ZIMBARDO: Can tests really tell us how smart we are, how well adjusted we are, and what goes on in our subconscious minds, or are they unfair barriers, narrowing our definitions of intelligence, competence, and mental health?

"Testing and Intelligence" this time on Discovering Psychology.

>> ZIMBARDO: Let me ask you a personal question.

How smart are you?

One way you could answer is to list your accomplishments, your grades, your job, your possessions.

Or you might point out all the things you're good at -- anything from solving crossword puzzles to inventing a better mousetrap.

Or you could give me just one number, your IQ.

If you say your IQ is 100, you've told me that you're average in certain aspects of intelligence compared to most people.

That means half the population for which the test was developed will have higher scores than yours and half will have lower.

But can the richness and complexity of intelligence really be reduced to a number?

And why do psychologists spend so much time and effort trying to get this number and others like it?

And perhaps, most importantly, what effect can such numbers have on your life?

Today we're going to try to answer these questions by focusing on the field of psychological assessment.

Its goal: to find out how people differ in their abilities, behaviors, and personalities.
Using a variety of testing methods, psychologists try to put a tape measure to the mind.

How people differ has long been an object of fascination, but psychologists have taken the measurement of these differences as their special province.

Their tool is psychometrics, mental testing in all its varieties.

While the test is being administered...

>> ZIMBARDO: The most well-known tests are those measuring some aspects of intelligence.

And then there are tests like the SAT, which measures your developed academic abilities.

And there are tests that measure your learning in specific subject areas.

Together, all these tests are called cognitive tests because they measure various aspects of your mental ability.

There are also personality tests, which try to measure the noncognitive parts of human personality: your interests, values, and personality traits.

Projective tests like the Rorschach help some clinical psychologists assess mental and emotional problems.

What you see in ambiguous inkblot patterns is then interpreted as reflecting your innermost thoughts.

While in the Minnesota Multiphasic Personality Inventory Test, you're asked to answer true or false to hundreds of statements about different attitudes and beliefs.

Your responses are then used to derive a set of personality characteristics and to assess whether they're more like the characteristics of people who are deemed normal or abnormal.

For all these tests, the bottom line is to gather a sample of responses conveniently and inexpensively and then to use these responses to predict with a minimum of error an individual's future performance in some real-life situation.

In fact, testing has become so convenient and inexpensive
and so pervasive, it's hard to imagine a time when someone wasn't being tested.

30 01:32:58:20 If we eliminate God's famous apple test of Adam and Eve's willpower, then the earliest recorded system of assessment was probably established by the emperors of ancient China.

31 01:33:11:06 As early as 2200 B.C., a civil-service testing program was instituted for government officials.

32 01:33:19:17 In modern times, Sir Francis Galton of England was one of the key figures in the development of mental measurements.

33 01:33:26:14 At the turn of the century, Galton tried to demonstrate the hereditary basis of intelligence.

34 01:33:32:12 His work triggered a long tradition of controversy over the relative contributions of heredity and environment to intelligence.

35 01:33:42:23 The other great early figure in assessment was a Frenchman named Alfred Binet, who published the first workable intelligence test in 1905.

36 01:33:53:11 Binet's test was devised to measure intellectual performance of school children so that teachers could classify and separate retarded children from normal ones.

37 01:34:03:11 The test was meant to replace the teachers' more subjective and often biased evaluations.

38 01:34:09:10 Binet believed that his test scores were only a practical estimate of current performance, not a measure of innate intelligence.

39 01:34:18:29 He believed that test results should only be used to help identify those areas in which children needed special help.

40 01:34:26:08 But the educators who administered his tests rarely made these distinctions, as we'll see later.

41 01:34:32:18 The key to Binet's approach was the quantification of a student's performance.

42 01:34:38:20 First, a great number of children of various ages were tested and an average score, or norm, was obtained for normal children in each age group.
Then an individual child's performance was compared to the average score for his or her age.

Test results were expressed in terms of mental age.

When a child's score was average for a group of five-year-olds, then the child was said to have a mental age of five, regardless of his or her actual chronological age.

Nowhere did Binet's test have a greater impact than in the United States.

Millions of people were streaming in from Europe and Asia and a vast public school system was emerging.

How could the mental abilities of these newcomers be assessed?

And there was also a flood of recruits enlisting in the army during the First World War.

They, too, had to be tested for intelligence and aptitude.

Assessment was being seen as a way to bring order to a chaotic society -- an inexpensive, democratic way of separating out those who could benefit from education or military training from those who were deemed incompetent.

Eventually, there was a growing acceptance of the idea that intelligence tests could be used to differentiate people in terms of their leadership ability and other socially important characteristics.

And there was also growing attention given to alleged intellectual differences according to race and ethnic background.

Statistics were used to support bogus arguments that blacks were inferior, as well as immigrants, like my grandparents and perhaps yours.

In 1916, Louis Terman of Stanford University adapted Binet's test for American school children.

The new version was commonly known as the Stanford Binet Intelligence Test.

With it, Terman would provide the foundation for a new
concept -- the intelligence quotient, or IQ.

58 01:36:48:13 This quotient, as it was measured then, was the ratio of mental age to chronological age multiplied by 100 to eliminate decimals.

59 01:36:59:23 Thus, a child with a chronological age of eight whose test score equaled that of a ten-year-old would have an IQ of 125.

60 01:37:10:17 The new Stanford Binet test soon became a standard instrument in clinical psychology, psychiatry, and educational counseling.

61 01:37:18:10 And more importantly, Terman and others came to believe, unlike Binet, that intelligence was an inner quality, that it had a large hereditary component and that so-called IQ tests could measure this inner quality.

62 01:37:32:03 The implicit message was that IQ characterizes something essential and unchanging about human nature -- our inherited intelligence.

63 01:37:42:28 The next milestone in intelligence testing was the work of this man, psychologist David Wexler.

64 01:37:50:04 In 1939, Wexler designed a new test to solve a major problem of intelligence tests in general, their dependence on language -- a problem for non-English speakers and small children.

65 01:38:03:10 So he developed new, nonverbal sections to go along with the verbal sections.

66 01:38:08:15 In this example, test takers are asked to put these pictures in a logical sequence.

67 01:38:16:20 Since the development of the original Wexler intelligence scales in 1939, later revisions have become the most frequently used assessment tool in the country.

68 01:38:27:03 More than three million people a year take it.

69 01:38:31:04 Psychological testing has become a big business in America.

70 01:38:34:29 There are thousands of psychologists whose main work revolves around using them.
But how do we know if they do what they're supposed to do?

A good place to start is to ask what is it that makes a test valuable to a tester?

Here at Howard University, the late William Curtis Banks looked at potential misuse and racial bias in aptitude testing.

>> A test first and foremost has to be valid.

The test has to be able to be demonstrated to measure what it is that you're interested in.

If it is intended to measure the ability of a young person to do well in college, then the test should yield a score which actually predicts how well the person does in college.

If the test is intended to predict how well an individual can perform a job, let's say of a police lieutenant, then the test score should actually predict and should be consistent with the individual's ability to actually perform that job when he is in it.

Secondly, the test needs to be reliable, because in order to be valid, it needs to measure whatever it is we're interested in in the same way each time it's used.

And three, we need to be careful in standardizing the test to make sure that the people that we use to establish its initial validity and reliability are the same people that we're going to use to test with to measure later on.

>> ZIMBARDO: So not only do the tests themselves have to be assessed, we also have to assess how they're used or misused.

Alfred Binet's dream, which is still shared by many, was that psychological assessment would replace subjective, biased evaluations with objective ones.

In the United States, testing has been championed as a democratic tool to eliminate arbitrary and biased evaluations of people based on such irrelevant criteria as race, sex, nationality, social class, and physical appearance.

But these lofty goals are often lost when test results become not the keys to the kingdom, but a passport to nowhere.
There is no area of psychology more enmeshed in controversy than that of assessment.

The way that competency tests have been used in a biased way in the past, and currently in some instances, is by using the results of them to tell us that certain people are not really capable of doing something when they actually are.

The reason why the tests tell us that sometimes is because they have been developed in such a way as not to take full account of the different ways, in some instances, in which different people do the same job.

Many educators also worry about the problem of teaching for tests, instead of for basic knowledge.

It's not that we're concerned if people are able to do well at the test.

What we're concerned about is if they're able to do well at the test, but they're not able to do well at anything else.

And if they're coached specifically for the test, regardless of whether we're talking about a test in the area of vocations or in the area of education, then we end up with a measure on the test that tells us not what we really want to know, and that is can this person do something in the rest of the world, outside of the test situation?

Think about what happens when a test score begins to define a person.

"I'm a C student with a 100 IQ" -- as if it's indelibly stamped on the forehead.

Does that kind of labeling then become a self-fulfilling prophecy?

Or can generally held beliefs about a person's gender, race, or age influence a person's performance?

Psychologist Claude Steele of Stanford University studies the effect labeling and stereotypes can have on a person's ability to perform in a given situation.

He calls this phenomenon stereotype threat.
Stereotype threat just... it is a situational threat.

It's usually tied to the social aspects of the situation you're in and it's just the sense that you could be judged in terms of some stereotype about an identity that you have -- your racial identity, your age, your sex.

>> ZIMBARDO: We don't have to believe in a stereotype.

Merely knowing that others could judge us because of our social identity is enough to distort our performance.

One of the pivotal studies Steele conducted was testing women's performance when compared to men on standardized math exams.

>> There is in our society a stereotype that women have a limited ability to perform especially difficult math.

And so a woman taking a difficult standardized exam in math, when she experiences frustration, that stereotype becomes relevant as a possible interpretation of the frustration both for her and for anybody who is going to see or interpret her performance.

That makes the experience of taking the test different for her than it does for a man in precisely the same situation.

>> Hi.

Are you two here for the psych study?

>> Our interest there was to see, if you brought women and men who were both very skilled in math and very dedicated to being good at math into the laboratory and you gave them a very difficult math test, how would they perform?

Our prediction was that the women would not perform as well as the men because they would be under stereotype threat; the men wouldn't be in this situation.

And that's exactly what happened.

Our prediction is that if you could turn the spotlight of stereotype threat off, their performance should go up to match that of equally skilled men in this situation.

We take the threat out of the situation in a very simple,
straitforward way.

112 01:44:31:28 Just before people get ready to take the test, we sit them down and say, "On this particular test, this is not a test that's capable of discriminating between men and women's performance.

113 01:44:42:03 It never has.

114 01:44:43:09 It never will."

115 01:44:44:19 The stereotype is now almost magically made irrelevant to her performance here.

116 01:44:49:29 And when we create that situation, women's performance goes up dramatically on this test to match that of equally skilled men in this situation.

117 01:44:59:27 So that's the kind of evidence we use to give us some sense of how this process, this stereotype-threat phenomenon -- just being subject to being seen that way -- can actually affect a very important performance in a person's life.

118 01:45:16:28 >> ZIMBARDO: In the view of Alfred Binet, intelligence is a dynamic, cognitive process.

119 01:45:23:01 But unfortunately, many of the psychometricians who've followed have seen intelligence as a single trait which is accurately measured by intelligence tests and which can predict certain kinds of academic and vocational performance.

120 01:45:40:00 In the last few years, however, cognitive scientists like Howard Gardner of Harvard have given us new perspectives on intelligence.

121 01:45:50:10 >> The theory arose because of a dissatisfaction on my part with the view of intelligence as being just one thing which you could measure with an IQ test.

122 01:45:58:22 And if you look around the world and you think about all the different kinds of things that people can do well that are valued in different cultures, one seems pretty simple minded.

123 01:46:06:05 And after a lot of analysis, I came up with a list of seven intelligences.

124 01:46:10:25 The first one is linguistic intelligence.
It's a kind of ability that a poet or writer or orator would have.

Second intelligence I call logical mathematical.

As the name implies, it's what you'd find in a mathematician, logician, scientist.

And it's important to mention language and logic first because those are the intelligences which usually are featured in school and in IQ tests.

And if you have language and logic, you're in pretty good shape scholastically.

Once you leave school, you might find that you could use some other ones instead.

And that's where the other five intelligences come in.

There's musical intelligence which is what you would find in a composer, performer, someone who can think musically.

Spatial intelligence: the ability to think about the whole world or about the world close to you in terms of its spatial aspects.

I talk about a kind of intelligence which I call bodily kinesthetic.

That's using your whole body or parts of your body, like your hand or your mouth, to solve a problem or to make something.

Dancer, athlete, surgeon, again, craftsperson would have a high degree of bodily kinesthetic intelligence.

And I talk about two kinds of personal intelligence.

Interpersonal is understanding of other people, what makes them tick, the kind of ability you would find in a teacher, a clinician, therapist, politician, actor.

And the final kind, which is very important, but difficult to talk about, is intra, or within a person, intelligence.

It means having a viable understanding of yourself, who you are, what kinds of abilities you have, what kinds of needs you have, what kinds of intelligences you have, how to use
those effectively to solve problems or to make something.

And I think if you take those seven into account and the particular combinations among them, you can account for most of the kinds of thinking abilities which people need to survive all around the world.

We now understand from psychology and from cognitive science that people think differently.

We have different strategies of learning.

We have different ways of representing things in our head, different ways of acquiring knowledge.

Many people feel that psychology's greatest contribution, greatest invention has been the IQ test, and certainly as far as fame and fortune is concerned that's an accurate description and IQs are what put psychology on the map in this country.

But I think IQ tests and their progeny, among which I include SAT and other kinds of achievement tests, have reached the highest point of utility that they're going to have and I think their utility will decline in the future.

I think that, to put it simply, testing has gone as far as it can.

We need to move to a much broader view of assessment, which is much more in touch with the actual thinking processes of the individual and which yields information which the teacher, the school district, and the student himself can make use of to have more successful learning in the future.

>> ZIMBARDO: At Yale University, psychologist Robert Sternberg has been studying the limitations of testing in measuring what he calls practical intelligence.

>> An example of practical intelligence is being able to find your way through busy city streets such as these.

Suppose, for example, that you wanted to go to the bank.

There are a number of alternative routes you might use to get to the bank.

The person who's higher in spatial ability would be able
quickly to figure out which of those routes will get him or her to the bank faster, whereas the person who's lower in spatial ability might have difficulty in figuring out which of those routes would be best.

154 01:49:39:07  >> ZIMBARDO: Sternberg recently devised a test specifically for spatial ability.

155 01:49:47:01 Subjects must diagram how to get between two places on this map.

156 01:49:51:20  >> And what we have found is that there is very, very little correlation or relationship between this practical aspect of intelligence and, say, the more academic one, traditionally measured by IQ tests or the SAT, so that people could be good in one and not very good in the other.

157 01:50:12:09 I view intelligence as the ability to adapt to existing environments -- in other words, to go into an environment and figure out what you need to know in it and then work effectively within that environment to shape environments into new ones -- in other words, to change the environment so as to make it more suitable for your interests, your abilities, or what you want, and where appropriate, to select a new environment.

158 01:50:40:28 If you doubt whether the standard view of intelligence is narrow, one thing you can do is look at how well standard intelligence tests predict not school performance, but performances of various kinds in everyday life: how well someone will do as a business executive, a lawyer, an accountant, a doctor, whatever.

159 01:51:00:18 And what you find is that the prediction level is very low.

160 01:51:04:01 The test we have for managers looks at a person's ability to pick up what you need to know to succeed in a particular job, in this case, management, that no one explicitly tells you and that often isn't even verbal.

161 01:51:21:10 So what we found is when you interview people who are highly successful in different careers, they say that what really matters to their success isn't so much what they learned in college, as it is what they pick up on the job.

162 01:51:34:23 So what we have done, we being myself and Richard Wagner, is to create a test of this practical intelligence in different fields, like for management, for salespeople, for
And what we find is that our test predicts performance on the job about twice as well as does the standard IQ test, but scores on our test are not themselves related to the IQ test scores.

>> ZIMBARDO: There's also another, very different kind of effort to measure intelligence, this one conducted by neuroscientists.

Unlike Robert Sternberg and Howard Gardner, they proposed to bypass both mind and environment and go directly to the brain itself.

Using special recording equipment tuned to certain brain-wave frequencies, researchers can measure the brain's response as it adjusts to interruptions by sudden stimuli.

The EEG recording of certain brain waves called P300 waves is particularly sensitive to these sudden interruptions.

The theory is that the brain of a smarter person will have smaller P300 waves over time because it adjusts faster to the interruptions and absorbs new inputs more readily.

But whether new, biologically based measures of intelligence will prove valid or will create new problems by further limiting the definition of intelligence remains to be seen.

Whatever the case, the controversies over what intelligence really is, how it should be assessed, and the relative contributions of heredity and environment are bound to continue.

The only thing we know for sure is the existence of individual differences between us.

In our next program, we'll look at the most profound differences of all, the ones between the sexes.

Are males and females really different psychological creatures?

Does anatomy determine behavioral destiny?

Or does society shape the XY genes into masculinity and the XXs into femininity?
It's nature versus nurture in the realm of the sexes next time.

I'm Philip Zimbardo.

Funding for this program is provided by Annenberg/CPB to advance excellent teaching.