy and Controversy

There has been a great deal of debate about how well intelligence tests work. The debate is not surprising, because the issue is inherently complicated. How accurate does a test have to be before we say that "it works"? There can be different standards for different purposes. How well do we understand why the tests work? People are uncomfortable using indicators that they do not understand. To the scientist interested in intelligence, though, this question ought to be a challenge. What are the consequences of basing important decisions about education, employment, and personal planning on test scores?

We will look at the question of test accuracy in considerable detail in Chapter 10. Here I will do three things: explain the criterion by which tests are judged as being accurate or not, provide just a few interesting statistics (saving the rest for Chapter 10), and then describe the controversies.

Tests would be perfectly accurate if, whenever a test score was used as a predictor, there was some critical score such that everyone who had a score equal to or higher than the critical score succeeded, and everyone with a lower score failed. No such intelligence test exists, and none ever will. Test scores are not perfect indicators of a person's cognitive power; other things besides intellectual talent determine success; and success itself is not an either/or thing. The question is not whether or not test scores can be used to make perfect predictions of success or failure; the question is whether or not using test scores reliably improves our ability to predict who will succeed or fail. Tests that do this are said to be *valid*. Validity is a matter of degree; the more using a test improves prediction, the higher the test's validity.

Here is an example that illustrates the issue.

In the 1930s the Scottish psychologist Godfrey Thomson conceived the impressive plan of testing virtually the entire nation of Scotland. And he managed to do it. In 1932 Thomson's *Moray House* intelligence test was given to almost every eleven-year-old in Scotland, more than 80,000 in all.

About seventy years later Ian Deary, a professor at the University of Edinburgh, and his colleague Lawrence Whalley traced 2,300 of the people whom Thomson had tested. Figure 1.2, taken from their report, shows the fraction of the original respondents to the 1932 test who were still alive at different times from the 1930s until the start of the twenty-first century. The data is shown separately for men and women from the upper quartile (top 25%) and lower quartile (bottom 25%) of the distribution of test scores. Intelligent people live longer!

An intelligence test score obtained in childhood was a valid statistical predictor of length of life. This ought to catch the reader's attention.

Now let's use this fact to uncover some of the complications dealing with test scores, and by implication, intelligence itself, as a predictor.

We first must determine whether or not we are dealing with an artifact due to

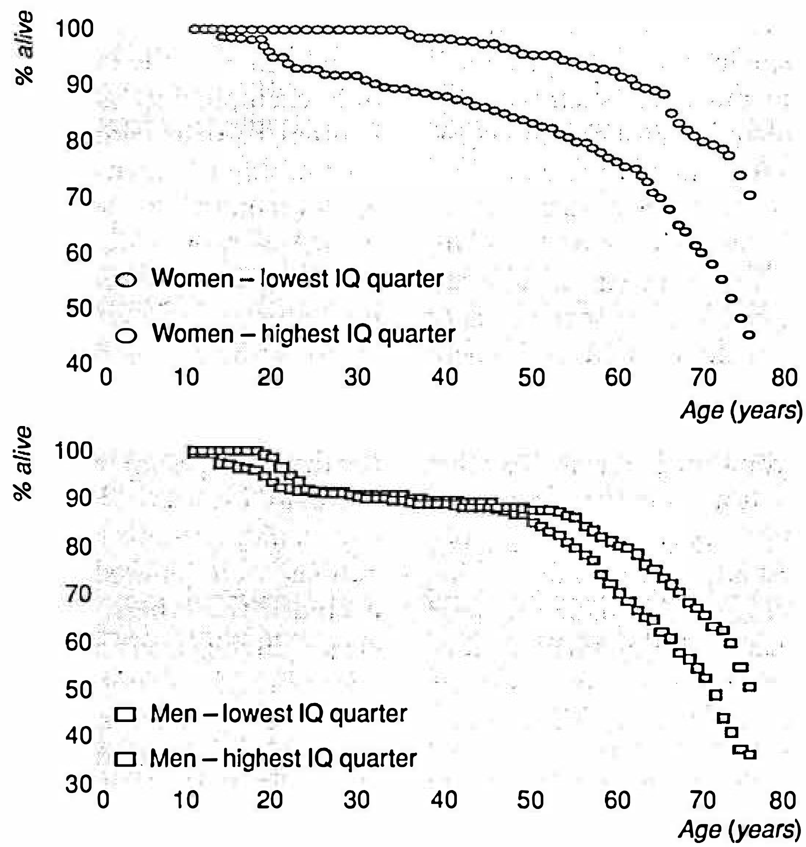

Figure 1.2. The fraction of individuals surviving to various ages for a cohort of approximately 2,200 Scottish children tested at age eleven, in 1932. From Whalley and Deary, 2001, Figure 1. By permission of the authors and the *British Medical Journal*.

some relationship that is forced between test scores and the outcome of interest. For instance, in the United States the SAT scores of people who enter college are higher than the test scores of people who do not enter college. That fact, alone, is uninteresting, because test scores are one of the determiners used to decide who gets to go to college. However, this sort of problem did not arise in the Scottish study, for it is unlikely that a score on a test taken at age eleven has any direct influence on mortality. The statistic indicates some meaningful relationship between mortality and whatever the test measures. But what?

I can imagine three different reasons why a test score might predict mortality. There might be a direct relationship. We know, today, that intelligence test scores are partly determined by the state of the brain. This statement will be documented in detail in Chapters 6 and 7. It could be that the test scores (imperfectly) revealed the state of the brain at age eleven, and this state carried forward over the years, producing an association between test scores and later mortality. This explanation would have appealed to both Galton and Juan Huarte de San Juan.

There could be an indirect relationship. Perhaps intelligence, as revealed by the tests, makes people less likely to do unhealthful things, like drinking alcohol to excess, or driving a motor vehicle after drinking to excess, or forgetting to take vaccinations against influenza, and so on. This explanation – that intelligence is related to finding smart solutions as you navigate your way through life – would have appealed to Binet.

Finally, it might be that the relationship is not due to intelligence at all; it is due to a third variable that both causes test scores (although it is not intelligence as we normally think of the term) and also influences mortality. Parental socioeconomic status (Parental SES) is a term used to refer to the general "place in society" occupied by a child's family. It includes such things as financial status, education, and generally beneficial lifestyle practices, such as having children inoculated against common diseases. It is also an indicator of the extent to which a child is likely to have "family connections" to help establish his or her own lifestyle. There is a positive statistical relationship between children's test scores and Parental SES. It might be that childhood test scores predict mortality because Parental SES assists in the development of a lifestyle that promotes longevity (or the opposite, for low test scores and Parental SES). The contemporary American psychologist Richard Nisbett, a professor at the University of Michigan, has made a strong argument that many of the relationships between test scores and social outcomes are actually relationships between SES and the outcomes.[20]

Which of these explanations do you think accounts for the Whalley and Deary data? Could two or more explanations both be true?

There are *probabilistic* relationships between test scores and a variety of interesting outcomes in life. In general, the higher a person's test score, the more likely that person is to have good things happen to them. These include educational achievement, on-the-job performance, income, marital status, health and longevity.[21] Most of the statistical relationships are high enough to be economically valuable evidence to guide decision making, both in personal life planning and in personnel classification. But because the relationships are probabilistic they have to be understood in statistical terms. This demands some expertise on the part of the person who tries to interpret the facts.

Test scores are not only correlated with outcomes in life, they are also correlated with a variety of other measures that are correlated with the outcomes. The Parental SES example just given is illustrative. It is often hard to determine why there is a relationship between intelligence test scores and outcomes such as educational achievement, job performance, and health. The fact that there is a correlation between an IQ score and X, where X is almost any life outcome, does not prove, alone, that "intelligence causes X." There will almost always be several plausible interpretations for the relationship, and they are often not mutually exclusive. In the intelligence–life expectancy example three alternative explanations were offered, and none of them excluded the others. The questions "What causes intelligence?" and "What does intelligence cause?" are not easily answered.

Do not expect simple answers about intelligence in this book. I shall try to be as clear as I can, but some things are necessarily complex.

If you encounter simple answers somewhere else, my advice is to be very, very suspicious.

1.3.1. *The Controversies*

Because test scores are used to make major social decisions, it is hardly surprising that they have been controversial. One of the first of these debates occurred in the 1920s, when the social commentator Walter Lippmann published a series of articles criticizing intelligence testing in *The New Republic*, an influential magazine in liberal intellectual circles. Lippmann's articles provoked a furious response by Terman and, in somewhat more measured terms, a critique by the Harvard professor E. G. Boring. Boring's reply included the somewhat famous definition of intelligence as whatever the tests test.

The debate between Lippmann and Boring is described in Panel 1.2. Their argument touched on concerns about intelligence testing that are still raised today. Other concerns have also been raised. Most of these concerns deal with complicated facts about what intelligence is worth, and who has it – topics that are discussed in detail in

20 Nisbett, 2009.

21 There are a great many studies backing up these assertions. See Neisser et al. (1996) for the statement of the situation by a review panel of the American Psychological Association, Gottfredson (1997, 2007b) for reviews, and Herrnstein and Murray (1994) for a comprehensive analysis involving the AFQT. Chapter 10 discusses the topic in much more detail.

Chapters 10 and 11. Here I will foreshadow the discussion of the controversies briefly, by introducing the objections to the tests and outlining the nature of the responses.

> Objection 1. *The tests cannot possibly work. It is unreasonable to believe that performance on a "Drop in from the Sky" test, made up by people who do not know the examinee, could possibly reveal important mental traits.*

A slightly more sophisticated way of saying this is that the objector believes that human thought is so subtle that its nature could not possibly be captured in a single interview. This makes the objection seem much less anti-intellectual, but it amounts to the same thing as my original, italicized version.

I do not know whether to call Objection 1 a Know-Nothing attitude (I don't want to learn) or a Know-All attitude (I already know the answer, so your tests are not necessary). In any case, this objection dominated the debate between Lippmann and Boring (Panel 1.2), and it still arises today. I do not believe that it is productive.

> Objection 2. *The tests don't work. There are people who have only modest test scores and do well, and people who have very high test scores and are not doing notably well.*

Here the objector asks that all predictions made on the basis of test scores be accurate. This is an impossible goal. Intelligence is not the only thing that determines success or failure, in academics or in life in general.

I conjecture that the objection is rooted in two facts – a psychological one and a sociological one.

The psychological fact is that when people have to deal with probabilistic relationships they try to get an intuitive feeling for the relationship, rather than analyzing numbers. This leads to an overinterpretation of exceptional cases. Because the tests do work pretty well, on the average, failures stand out. People remember the case of the person with low test scores who did brilliantly, or the one with high test scores who did miserably. The fact that most people perform about as well as their test scores predict, with a small amount of variation, just does not stick in memory.

The sociological fact is that many of us live in a cognitively stratified society. Children of parents with high, middle, or low SES go to schools with children of the same background as their own. Academic testing mechanisms stratify college students on the basis of SAT or similar test scores. In the workplace our coworkers tend to be of roughly the same intelligence as we are. When we deal with people of greatly varied intelligence we tend to do so in a stereotyped way. A sales clerk, for instance, encounters customers of highly varied intelligence, but does not interact with them in a way that would reveal their intelligence. Stratification occurs at home, because there is a great deal of residential segregation by socioeconomic status; therefore, neighborhood life is "stratified by intelligence."

The result is that within a person's local society there will not be a great deal of variation in intelligence, so other factors, such as variations in personality, will play a role in determining social success. Therefore, when people try to evaluate intelligence by referring to their personal experiences they are likely to undervalue the role of intelligence in the general society.

The poor statistician has an uphill climb when trying to show that intelligence really is an important part of the big picture.

> Objection 3. *The tests work only in the academic arena.*

This objection flies in the face of the evidence. The point will be addressed directly in Chapter 10, where evidence will be given showing that test scores predict success in both the academic and the industrial/economic sphere. Because the argument has to be made statistically, it has to counter the results of the intellectual stratification of the workplace, as described in the response to objection 2.

There is also another problem. Leaving aside spectacular cases, like billionaires heading companies that pioneer new technology, it is hard to establish who is succeeding in the workplace on an individual basis. This makes the statistical analyses more complex than they are in the case of analyzing academic success, and hence harder to explain to the nonstatistician.

> *Objection 4. The tests work only for certain demographic groups, notably middle-class whites. The tests do not predict well for other groups.*

This issue will be addressed directly in Chapter 11. Foreshadowing, while some data is lacking, it is usually found that test scores have about the same power to predict achievement in all demographic groups.

> *Objection 5. The tests should not be used because they are prejudiced against minority group members, who tend to get low scores.*

This objection raises one of the most incendiary topics in psychology: the possibility that there are individual differences in intelligence across racial and ethnic groups. The topic is discussed in detail in Chapter 11. Here I present a brief statement of the issue.

In the United States, and in other industrially developed nations, there are differences in the average intelligence test scores of different racial/ethnic groups. In the United States the order of scores, from highest to lowest, is Asian-derived groups, European-derived groups, Latin-American–derived groups, and then African-derived groups. The Asian-European gap is reliable, but rather small. The European-Latino and European-African gaps are substantial. The debate is not over the existence of differences, it is over their meaning and their implications for action. These are two separate issues.

If the tests predicted performance only in the white group, discrepancies in test scores across groups could simply be disregarded. However, that is not the case; the test scores do predict performance in all racial/ethnic groups. Although many strong opinions have been expressed, there is no consensus about the causes of group differences in IQ scores. Because this debate requires a good understanding of intelligence itself, I have postponed the discussion of racial and ethnic differences until the next-to-last chapter.

Should test scores be used in personnel screening if doing so would reduce the proportion of minority applicants who are accepted for employment or education? There is no justification for using a test that would have an adverse impact on one group or another if that test is *not* a valid predictor of performance. If the test is a valid indicator – as intelligence tests usually are, to some degree – then the policy maker is faced with a trade-off. Should the best people be selected, regardless of group membership, or should some effort be made to balance rates of acceptance across racial and ethnic groups? This is a policy issue, not a scientific one. Scientific research can provide information about the costs and benefits of a policy, but the decision is up to the policy maker.

1.4. A Framework for Thinking about Intelligence

1.4.1. *Manifest and Latent Variables: Definition and Diagramming Conventions*

Many of the arguments over intelligence have been less informative than they might be, because the arguers have not been clear about what their concept of intelligence is, or about what they think the relation is between test scores and intelligence, in the more general sense of individual differences in mental competence. Here I present my own views. They represent an expansion of a model of the causes and consequences of mental competence that I developed together with Jerry Carlson, a professor at the University of California, Riverside.[22] We did not present a new theory of what intelligence is, nor do I here. The goal is to provide

22 Hunt & Carlson, 2007a.

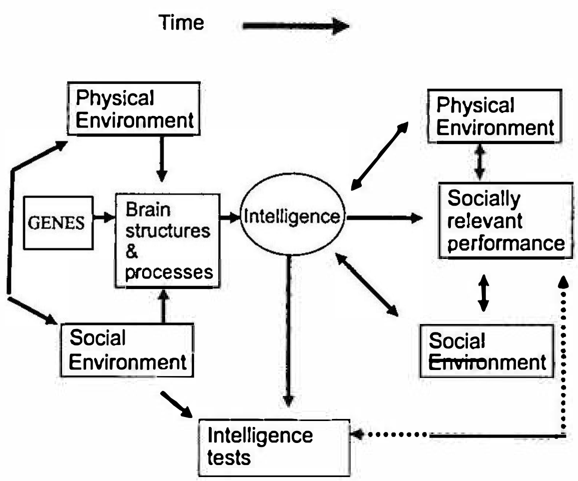

Figure 1.3. A model of the development and utilization of cognitive abilities ("Intelligence"). Intelligence is a latent variable, indicated by manifest variables that include but are not limited to intelligence test scores. Intelligence influences, and is influenced by, a number of environmental and biological variables.

a framework for discussing theories and facts about intelligence.

Figure 1.3 presents the model I will use throughout this book. Before discussing its content a word is in order about the notation used in the figure, for these notational conventions will be followed throughout this book.

Rectangles will be used to represent *manifest variables*, things that can be seen and, in principle, measured. Test scores, grades, and salaries are examples of manifest variables. Ellipses will represent *latent variables*, conceptual entities that are used in theories. Intelligence, cultural knowledge, and socioeconomic status (SES) are latent variables. They can be defined only by their (imperfect) manifest indicators. Latent variables are never defined or measured directly; their values are inferred from observation of the appropriate manifest variables. For example, SES is inferred from education, wealth, and residence. Intelligence is inferred from test scores and, in many cases, other indices of cognition.

We also have to distinguish among three separate relations between variables: causation, reciprocal causation, and correlation. Causation will be indicated by a single-headed arrow. For instance, in Figure 1.3 there is a single-headed arrow between genes and brain structure, because genetic makeup does determine brain structure. Reciprocal causation, in which one variable causes another and then the second feeds back into the first, will be indicated by a double-headed arrow. The double-headed arrow between intelligence and the social environment indicates that a person's intelligence partly determines his or her social environment, which in turn influences the further development of intelligence. Correlation, in which two variables tend to occur together, without any implication of causation, will be indicated by a double-headed arrow with a dashed line between the arrowheads. There is an example of this in Figure 1.3, and we will encounter examples of correlation without causation in other diagrams.

The debate between Lippmann and Boring can be cast in terms of manifest and latent variables. Lippmann thought of intelligence as a latent variable, but failed to grapple with the question of how it should be related to manifest variables. Boring either did not make the distinction, or wanted to assert that his manifest variable, the test score, was a perfect indicator of the latent variable. I do not think that either position can be maintained.

1.4.2. *The Causes of Intelligence*

Now let us turn to substance, working from "the beginning," at the left side of Figure 1.3, toward the expression of intelligence represented on the right-hand side.

It all starts with the genes. A person's genetic makeup determines his or her potential for the development of brain structures that support all activities, including cognition. There are individual differences in genetic makeup; otherwise we would all be clones. These differences do have implications for the development of intelligence, even though most people probably operate well below their genetic potential.

Although the genotype is established at conception, parts of the genotype may not be expressed until certain ages are reached.

For example, there are a number of medical conditions, such as Alzheimer's disease, that have some genetic basis, but are not displayed until well past childhood.

The extent to which the genetic potential is realized depends upon the extent to which the environment encourages the development of intelligence. In modern developed societies it appears that somewhere upwards of 50% of the variance in IQ scores can be statistically associated with genetic variation. However, no one, ever, inherited a test score in the same sense that they inherited the color of their eyes. People do differ in the extent to which they have inherited brain mechanisms that allow them to deal with their society in a way that produces the mental capabilities required to solve the problems proposed on a cognitive test. My apologies for a complicated sentence, but I cannot think of any other way to make this very important point. Genetic capacities unfold throughout life. In fact, the association of genetics with intelligence test scores is higher in old age than in adolescence.[23]

The extent to which the genetic potential is realized is determined by the physical and social environment. Here is a striking example.

If a pregnant woman abuses alcohol her child may be born with *fetal alcohol syndrome*, a serious form of mental retardation. The likelihood that this physical condition will occur depends on the social environment. Fetal alcoholism does not incur in societies that enforce abstinence. It can be a substantial problem in societies where alcohol is freely available as a recreational drug, especially if the social stresses that lead to alcohol abuse are present.

Social influences also act upon the developing mind. At this point, though, we have to shift from the concept of *brain* to the concept of *mind*. Ultimately everything is in the brain. Nevertheless, it is often useful to distinguish between mental capabilities that are determined by a person's capacity to process information in the abstract and capabilities that are determined by the possession of information, either about the world or about how to solve problems in the world. Information and problem-solving styles are heavily influenced by the society within which a person lives.

Arguably the most striking example is literacy. The ability to read appears to be associated with willingness to evaluate abstract arguments and to take multiple perspectives, a behavior that is tested in many intelligence tests.[24] These can be considered direct influences of literacy upon intelligence. Literacy also has an indirect influence, for it opens the door to formal education, and education is, by definition, a major avenue for passing on culturally acquired knowledge. Literacy facilitates the transmission of cultural knowledge, and cultural knowledge allows us to behave more intelligently.

Medicine offers an example. Modern health workers wash their hands when they move from patient to patient. Prior to the mid nineteenth century this was not seen as necessary. Does this make today's physicians more intelligent than the physicians of George Washington's day? In an important sense, they are. Knowledge is cognitive power.

There is a two-way interaction between brain and mind. One of the major limitations on knowledge acquisition is the ability to concentrate mental effort on a topic, in the face of distractions. The ability to concentrate depends on how well certain brain structures work, primarily but not exclusively in the forebrain. It also depends upon how efficiently information about the external world is coded inside the nervous system. The coding systems people use depend very much upon their experience with the topic at hand.

Intelligence is a personal trait produced by an interaction between genetic potential and environmental support. Where do IQ scores fit into this picture?

23 McGue et al., 1993.

24 Cole, 2005; Wolf, 2007.

1.4.3. *The Measurement of Intelligence*

The center portion of Figure 1.3 makes a point that is obvious when you think about it, but that is surprisingly often forgotten.

A person's intelligence test score is produced by two things: the social decision to construct an intelligence test in a particular way and the examinee's ability to deal with the test once it has been constructed. Different societies might construct different tests, depending upon the mental capabilities that each society values. This does not mean that tests are narrowly culture-bound, for some mental qualities are seen as vital in all societies. For instance, all societies demand that their members learn their native language. On the other hand, societies may differ in the emphasis that they place on other aspects of cognition.

Both these points were illustrated nicely by an anecdote told me by Manuel de Juan-Espinosa, a Spanish psychologist who has studied conceptions of intelligence held by the Fang, a society of mixed agriculturalists and hunters in Equatorial Guinea. When you ask people in Western society to list the attributes of an intelligent person, you generally get statements about the ability to solve abstract problems and the ability to comprehend and use language. Spatial orientation, the ability to locate oneself in the physical environment, is either not mentioned or mentioned far down on the list. When the topic is brought to their attention people do agree that it is intelligent to be able to find your way around the neighborhood. In our society, though, this is usually not an important skill. Our intelligence tests include only a few evaluations of spatial orientation.

When the Fang list the qualities of an intelligent person they say, "Intelligent people do not get lost in the forest." This does not mean that the Fang devalue the sorts of verbal skills that Westerners mention. In fact, they stress verbal skills, in much the same way that Westerners do. If the Fang were to construct intelligence tests they would include tests of verbal skills, as ours do. The Fang might look into the matter of spatial orientation in more detail than we do.

IQ tests, and tests related to IQ tests, such as the SAT and AFQT, are artifacts of the cultures in which they arose. They test some aspects of intelligence but not others. However, the tests are not arbitrary.

IQ tests would not have survived, as artifacts, unless test scores could be used as (imperfect) predictors of what our society sees as socially important behaviors, such as academic and social achievement. Because the test scores do meet this criterion, the tests must either evaluate mental skills that are used by the society or they must evaluate mental skills that are not used, in themselves, but whose possession is highly correlated with the possession of skills that can be used. That is, the tests might be like an Army physical examination where the candidate is required to do push-ups. Soldiers are not going to do push-ups in combat, but the ability to do push-ups is correlated with the ability to move heavy objects (e.g., carrying artillery shells to a gun position), which soldiers may have to do. The same argument may apply to the mental gymnastics required to perform well on intelligence tests.

While all human societies are not identical, there is a core set of cognitive skills that all societies rely upon. All societies demand that their members learn to speak the native language, be able to control attention, and, by the standards of all other animals in the world, remember events very well. Therefore, an intelligence test that is valid in one culture is unlikely to be entirely invalid in another, although its validity may be reduced.

Individual differences in cognitive skills might, in principle, either be due to possession of a great many special-purpose brain mechanisms or be due to possession of very general information-processing capabilities that can be applied to all mental challenges. Intermediate solutions are possible; we might have a single general processing capacity, augmented by special processors. To the extent that the evolution of our species has produced a general problem-solving brain, it does not matter precisely how that brain is evaluated, for a person's behavior in one cognitively challenging situation will predict how well he or she deals with other situations.

There is a surprising amount of evidence that evolution has taken the general problem-solving solution.

1.4.4. *The Uses of Intelligence*

We now come to the right-hand section of the diagram, dealing with the results of intelligence.

People use their cognitive abilities to define their environments. Consider good health practices. Sometimes very bright people become alcoholics and/or crash cars, both activities that can lead to brain damage. But, on the average, people with high intelligence test scores do not do such things, and so enjoy better health.[25] Once again, we see how intelligence and the environment interact.

Cognitive abilities are not the only abilities that we have. Differences in behavior are also produced by individual differences in a variety of emotional-motivational traits. These are lumped together loosely under the title "personality." It is not clear where to draw the line between intelligence and personality. For instance, *conscientiousness*, the tendency to fulfill obligations to others (including one's employers) is usually considered a personality trait. However, it is possible to see conscientiousness as an offshoot of intelligence. It makes sense to be conscientious in fulfilling obligations to people who control resources that you want.

This is hardly a new observation. Machiavelli's famous sixteenth-century discourse on political behavior, *The Prince*, contained cogent arguments for displaying certain personality traits, including honesty and conscientiousness, because such behavior is in one's enlightened self-interest. It is intelligent to be thought trustworthy and reliable.

We can make a distinction between cognition and personality by asking if we are talking about whether a person *can do* or *will do* a certain behavior. To the extent that the answer is "can do" the mental acts controlling the behavior are part of cognition, and hence individual differences in them are part of intelligence. To the extent that the answer is "will do" the mental control is part of motivation, and individual differences are part of personality.

Any particular action has to satisfy both "can do" and "will do" requirements. Although this is hardly a profound statement, it is surprising how often explanations of behavior focus on personality to the exclusion of intelligence or vice versa.

1.4.5. *The Results of Intelligence*

Intelligence shows itself in two ways: by test scores and by socially relevant behaviors. Test scores are easy to analyze; socially relevant behavior is hard to analyze. Nevertheless, socially relevant behaviors are far more important than test scores.

There are statistical associations (correlations) between intelligence test scores and measures of socially relevant behavior including academic achievement, income, health, and occupation of prestigious positions in society.[26] These associations are facts. Arguing about why they occur is a reasonable thing to do. As every student in elementary courses on experimental design is told, correlation does not imply causation. Rather, the fact that intelligence test scores and measures of socially relevant behaviors are correlated suggests a number of possible causes, all of which are worth investigation.

Performance on an intelligence test and performance in socially relevant situations might both depend upon the same cognitive processes. This is what most researchers on intelligence believe. However, the statistical associations, although not negligible, are not compellingly large. They generally range from a correlation coefficient of .2 to

25 For a typical study recent study, see Batty et al., 2006.

26 Gottfredson, 1997; Herrnstein & Murray, 1994. See the extensive discussion in Chapter 10.

.5, and in a few situations correlations as high as .8 have been reported. Just what these statistics mean is explained in more detail in Chapter 2. Temporarily, I ask readers to take on faith the statement that the correlations are high enough to indicate that test scores and behaviors tap some common traits, but low enough to indicate that the traits that affect the tests and the behaviors are not exactly the same.

It could be that the test scores themselves are causing the behaviors but that the cognitive capabilities required by the tests are not. This might happen in two ways. One is that knowledge of a high or low test score could cause other people to treat the examinee differentially. It has been claimed that this is the case in education, because teachers may respond differently to students with high or low IQ scores.[27]

Test scores will influence behavior if the scores are used to decide who is offered a chance to behave in a certain way. The SAT, the ASVAB, and the Department of Defense's Officers Qualifying Test (OQT) are examples. Other things being equal, the people with the highest test scores will be allowed to enter a prestigious university, be commissioned as an officer in the armed services, and so forth. Subsequent behavior will be guided by the experience of having been in the university, been a commissioned officer, and so forth. On the one hand, it can be argued that this is appropriate. If a test correctly identifies people who have the cognitive capability to, say, benefit intellectually from an Ivy League education, then the test is an appropriate gatekeeper. On the other hand, it could be argued that the benefits one gets from being allowed to enter a certain social stratum depend very little on cognitive abilities. Colloquially, it's not what you learn

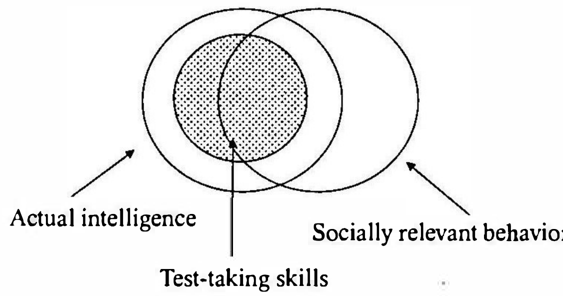

Figure 1.4. People possess cognitive skills (conceptual intelligence) to varying degrees. Some of these skills are evaluated by the tests. Of the skills evaluated by the tests some are specific to the test situation, and others are relevant to social behavior in general.

at Harvard (Yale, Stanford, Duke, Oxford, or Cambridge) that determines your subsequent success in life, it is who you meet in college. Either of these mechanisms could produce a correlation between test scores and later success, on a between-institution basis, because the test scores were used to determine who got in, but not on a within-institution basis, because the test scores did not tap the abilities required to benefit from the experience.

The two explanations are not mutually exclusive. My own bet is that both factors operate.

Figure 1.4 summarizes the argument. Socially relevant behaviors are partly determined by cognitive skills. Variations in these behaviors are, conceptually, intelligence. Some of the skills that define conceptual intelligence are reflected in test scores. Other socially important cognitive skills that are part of intelligence are not reflected in test scores, and test scores also reflect test-specific skills that are not related to socially relevant behavior.

Lippmann was right that IQ scores do not measure everything, Boring was right that they do measure something.

Intelligence tests do a good job of evaluating an examinee's ability to respond rapidly to problems of varying degrees of complexity. In the following chapters I discuss evidence showing that IQ and similar tests are statistically related to the cognitive abilities that underlie socially relevant behaviors, ranging literally from becoming a judge to

27 Should teachers know their students' IQ scores? There are rational arguments for and against the practice. If teachers know students' IQ scores they might offer special instruction to students they think are bright, or they might give up on teaching low-scoring students. Looked at another way, the information could be used diagnostically, helping teachers tailor instruction to the student.

becoming a criminal. However, important cognitive abilities required for socially relevant behavior are not tapped by the tests. What might these be?

Any testing method that relies on the conventional "Drop in from the Sky" paradigm cannot evaluate abilities that reveal themselves over a relatively long period of time, such as the ability to plan, to allocate time for extended courses of action, or to integrate information from multiple sources. It is difficult, if not impossible, to tap these skills in a test that seldom takes more than three hours and that consists of unrelated problems that, individually, take only a few minutes to solve. Overreliance on conventional testing has greatly limited modern research on intelligence.

1.4.6. *Cause and Effect in Intelligence Research*

The arrows going back and forth in Figure 1.3 highlight how hard it is to determine cause and effect when studying intelligence. If everything were simple we would place causal variables on the left, intelligence in the center, and the effects of having intelligence on the right. To some extent this can be done. Genetics and a child's physical and social environment do produce intelligence, and a person's intelligence does produce socially relevant behavior, and thus alters the person's environment. The problem is with the feedback. Intelligence influences a person's environment, and feedback from the environment alters intelligence.

Take the case of aging. As we grow older two different processes influence intelligence. There is a decline in brain function. At the same time, as people live they acquire "wisdom," better and better knowledge about how the culture runs. There are huge individual differences in both of these processes. Some people remain cognitively fit until great age; others descend into near-senility at little past fifty. Some people acquire wisdom as they pass through the world; others only have experiences. In both these situations intelligence appears to act as both cause and effect. Other things being equal, more intelligent adults are more likely to maintain healthy lifestyles than less intelligent ones, and are more open to engaging with, and thus extracting wisdom from, the world about them.

The interaction between intelligence and the environment has posed a major problem for researchers. We know far less than we need to know about the development of cognitive power over the adult life span. It is reasonably easy to measure cognition and other psychological variables as long as people are in the educational system. Similarly, it is relatively easy to obtain access to retired people, through social and medical support institutions, ranging from gray-haired hikers pausing at an elder hostel to the patients in a medical care facility. It is much harder to obtain access to people in the working years, for the simple reason that they are busy at work and raising families.

This is a serious situation, for adult intelligence is not inert. It rises to meet environmental challenges.

1.5. Reaction Ranges and the Challenge Hypothesis

Imagine a hypothetical person, Harry P., at age fifty. We want to estimate Harry's intelligence without actually measuring it. Here is what we can do.

1. Our initial estimate is the average intelligence of a person at age fifty.
2. Collect facts about Harry's physiological status, including genetic background, medical history, present eating and drinking habits, and so on. Use these to compute a biological correction factor to the initial estimate.
3. Collect facts about Harry's social environment, including education, marital status, hobbies, profession, and so forth. Compute a social correction factor, and add it in. The twice-adjusted figure is our estimate.

The result will be an estimate of current intelligence. We would not know how that

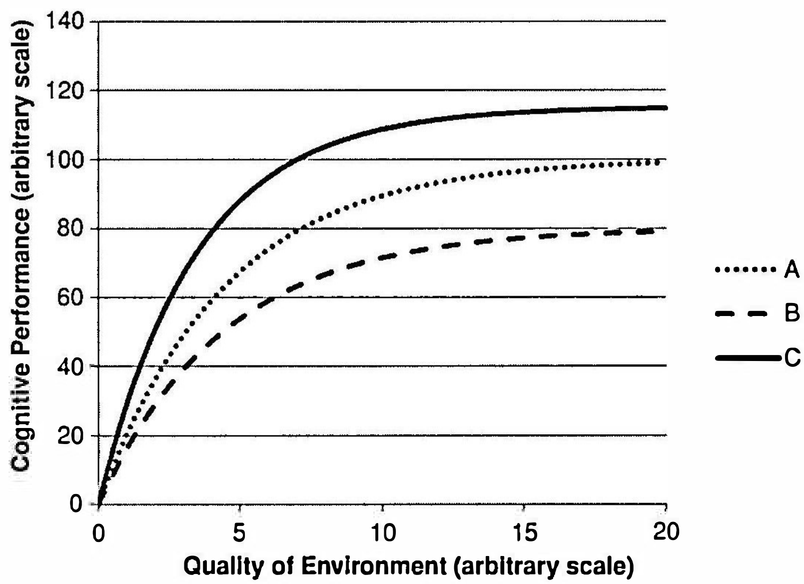

Figure 1.5. The concept of reaction range. The ordinate represents the quality of observable performance, and the abcissa represents the extent to which the environment supports development of cognitive skills. A person's level of cognitive performance is determined by the combination of reaction range and quality of the environment. Persons A, B, and C each have a unique potential for cognitive performance, indicated by the three lines. Their actual performance will be determined by the quality of the environment. If the environment goes from very poor (at the far left of the figure) to moderate (toward the middle), there will be considerable improvement of performance. Going from moderate to good environments (towards the right) does not result in a great deal of improvement.

intelligence had been acquired. Estimation alone does not explain the dynamics of intelligence.

A person's genetic makeup does not provide that person with a certain number of "intelligence units," any more than it provides a person with a certain number of points on an IQ test. Genetic status provides a *reaction range*, a range of levels of intelligence. Environmental factors determine where the person will operate within that range.

Figure 1.5 displays three hypothetical reaction ranges. Cognitive functioning is shown on the ordinate (y-axis), the level of favorableness of the environment on the abscissa (x-axis). Three different reaction ranges are shown. Where a person actually functions is determined by the point at which a vertical line drawn from the environment's rating crosses the reaction range line. Observable cognitive performance is determined by the combination of genetic reaction range and environmental quality. The level of genetic inheritance cannot be inferred from IQ scores unless environmental quality is known, nor can environmental quality be inferred from IQ scores unless genetic inheritance is known.

Varieties of this theme will appear throughout our discussion of intelligence. For example, any difference in the cognitive performance of identical (*monozygotic* or *MZ*) twins can be attributed to the environment because MZ twins, having identical genotypes, will have identical genetically determined reaction ranges.

The concept of reaction range applies to the environment as well. This can be seen in Figure 1.6, which is simply a "cleaned up" version of Figure 1.5, in which three vertical lines have been drawn upward from

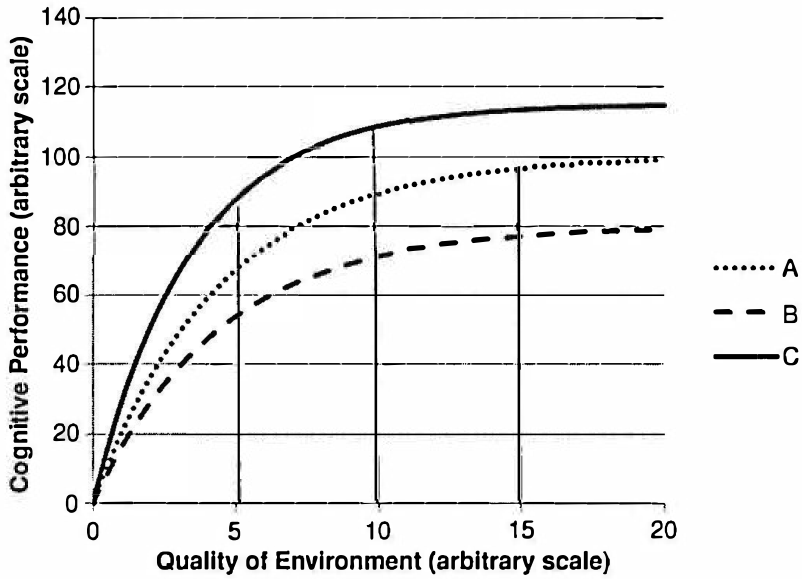

Figure 1.6. Environmental reaction ranges. Suppose the environmental quality is fixed (arbitrarily) at a quality of 10 for all individuals. Differences in cognitive performance will occur due to individual differences in genetic potential. However, if individual A is placed in an environment of quality 15 and C in an environment of quality 5, A will outperform C even though C has greater genetic potential.

the environmental quality axis, each line representing a different quality of environment. Each of the lines intersects the different genetically determined reaction ranges at different points, resulting in a variety of cognitive performances, even though the environment is identical for all individuals.

Figures 1.5 and 1.6 show the genetic reaction ranges as negatively accelerated curves that rise steeply at first, and then flatten out as they approach an asymptotic value. This was an arbitrary choice, for both the cognitive performance and environmental quality axes are shown on arbitrary scales. I chose to display negatively accelerated curves because of my conjecture that, in fact, attempts to improve cognition will result in negatively accelerated curves. Why do I make this conjecture?

There are a number of ways of producing physical environments that greatly constrain the development of intelligence. Prolonged famine is one; infection of the brain is another. Once these disaster states are avoided, the incremental effects of improving the physical environment are probably rather small. The same thing is true for the establishment of genetic potential. There are many "catastrophic" genetic conditions, where the presence of a single anomaly greatly restricts cognitive development, but if there is no anomaly genetic potential seems to depend upon the combined effects of a large number of genes, no one of which contributes a great deal. These matters are discussed in some detail in Chapters 8 and 9. There are lots of ways to restrict intelligence, but we know of relatively few ways to expand it. This implies negatively accelerated growth curves.

Having a reaction range does not mean that you use it. Cognitive skills are acquired by investing time and energy. Willingness to invest depends upon one's perception of the likely outcome of the investment. I propose the following conjecture.

The Challenge Hypothesis:

Intelligence is developed by engaging in cognitively challenging activities. Environments vary in the extent to which they support such challenges, and individuals vary in the extent to which they seek them out.

The development of intelligence depends upon the extent to which the individual wishes, is allowed to, or is required to meet environmental challenges.

Consider one of the linchpins of human cognition – language. Every human being is required to learn a spoken first language, and all normal people do so. Only in the last two hundred years have societies begun to demand literacy, which is a much harder skill to acquire. It is worth the effort. A literate person has acquired a cognitive skill that makes him or her more intelligent than an illiterate one with an identical, but unrealized, biological potential for intelligence.[28]

The example expands. Literate societies dominate the globe. They have developed formal education systems that force further intellectual development. Mechanisms of education such as the school, the newspaper, and the World Wide Web are different from, and far more challenging than, the educational systems of the nineteenth century. As a result, today's children are more intelligent than children in the past. As a reflection of this phenomenon, IQ scores rose throughout the twentieth century.

Literacy is an example of a compulsory challenge; individuals have to meet it or suffer severe consequences. In other cases there may be options. Learning about statistics and probability provides a good example. On a population basis, relatively few people study statistics. Students who graduate from, and understand, elementary statistics courses can solve problems that Pascal and Gauss could not solve. I do not think that the typical student in a modern statistics class has the biological potential for intelligence that Pascal and Gauss had. There is a sense in which they are more intelligent, for they are more powerful problem solvers.

Robert Sternberg has identified three ways a person has of responding to cognitive challenges.

1. *Adapting*: Changing your own cognitive behaviors to meet the challenge. Studying is a way of adapting, though certainly not the only one.
2. *Shaping*: Changing the environment to adapt it to your current capabilities. If you have difficulty doing arithmetic, buy a hand calculator.
3. *Selecting*: Finding a new environment that does not present the challenge you want to avoid. If you are a college student majoring in engineering, but math classes are difficult for you, consider switching majors.

Sternberg's three strategies have different cognitive demands of their own. Both adapting and shaping require a willingness to engage with the environment. This is in itself a reliable individual trait. Philip Ackerman, a professor at the Georgia Institute of Technology, has conducted research showing that willingness to engage in intellectual challenge is characteristic of the people who hold the occupations and avocations that we think of as requiring intelligence.[29] Selecting can be rational in some situations, but it has the danger of becoming a way to avoid intellectual challenge, and hence, to avoid the development of one's biological potential for intelligence.

1.6. Intelligence Is Part of a System

Defining intelligence solely in terms of test performance is an impoverished view. It focuses our attention on explaining variations in test scores, which are not important in themselves, at the expense of studying individual variations in socially relevant behavior, which are important. However, it would be foolish to disregard the considerable amount of information about intelligence that is incorporated in test scores.

One of the major reasons for studying intelligence is to understand how individual differences in cognitive competence, intelligence in the broad sense, are related to individual differences in the display of socially

28 See Wolf (2007) for a very well-argued, extensive expansion on this point.

29 Ackerman, 1996; Ackerman & Beier, 2005.

relevant behaviors. In practice, there are two problems. Variations in socially relevant performance are determined by noncognitive as well as cognitive factors. Success (or failure) is based on both "can do" and "will do." In addition, the display of socially relevant behaviors depends upon the opportunity to display them, which may be quite beyond a person's control, no matter what his or her personal characteristics are. Here are two historical examples.

In the early twentieth century Joseph P. Kennedy, an American financier, amassed a considerable fortune. Subsequently three of his sons – John, Robert, and Edward – became United States Senators. In 1960 John became President of the United States. One can argue, probably correctly, that Joseph Kennedy provided his sons the genetic capability to become very intelligent. But his fortune certainly helped their political careers, if only to relieve them of the need to earn a living outside of public service.

Now let us look at an example of restriction. The Declaration of Independence of the United States, written in 1776, contains the statement, "All men are created equal." At that time women were disenfranchised, and the slavery of Africans was condoned. Over two hundred years later Condoleeza Rice, an African American woman, who had served in many high-level positions, including Secretary of State, observed that when the Declaration was written, "They weren't talking about me." Dr. Rice's impressive achievements, which depended very largely on her intelligence, would have been impossible in 1776.

At the start of the twenty-first century the Darfur region of the Sudan was wracked by drought, famine, and vicious ethnic warfare. Children born in Darfur in the year 2003, no matter what their genetic potential, did not have good life prospects.

The point of these examples is that both the causes and effects of intelligence are embedded in a matrix of other variables. This makes good research on intelligence hard to do. The task is difficult, but it is not impossible. While an ideal study may be impossible to implement, a great deal can be learned from less-than-ideal studies. Progress can be made by investigating the issues that can be studied, and hopefully by holding down excessive interpretations and conclusions where we do not have the right evidence.

It is fairly easy to determine the correlation between test scores and other measures of interest, such as grade point average (GPA) and performance at work. Such studies provide important data. However, they bias our knowledge toward finding out the role of intelligence in certain institutions, such as the schools and the military, that, while certainly important, are not the whole of society. In addition, the study of bivariate correlations, in isolation, fails to stress an important fact. Intelligence is just one of the variables in the system defined by human society. What does this mean?

A *system* is a set of interdependent variables, in which each variable influences the others. In *closed systems* the interdependence is complete. The value of each variable is completely determined by the other variables in the system. In *open systems* some variables are influenced by conditions outside the system. The real-world systems we study are always open. Therefore, it is important to distinguish between *system variables*, which exert measured reciprocal influences on each other, and *external variables*, which influence the system variables but are not (to any great extent) influenced by the system variables.[30]

To illustrate, consider a hypothetical study of the roles of genetics, family influence, and intelligence during primary school, middle school, and high school. The following relations hold:

1. A child's intelligence on the first day of school has been determined by genetic inheritance and family environment prior to entering school.
2. Intelligence on entering middle school is determined by intelligence on entering

30 In economics the terms *endogenous* and *exogenous* variables are used.

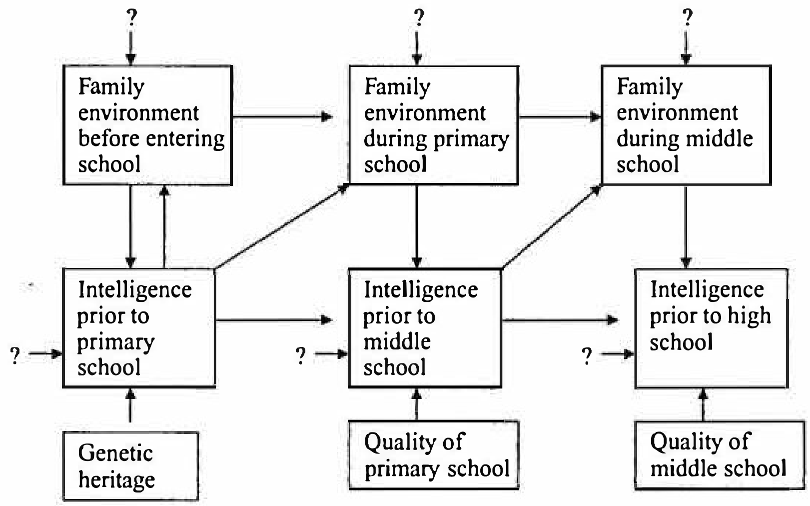

Figure 1.7. A systems' view of the relationship between intelligence, family environment, and the quality of schooling. Intelligence and family environment are system variables. Genetic heritage and quality of school are external variables. Unknown sources that influence the system variables are indicated by a question mark.

primary school, the quality of the primary school, and family environment during primary school, but not by the environment prior to entry into primary school.

3. A child's intelligence on entering high school is determined by intelligence on entering middle school, family environment during middle school, and the quality of the middle school.
4. At any time family environment is determined by prior family environment and by the child's intelligence. (People, including children, influence their environment.)

Figure 1.7 diagrams the system. Intelligence and family environment are system variables, because they influence each other. Genetic potential and quality of schooling are external variables, because they exert influences on other measures but are not influenced by them. If we can obtain measurements of all these variables, we can use modern statistical methods to evaluate the relative influences of each variable upon the others.

Then we come to the external, unmeasured variables. Figure 1.7 makes no provision for extrafamilial influences on the family environment, such as financial emergencies. Nor is there any provision for extrafamilial and extra-educational influences on intelligence, such as physical injury. Therefore, we have to allow for the "unknown unknowns," the influence of unmeasured external variables. These are indicated by the "?" symbols in the figure. While we cannot identify these variables, modern statistical techniques do allow us to estimate the size of their influence compared to the influence of the measured system variables on each other. Hopefully the influence of the "unknown unknowns" will be small; if it is large, analyses within the system will account for only a small part of what we need to know.

We can learn a great deal by comparing systems models to each other. The intelligence-education system displayed in Figure 1.7 treats genetic influences as a one-shot effect – genetics influences intelligence prior to entering school – but has no direct influence subsequently. In fact, though, some genetic effects unfold over time. For example, individual differences in the rate at which connections are developed in the forebrain during adolescence may result in differences in the ability to control impulsive behavior, which may influence how much a student learns in middle school, and so forth. These influences could be modeled by extending arrows from the genetic inheritance box to the boxes representing measurements of intelligence at each time period, not just at the point of school entrance. The extension is needed only if the added arrows produce a system that (reliably) explains more of the variation in intelligence than did the simpler system.

System analyses of this sort allow researchers to move far beyond arguments over the meaning of a correlation between IQ scores and just one other variable, in isolation from the system in which both occur.

Still, problems remain. Systems incorporating human intelligence are so complicated that no one study can ever include adequate measures of all relevant variables. The size of the influence of unknown variables can be evaluated, but what those variables are and how they exert their influence will remain unknown. Therefore, we cannot explain all the causes and ramifications of intelligence, all at once. We can identify and analyze reasonably closed subsystems, dealing with a particular aspect of intelligence. No one such study will tell us all about intelligence, but, taken together, they will tell us quite a lot.

1.7. Summary and Prospectus

Individual differences in cognitive capacities – *intelligence*, for short – are an important part of human variation. Intelligence is partially tapped by intelligence tests (IQ) and by other tests of cognitive achievement, but other important cognitive abilities lie outside of the tested realm. While intelligence does not completely determine success in life (or lack of the same), it certainly contributes to it. In order to get a complete picture we have to consider personality and motivational factors, and the extent to which the social and physical environments encourage some behaviors and discourage others.

Intelligence has both multiple causes and multiple consequences. In order to study intelligence it is necessary to isolate relatively closed systems of variables. This poses a challenge, because our ability to measure some variables, and hence to study systems involving them, is much greater than our ability to measure other variables. Test scores and measures of genetic variation are much easier to obtain than measures of success in society, or measures of variation in the physical and social environment. Therefore, we have to be vigilant against the error of studying what is easy to analyze, at the expense of missing amorphous but important effects. This tension will be reflected throughout the discussions in this book.

Some typical tests will be described in Chapter 2. Subsequent chapters deal with the description, causes, and consequences of intelligence. The book closes with a discussion of how intelligence is distributed across our society and, finally, some speculations about the development of research on intelligence in the future.