

506 F.Supp. 831
United States District Court, N.D. Illinois, Eastern
Division.

PARENTS IN ACTION ON SPECIAL
EDUCATION (PASE), an incorporated
association; Lue B. B., on her own behalf and as
next friend of Barbara B.; and Onollie J., on her
own behalf and as next friend of Angela J., on
behalf of themselves and all other persons
similarly situated, Plaintiffs,

v.

Joseph P. HANNON, Individually, and in his
capacity as General Superintendent of Schools in
Chicago; Elberta Pruitt, Individually and in her
capacity as Director of Special Education for the
Chicago Board of Education; Louise G. Daugherty,
Individually and in her official capacity as Director
of Pupil Personnel Services for the Chicago Board
of Education; William Canning, Individually and
in his capacity as Director of the Bureau of Child
Study for the Chicago Board of Education; the
Chicago Board of Education, a body corporate and
politic; John D. Carey, Dr. Edgar G. Epps, Dr.
Bernard S. Friedman, Herbert E. Johnson, Henry
W. McGee, Mrs. Louis A. Malis, Thomas J.
Nayder, Patricia O'Hern, Carey B. Preston, Mrs.
William L. Rother, Gerald L. Searboro, Carmen
Velasquez, and Mrs. Lydon Wild, Individually and
in their official capacity as past or present
members of the Chicago Board of Education;
Joseph Cronin, Individually and in his capacity as
Superintendent of the Illinois Office of Education;
the Illinois State Board of Education, a body
corporate and politic; and the Illinois Office of
Education, the State Educational Agency of
Illinois, Defendants.

No. 74 C 3586.

|
July 7, 1980.

Synopsis

Two black children who were placed in classes for the educable mentally handicapped after achieving low scores on standard intelligence tests brought suit on behalf of all similarly situated black children against city board of education and its officers responsible for administration

of the relevant programs. The plaintiffs alleged that the intelligence tests administered by the city board of education were culturally biased against black children and that the use of such tests violated the equal protection clause of the Fourteenth Amendment as well as various federal statutes. After a three-week trial to the court, the District Court, Grady, J., held that: (1) though one item on one of the challenged intelligence tests and eight items on two other tests were culturally biased against black children, or at least sufficiently suspect that their use was inappropriate, the items would not significantly affect the score of a child taking the test and did not render the tests unfair, and (2) the challenged tests, when used in conjunction with the statutorily mandated other criteria for determining an appropriate educational program for a child, did not discriminate against black children in the city's public schools.

Judgment for defendants.

Attorneys and Law Firms

*833 Wallace C. Winter, Legal Advocacy Service, David A. Goldberger, American Civil Liberties Union, Linda Lipton, Better Government Ass'n, James L. Pittman, Lord, Bissell & Brook, Chicago, Ill., for plaintiffs.

Michael J. Murray, Patricia J. Whitten and Christine Cheatom, Chicago, Ill., for defendants.

MEMORANDUM DECISION

GRADY, District Judge.

This case presents the question whether standard intelligence tests administered by the Chicago Board of Education are culturally biased against black children. The action is brought on behalf of all black children who have been or will be placed in special classes for the educable mentally handicapped ("EMH") in the Chicago school system. The defendants are the Chicago Board of Education and its officers responsible for administration of the relevant programs. The named plaintiffs are two

black children who were placed in EMH classes after achieving low scores on standard intelligence tests.

The Illinois school code requires classes for the educable mentally handicapped, who are defined as:

. . . children between the ages of 3 and 21 years who because of retarded intellectual development as determined by individual psychological evaluation are incapable of being educated profitably and efficiently through ordinary classroom instruction but who may be expected to benefit from special education facilities designed to make them economically useful and socially adjusted.

Ill.Rev.Stat. ch. 122, s 14-1.04 (1977).

There are 483,209 children enrolled in the Chicago public school system. Of those, 299,590, or 62 per cent, are black. For the 1978-79 school year, 13,225 children were enrolled in EMH classes. Of these, 10,833, or 82 per cent, were black. Of the 106,581 white children enrolled in the system, 1,404 were attending EMH classes. Three and 7/10 per cent of all black students enrolled in the system are in EMH, whereas only 1.3 per cent of the white students are in EMH.

The EMH curriculum is designed for the child who cannot benefit from the regular curriculum. It is designed for children who learn slowly, who have short attention spans, slow reaction time and difficulty retaining material in both the short term and the long term. The curriculum also recognizes the difficulty an EMH child has in ***834** seeing similarities and differences, in learning by implication, in generalizing and in thinking abstractly. The curriculum thus involves much repetition and concrete teaching. Subjects are taught for short periods of time, in recognition of the children's short attention spans. The subject matter of the EMH courses is oriented toward socialization, language skills and vocational training. Academic subjects are taught, but on an elementary level and with the objective of helping the child become economically independent. The assumption of the EMH curriculum is that the child will not go on to college, and,

in fact, children who graduate from EMH programs in the Chicago school system are given special diplomas which do not qualify them for college entrance.

These characteristics of the EMH program were described by plaintiffs' witness Dale Layman, a professor at the University of Illinois who specializes in training special education teachers and designing special education curricula. Dr. Layman had no argument with the EMH curriculum in Chicago, and believes it is well suited for EMH pupils. She testified that the underlying assumptions about the learning abilities of EMH students are valid, and that it is not realistic to expect a child who is genuinely retarded to be able to cope with the regular curriculum.

Dr. Layman and several other witnesses testified about the social stigma which attaches to a child who is assigned to a classroom for the retarded. While the teachers and school administrators attempt in various ways to protect the children, the evidence establishes without doubt that EMH pupils suffer from feelings of inferiority and that the label they receive in school often follows them throughout their lives.

An erroneous assessment of mental retardation, leading to an inappropriate placement of a child in an EMH class, is clearly an educational tragedy. However beneficial such classes may be for those who truly need them, they are likely to be almost totally harmful to those who do not. The two named plaintiffs in this case are examples of what can happen. Each of these children had learning disabilities but was erroneously diagnosed as being mentally retarded. Each of them scored low on a standard intelligence test administered as part of the assessment process. The two plaintiffs were assigned to EMH classes, where they spent several years. As a result of a belated re-evaluation, it was determined that these two children were not mentally retarded but rather were children in the normal range of intelligence whose learning was hampered by disabilities which are remediable.

The two named plaintiffs claim that their misassessment as retarded children was caused by racial bias in the standard intelligence tests they took, causing them to achieve low scores. It is claimed on behalf of the two named plaintiffs and the class they represent, consisting of all black children in the Chicago school system who are or might be assigned to EMH classes, that the use of racially biased intelligence tests in EMH placement violates the Equal Protection Clause of the Fourteenth Amendment as well as various federal statutes.¹ Plaintiffs seek declaratory and injunctive relief. The principal relief

sought is a permanent injunction against the use of standard IQ tests in the evaluation of black children for EMH placement.

The case was tried to the court over a period of three weeks in January 1980. Each side presented a number of expert witnesses, mostly psychologists and educators. *835 Each side also offered a large volume of reading material, including a number of articles in psychological and education journals. At the conclusion of the testimony, I recessed to examine this voluminous material prior to final arguments. During the recess, the United States Department of Justice filed a motion for leave to file an amicus curiae brief and to participate in oral argument. Defendants objected to the motion, pointing out that the case had been pending for six years and that the Department of Justice had at all times been aware of it. The objection was overruled and the Department did file a brief and participate in the oral argument, which took place on March 11, 1980. The government sided with the plaintiffs on all issues.

It will be helpful at this point to indicate the organization of this opinion. The early sections will be devoted to a description of the factual contentions of the parties and an extensive description of the specific items on the three intelligence tests which are in issue. I will in some instances comment upon the merits of the parties' respective positions during the course of these descriptions, but generally I will save the statement of my own conclusions until later, *infra* at pp. — et seq.²

The disagreement between the parties can be summarized briefly. It has been known since the early days of standard intelligence tests, around the time of World War I, that blacks as a group score about one standard deviation 15 points lower than whites. On the Stanford-Binet test, for instance, the mean white score is 100 and the mean black score is 85. While there is no disagreement as to the existence of this phenomenon, there is considerable disagreement about what causes it.

The psychologists who developed the Stanford-Binet test in this country, Terman, Yerkes and Goddard, believed that they were measuring innate mental abilities which were not subject to change. This was their concept of "intelligence." They explained the relatively poor performance of blacks, as well as that of many other groups such as recent immigrants to this country from southern and eastern Europe, on the basis of genetic inferiority.

The genetic view had wide acceptance among psychologists for a time, but it lost ground in the light of studies which showed that IQ scores were in fact not constant but were subject to change. The genetic view never did take account of the fact that many blacks scored above the white mean, but this seems to have been overlooked. However, it was more difficult to overlook the fact that northern blacks scored higher on the average than southern whites, and that blacks who moved north often experienced an increase in their IQ scores. Black infants adopted by white families tended to achieve IQ scores in later years which correlated highly with the scores of the natural children of the adoptive parents.

The current view of most psychologists is that IQ tests measure something which is changeable rather than something that is fixed for all time, something which can be increased and improved. The parties in this case agree on that much.

The question remains, what does the IQ score measure? Dr. Leon J. Kamin, a psychologist from Princeton University, testified for plaintiffs that the IQ tests measure *836 "current performance." Dr. Kamin is the country's leading exponent of the view that the tests measure nothing innate. He writes and lectures extensively on the subject. He testified that in his opinion differences in performance on the tests are due solely to differences in exposure to "information" called for by the tests.

Dr. George Albee, a Professor of Psychology at the University of Vermont, was another witness for the plaintiffs who stated a similar view. He testified that the IQ tests measure a child's "sharing of the dominant white culture." Poor performance by a black child simply indicates his lack of exposure to white culture. It denotes nothing about the intellectual functioning of the child.

Defendants contend that the tests measure the child's current level of abilities which correlate significantly with his prospects of succeeding in school. Two school psychologists, Dr. Terrence Hines and Mr. Elmer Smith, testified for defendants that the IQ tests afford an indication of the areas of the child's mental strengths and weaknesses. According to Mr. Smith, who has a masters degree in psychology from Northwestern University and some 30 years experience as a school psychologist, the tests give an indication of the child's ability to retain factual information, to attend, to concentrate, to formulate new associative learning, and to perform simple arithmetic processes. These abilities are called for by the regular school curriculum, and accordingly the test results

have some predictive value. Defendants' witnesses concede a slight amount of cultural bias in the tests but deny that this results in erroneous placements or deprives the tests of their usefulness. They point out that a diagnosis of retardation is not based solely upon an IQ score but upon a combination of relevant factors. These witnesses also emphasized that the IQ score affords a criterion that is relatively objective. They fear that, lacking the student's score on a standardized test, they would be forced to make the assessment upon a largely subjective basis.

This testimony, standing alone, does not preponderate in either direction. I have seen cases in which one set of experts is clearly more credible than the other and will, by their demeanor, appearance, credentials, and the reasonableness of their testimony, carry the day. This is not such a case. None of the witnesses in this case has so impressed me with his or her credibility or expertise that I would feel secure in basing a decision simply upon his or her opinion. In some instances, I am satisfied that the opinions expressed are more the result of doctrinaire commitment to a preconceived idea than they are the result of scientific inquiry. I need something more than the conclusions of the witnesses in order to arrive at my own conclusions.³

*837 Plaintiffs produced only one witness who made any attempt to demonstrate racial or cultural bias in specific test items. This was Dr. Robert Williams, whose testimony we will discuss in detail at a later point in this opinion. The other plaintiffs' witnesses who expressed the opinion that the tests are biased did not attempt to demonstrate or illustrate their point by any reference to specific items on any test.

It is obvious to me that I must examine the tests themselves in order to know what the witnesses are talking about. I do not see how an informed decision on the question of bias could be reached in any other way. For me to say that the tests are either biased or unbiased without analyzing the test items in detail would reveal nothing about the tests but only something about my opinion of the tests.

Plaintiffs were ambivalent in their attitude toward the need to analyze the specific test items. On the one hand, they recognized the relevance of such an inquiry by presenting Dr. Williams' testimony concerning bias in particular test items. However, he testified about only a few of them. None of the attorneys for plaintiffs nor the attorneys for the Department of Justice were prepared to

discuss specific test items during the day-long oral arguments at the conclusion of the case, even though I had indicated long before the conclusion of the evidence that I felt analysis of specific test items was essential to a proper understanding and decision of the case. I am not satisfied that any of the dozen or so attorneys who participated in the trial of the case have even read the tests. In response to a direct inquiry during final argument, some of them admitted they had not and the rest said they had "at one time, but not recently." Plaintiffs' attorneys, as well as one attorney for defendants, stated that they felt it was unnecessary to look at the tests.

I have said enough to indicate my belief that an analysis of the tests is essential. I will now proceed to that task. Plaintiffs' criticism of specific test items will be discussed as we go along.

Three tests are challenged in this case. They are the three intelligence tests most often used in the assessment of mental retardation in the Chicago public school system. Most children referred for evaluation are given one or two of these tests, and the one most frequently given is the Wechsler Intelligence Scale for Children, Revised. This is known as the "WISC-R."

THE WISC-R

The test is divided into twelve sub-tests, and each sub-test consists of a series of similar items of increasing difficulty. There are suggested starting points for children of different ages, and if a child answers the initial items for his age correctly, he is given credit for all items prior to that level. If a child is unable to answer the items for his age, the examiner goes backwards in the sub-test until he finds the child's level of performance. If a child is suspected of mental retardation, the test instructions require that the examiner start with the first item in each sub-test. In some evaluations only ten of the twelve sub-tests are used, but when a child is suspected of retardation all twelve are used.

The first sub-test on the WISC-R is "Information." Items 1 through 4 are suggested for ages 6 and 7. These four questions are as follows:

1. (The examiner, showing the child his thumb, asks)

“What do you call this finger?”

2. “How many ears do you have?”

3. “How many legs does a dog have?”

4. “What must you do to make water boil?”

The acceptable “responses” for the first three questions are “thumb,” “two,” “four,” and, for No. 4, “heat it . . . put it on the stove . . . put fire under it . . . turn the stove on . . . cook it (or any such response indicating that the water must be heated.)”

There are thirty questions on the “Information” sub-test, and the maximum number of points the child can obtain is thirty, one point for each question answered correctly. *838 The examiner continues to ask the child questions, even beyond those designated for his age group, until the child misses five consecutive questions.

Questions 5 and 6 on the Information sub-test are the suggested starting points for 8 to 10 year olds. These questions are:

5. How many pennies make a nickel?

6. What do we call a baby cow?

The correct responses are “five” and “calf.”

Questions 7 through 10 are suggested as the starting points for 11 to 13 year olds. They are as follows:

7. How many days make a week?

8. Name the month that comes next after March.

9. From what animal do we get bacon?

10. How many things make a dozen?

The acceptable responses are “seven” (except that if the child answers “five” he is to be asked “how many counting the weekend?”); “April,” “pig . . . hog . . . piggy,” and “twelve.”

Questions 11 through 30 are for 14 to 16 year olds. They become progressively more difficult. Question 11 asks for the four seasons of the year. They can be named in any order. Question 12 asks, “Who discovered America?” Acceptable responses are “Columbus . . . Leif Erickson, Vikings (Norsemen), Amerigo Vespucci. (If a child says

‘Indians,’ say ‘yes, the Indians were already there, but who sailed across the ocean and discovered America?’).”

Item 12 is the first on the WISC-R to draw the fire of the plaintiffs. Their witness Williams stated that this item is “absolutely insulting” to Native American children, since it implies that the land where their forebearers resided needed to be “discovered” by someone else. Dr. Williams also criticized the question as confusing, since it is a contradiction to say that something was “discovered” when it was already occupied. Whatever the merits of these reactions may be, Dr. Williams did not explain how Item 12 discriminates against black children as opposed to white children, and we assume that he was not attempting to establish any such discrimination.

Item 13 asks, “What does the stomach do?” The acceptable responses all center on the function of the stomach in digesting or holding food. Dr. Williams testified that many black children answer, “It growls.” He attributes this to the fact that many black children come from poverty level families and simply do not have enough to eat. The point seems well taken, and, if no credit would be given for such an answer, it is clear that a black child would be penalized unfairly. The WISC-R manual instructs the examiner to give the child credit for any response which is similar to or better than the acceptable responses listed. (Manual, p. 61) On the other hand, the examiner is cautioned to adhere strictly to the test procedures outlined in the manual, since “to change the conditions of administration is to change the test results to an unknown degree.” (Manual, p. 53) The test is to be administered by a trained psychologist, and one would hope that the child who answers that his stomach “growls” would either be given credit for an “equal to” response or at least that the response would come up for later discussion when the child is evaluated for placement. The test kit includes a record form which the examiner is to complete as he administers each item to the child. The manual instructs the examiner to record the response in sufficient detail so that it is preserved for later discussion. (Manual p. 63)

Item 14 asks, “In what direction does the sun set?” The answer is west, and if the child points, he is asked what direction that is. Dr. Williams did not criticize this question, but it was cited by Dr. Dale Layman in the only instance where a witness other than Dr. Williams mentioned a test item during plaintiffs’ case. Dr. Layman said that this item would be unfair for a child who lives in a high-rise housing project and has never been on the west side of the building to see the sun set. It was not clear to

us whether Dr. Layman was basing her opinion upon actual experience with this test item or not. She did not specifically say that she has encountered any such difficulty with this question.

***839** Questions 15 through 20 ask what month has an extra day during leap year; who invented the electric light bulb; from what country did America become independent in 1776; why oil floats on water ("because it is lighter . . . it's not as heavy . . . it's less dense . . . it's not heavy enough to go to the bottom . . . it's more buoyant . . . water is heavier"); name the two countries that border the United States (both must be named for credit); and how many pounds make a ton (credit is given for 2,000 or any answer from 2,200 to 2,240, since the child may be referring to a metric ton). No witness referred to any of these questions.

Question No. 21 asks, "In what continent is Chile?" Dr. Williams criticized the question by saying, "That's hardly our environment." Since the WISC-R manual defines intelligence as ". . . the overall capacity of an individual to understand and cope with the world around him," Dr. Williams says the question does not measure intelligence. Considering that this test item is the twenty-first of thirty items on the Information sub-test, it is doubtful that a school-age child suspected of retardation would even reach it. One problem we have with Dr. Williams' testimony is that he never referred to the age level for which the various items are recommended. In any event, it does not appear to us why the location of Chile would be more familiar to a white 14 year old child than to a black 14 year old child.

Question 22 asks, "What is the main material used to make glass?" The acceptable responses call for a reference to sand, silica or quartz.

Item 23 asks, "What is the capital of Greece?" Dr. Williams criticizes this question on the same basis he found fault with the question about Chile; "Athens is not a part of our environment."

Item 24 asks, "How tall is the average American man?" Any answer from 5 ft. 7 in. through 5 ft. 11 in. is acceptable. The examiner is specifically instructed not to give credit for 5 ft. 6 1/2 in. or 5 ft. 11 1/2 in. Dr. Williams criticizes this item on the basis, again, that it does not test a child's ability to cope with his environment.

Item 25 asks, "What is a barometer?" Any answer which indicates that it measures air pressure or forecasts rain or

weather is acceptable.

Item 26 asks, "What causes iron to rust?" Any answer indicating oxidation is acceptable.

Item 27 asks the distance from New York to Los Angeles, and any answer from 2,500 to 3,500 miles is acceptable.

Item 28 asks, "What are hieroglyphics?" The acceptable responses are anything indicating ancient writings or symbols. Dr. Williams expresses dissatisfaction with this item, again because it is "still not a part of this culture."

Item 29 asks, "Who is Charles Darwin?" Dr. Williams testified:

If you're going to ask kids questions about their environment, the better questions would be, would relate to folks that they're familiar with, people with whom they come in contact or they read about and who are part of their particular black culture.

That would be a better estimate. For example, why not ask them who is Malcolm X, who was Martin Luther King, or who were some of these people that they would have had an opportunity to learn about, rather than Charles Darwin who was an Englishman.

Dr. Williams finds it ". . . interesting, again, that Darwin would be included in a question, here. As I indicated before, he wrote Origin of the Species, and that he also had a very negative theory against black people."

Finally, Item 30 on the Information sub-test is, "What does turpentine come from?" Any answer indicating a fir or evergreen tree is acceptable.

To summarize plaintiffs' specific criticism of the thirty information items, No. 13, the stomach question, is said to be biased against black children because they would be more likely than white children to think ***840** of a stomach as something that growls rather than something that digests or stores food. Item 14, the direction in which the sun sets, may also have been criticized by Dr. Layman as a question that penalizes black children who live in high-rise projects. This is not clear. These are the only two items of the thirty as to which some rationale has been suggested to explain why a black child would respond differently than a white child. Six other questions, 12, 21, 23, 24, 28 and 29, are criticized as not relevant to a child's "environment." Dr. Williams stated why certain other persons would be better known than

Charles Darwin to black children, but it seems likely that there are names which would be better known to white children as well. For instance, it may be that white children would be more apt to know about Malcolm X and Martin Luther King than Charles Darwin. If not, it would still be easy to think of other names that are better known than Darwin. But the object of the test is not always to ask for information that is well known to everyone. The Darwin question is 29th in a list of 30 items arranged in order of increasing difficulty. It is a question which most children are expected to miss. It is not a realistic possibility that any child, white or black, would fall below a score of 80 on the test the dividing line defendants use for EMH eligibility because of missing this question. A child who is going to score low on the test will simply not reach the question.

The other items which are criticized for not being part of a child's "environment" are not claimed to be any more directly related to the everyday experience of a white child than they are to that of a black child. Hieroglyphics are not a pressing issue for any American child, white or black.

The next sub-test of the WISC-R is entitled "Picture Completion." It contains twenty-six items, with a point for each item. The test consists of twenty-six cards, each with a picture of an object. Something is missing in the picture and the child is to say what is missing. He continues with the test until he has four consecutive misses.

Items 1 through 4 are for 6 and 7 year olds. Item 1 shows an ordinary hair comb with some missing teeth. Dr. Williams criticized this item as unfair to black children because they may have been exposed only to an Afro-type comb and do not recognize the article shown in the picture.

Item 2 is a picture of a black woman with no mouth.

Item 3 shows a fox without a left ear. The right ear is clearly shown.

Item 4 shows the back of a hand. The little finger has no nail, whereas the nails on the other four fingers are vividly colored.

Items 5 through 26 are for children ages 8 to 16. Item 5 shows a cat with whiskers on the right side of its face but no whiskers on the left side. Item 6 shows a girl holding a doll, looking into a mirror. The mirror image shows the

girl but not the doll. Item 7 is a clock showing all numbers but "8." There is a blank space where the "8" should be. Item 8 shows an elephant with only three legs. Item 9 shows a step ladder with one step missing. Item 10 is a picture of a dresser with four drawers. There are two knobs on three of the drawers but the fourth drawer has only one knob. Item 11 shows a belt with a buckle but no holes. Item 12 is a front view of a white man's face. Part of his nose is missing.

Item 13 is a picture of a door. There is a hinge shown toward the top of the door, but no other hinge is shown. Dr. Williams testified that a black child from poor economic circumstances might be accustomed to doors with missing hinges, so that he would not understand what element is missing in this picture.

Item 14 is a 5 of diamonds playing card. A diamond is shown in each of the four corners. The fifth diamond, which belongs in the center of the playing card, is missing.

Item 15 shows a black girl with a shoe and sock on her right foot but only a shoe on her left foot. Item 16 is a front view of a man's jacket showing three buttons but no buttonholes. Item 17 shows a boy wearing *841 a wristwatch with the band missing. Item 18 shows a pair of scissors in the open position. The screw which connects the two blades is missing.

Item 19 shows a profile view of a white girl without an ear. Item 20 is a screw without a slot in the head. Item 21 shows a cow with a cleft in three of its hooves but no cleft in the fourth hoof. Item 22 shows a thermometer which has no mercury in its bulb.

Item 23 shows the sun shining on a house and tree. The house casts a shadow, but there is no shadow cast by the tree. Item 24 is a telephone which has no cord connecting the receiver to the base. Item 25 is a profile view of a white boy without an eyebrow. Item 26 shows an open umbrella without spokes.

The third sub-test on the WISC-R is "Similarities." All children begin with Item 1 and discontinue after missing three consecutive items. The seventeen items have different point values. Items 1 through 4 are one or zero, Items 5 through 17 are scored either two, one or zero.

On each item, the child is asked to tell how two things are alike. Item 1 is "wheel-ball." Acceptable responses are that they are both round, they both roll, they are both

circles, they are both shaped like a zero, or an “o.” Items 2, 3 and 4 are “candle-lamp,” “shirt-hat,” and “piano-guitar.” The acceptable responses are the obvious ones.

The two-point items begin with Item 5. Two points are given for answers which identify a general classification into which both items fit. One point is given for identifying any specific properties or functions which are common to both items. For example, Item 5 is “apple-banana.” Two points are given for any response indicating that both are fruits, one point is given for any response indicating specific properties, such as that both can be eaten. Item 6 is “beer-wine.” Two points are given for any response indicating that both are alcoholic beverages, one point is given for a response indicating that both are drinks or both are liquid. The remaining items are as follows, with the two-point responses indicated in parentheses:

7. Cat-mouse. (Animals)
8. Elbow-knee. (Joints)
9. Telephone-radio. (Means of communication)
10. Pound-yard. (Units of measurement)
11. Anger-joy. (Emotions)
12. Scissors-copper pan. (Made of metal or utensils)
13. Mountain-lake. (Natural features of the earth)
14. Liberty-justice. (Ideals)
15. First-last. (Extremes or positions in a series)
16. The numbers 49 and 121. (Both are perfect squares, or have odd-number square roots)
17. Salt-water. (Necessary for life or chemical compounds)

Plaintiffs presented no evidence that any of the Similarities items are racially biased. Dr. Williams did not mention this sub-test.

The fourth sub-test is “Picture Arrangement.” This consists of thirteen sets of pictures. Each set is presented to the child in a mixed-up order, and he is asked to arrange the cards in a sequence that tells a sensible story. The mixed-up order is the same for each child.

The child is allowed 45 seconds for each of Items 1 through 8 and 60 seconds for Items 9 through 12. The child continues with the items until he has three consecutive failures. Bonus points are given if the child finishes in less time than allotted for the item. The maximum score is 48 points.

Items 1 and 2 are for 6 and 7 year olds. Item 1 consists of three cards about a boxing match. One card shows the boxers in the ring, sparring. One is in black trunks, one is in white trunks. The second card shows the boxer in white trunks obviously winning the fight, and the third card shows the boxer in black trunks being carried away from the ring while the boxer in white trunks stands in the ring with his arms raised in victory.

***842** Item 2 consists of three cards about a picnic. One card shows a man and woman carrying a picnic basket and being followed by a dog which is looking hungrily at the basket. Another card shows the dog pulling food out of the basket as the couple, unaware of his presence, continues walking. A third card shows the couple with their picnic blanket spread, reacting in consternation as they view their empty picnic basket.

Before the child starts each item, the examiner tells him something about the pictures. In Item 1, the examiner says, “These pictures tell the story about a fight, a boxing match. The pictures are in the wrong order now. See if you can put them in the right order so they tell a story that makes sense.” In Item 2, the examiner says, “These pictures tell a story about a picnic. These pictures are in the wrong order now. See if you can put them in the right order so they tell a story that makes sense.”

Items 3 through 12 are for 8 to 16 year olds. Item 3 consists of four cards. One shows a boy playing with matches and being scolded by his mother. Another card shows the match box and the window curtain on fire, and the boy running away. A third card shows a manned fire truck racing along the street, and a fourth card shows the little boy crying outside the burning building while the firemen fight the fire.

Item 4 consists of four cards which tell the story of a boy who used a nearby lumber pile to make a bridge to cross a stream.

Item 5 is four cards showing a burglar entering a window, stealing some items from a dresser and being confronted by a policeman as he comes back out the window.

Item 6 shows a man waking up to his alarm clock, wolfing down his breakfast, running down the sidewalk to work and then sleeping again at his desk at work.

Item 7 is four pictures showing an artist setting up her easel and painting a picture.

Item 8 shows a western gunman looking at a lasso in a store window and then shows the store proprietor tied up with the lasso while the gunman rifles the cash register.

Item 9 is a five-card story about a man whose boat drifts away from him before he can board it.

Item 10 tells a story about a boy whose mother handed him a spade with instructions to plant a tree. He used the spade to dig worms to go fishing.

Item 11 shows a man who purchased a bench, and, while walking down the street with it, accidentally bumped another man in the head. The two men then become engaged in a fight while two spectators sit on the bench and watch.

Item 12 starts with a girl who is refusing an umbrella being offered by her mother. The mother is pointing to rainclouds visible through the window. The girl then goes out and the rain starts. The girl comes back inside, dripping water, and goes out again carrying the umbrella.

Dr. Williams did not comment on the picture arrangement sub-test in the WISC-R. He did criticize two of the items as they appeared in the earlier WISC test, and we will note those criticisms when we discuss the WISC.

Sub-test five is "Arithmetic." There are eighteen items, with one point for each, and a maximum score of 18 points. Children 6 and 7 years of age start with Item 1, 8 to 10 year olds start with Item 5, 11 to 13 year olds start with Item 8, and 14 to 16 year olds start with Item 10. There are 30 second time limits for Items 1 through 13, 45 seconds for Items 14 and 15, and 75 seconds for Items 16 through 18. The test is discontinued after the child has had three consecutive failures. A child may not use pencil and paper for any problem.

Items 1 through 4 are done with two cards. One card is blank and the other has 12 trees in a horizontal line.

In Item 1, the examiner places the tree card before the child and says, "Count these trees with your finger. Count

them out loud so I can hear you." If the child counts the 12 trees, he scores one point.

In Item 2, the child is given the blank card and asked to cover up all of the trees *843 on the other card except four. "Leave four trees showing."

In Item 3, the child is asked to cover up all trees but nine. In Item 4, the child is asked how many trees there would be on the tree card if one tree were added at each end of the line.

In Items 5, 6 and 7, the child is asked how many pieces he would have if he cut an apple in half, how many ribbons Barbara would have if she started with five and lost one, and how many pennies John would have if he started with four and his mother gave him two.

In Item 8, the child is told that Jim had eight marbles and bought six more. "How many marbles did he have altogether?" Item 9 asks how many newspapers a boy would have if he started with twelve and sold five.

Item 10, the starting point for 14 to 16 year olds, asks how much three candy bars would cost if they cost 8 cents each. Item 11 states that Bill, Dave and Tom each earned \$9.00 working in a supermarket and asks how much they earned altogether. Item 12 states that a milkman had 25 cartons of milk and sold 14. "How many cartons did he have left?" Item 13 asks how many hours a workman worked if he was paid \$4.00 an hour and earned \$36.00. Item 14 asks how much change you would get back from a dollar if you bought two dozen pencils at 45 cents a dozen. Item 15 concerns four boys who had 72 pennies. If they divided them equally, how many pennies did each boy receive?

Items 1 through 15 are read to the child by the examiner. In Items 16 through 18, the child reads the problem aloud from a book. However, if he cannot read, the examiner will read the problem to him.

Item 16 asks, "If three pieces of bubblegum costs 5 cents, what will be the cost of 24 pieces?" Item 17 reads, "Tony bought a secondhand bicycle for \$28.00. He paid two-thirds of what the bicycle cost new. How much did it cost new?"

Item 18 asks, "A jacket that usually sells for \$32.00 was on sale for $\frac{1}{4}$ less. When no one bought it, the store owner reduced the sale price by $\frac{1}{2}$. How much did the jacket sell for after the second price reduction?"

(Emphasis in original).

Plaintiffs offered no criticism of the Arithmetic sub-test. Dr. Williams did not mention it. Of the seven sub-tests still to be discussed, only one was referred to by any of the witnesses.

The sixth sub-test is "Block Design." The materials consist of nine small cubes colored red on two sides, white on two sides, and red/white on two sides. The remaining materials are eleven cards printed with pictures of the colored cubes arranged in various patterns. Each card is an item on the test. The child is shown the card by the examiner and asked to arrange the cubes to show the same design that appears on the card. The examiner demonstrates how it is done on the first three items before the child tries for himself. There is a time limit of 45 seconds for each of the first four items, 75 seconds for Items 5 through 8, and 120 seconds for Items 9 through 11. The child is allowed two tries on Items 1, 2 and 3. For each of Items 4 through 11, there are bonus points given for completion within various intervals under the allotted time limit. The maximum score for the eleven items is 62 points. The child continues until he has failed on two consecutive items.

As with the other sub-tests, the items become progressively more difficult. Items 1 and 2 on this sub-test are for 6 and 7 year olds, and Items 3 through 11 are for ages 8 through 16. It is unnecessary to describe every item. A description of the first three items and one of the advanced items will suffice to indicate the nature of the test.

Item 1 shows four cubes arranged in a square. The bottom two cubes are solid red and the top two are solid white. Item 2 again shows four cubes stacked in a square. This time, the white ones are at the top left and bottom right while the red ones are at the top right and bottom left. Item 3 is again an arrangement of four cubes in a square. The top two cubes and the bottom right cube are red. The bottom cube on the left is divided diagonally into a red half on the right and a white half on the left.

***844** Item 11 is an exotic design which looks something like a pinwheel. It requires all nine cubes and is difficult to construct within the time limit.

The seventh sub-test on the WISC-R is "Vocabulary." The items are thirty-two words, with a maximum score of 64 points. The examiner says the words to the child. ("I am going to say some words. Listen carefully and tell me

what each word means.") The child can score either one or two points on each item, depending upon the quality of the answer. The manual provides detailed scoring instructions, with sample answers, indicating how various responses should be scored. The examiner is instructed to disregard "elegance of expression." (Manual, p. 161). Generally, a two-point answer is one which shows in some way that the child is thoroughly familiar with what the word means, whereas a one-point answer is one showing less understanding. An obviously wrong answer results in zero points.

Six and 7 year olds start with Item 1, 8 to 10 year olds start with Item 4, 11 to 13 year olds start with Item 6, and 14 to 16 year olds start with Item 8. A child continues until he has five consecutive failures.

Item 1 is "knife." The following are given as samples of two-point responses: "Something you cut with . . . has a blade and a handle . . . silverware, it cuts . . . a weapon . . . to stab with . . . you can peel an apple with it" Samples of one-point responses are listed as: "eat with it . . . to kill people . . . sharp . . . made of steel . . . you can scare people with a knife . . . to hunt with" Zero-point responses are: "I play with it . . . I have one . . . put in your pocket."

Item 2 is "umbrella." Two-point responses are: "Use it to keep the rain off . . . protects you when it rains . . . put it over your head when it rains . . . so you don't get wet when it rains." One-point responses are: "Carry it when it rains . . . big round thing that can fold up . . . put it over your head . . . to keep off the sun . . . you hold it up (gives appropriate demonstration) . . . helps you if it starts raining . . . keeps you dry." All of these one-point responses are marked with "Q" in the manual, indicating that the examiner should follow up the response with another question as to what the child means. If a child says, "Put it over your head," the examiner should ask, "Explain what you mean." If the child says something like, "You know, like when it rains," he is given two points for the response.

Dr. Williams criticized "umbrella" as a vocabulary word for black children because a black child might call the object a "parasol" and not know the meaning of the word "umbrella." According to Dr. Williams, the object is called a "parasol" in the black community. He did not indicate whether the word "umbrella" is also known in the black community, in the same way "parasol" is known but not commonly used in the white community.

The “umbrella” item is the only one on the Vocabulary sub-test which drew any comment from plaintiffs.

Items 3 through 32 of the vocabulary test are as follows:

- | | |
|--------------|----------------|
| 3. Clock | 18. Fable |
| 4. Hat | 19. Hazardous |
| 5. Bicycle | 20. Migrate |
| 6. Nail | 21. Stanza |
| 7. Alphabet | 22. Seclude |
| 8. Donkey | 23. Mantis |
| 9. Thief | 24. Espionage |
| 10. Join | 25. Belfry |
| 11. Brave | 26. Rivalry |
| 12. Diamond | 27. Amendment |
| 13. Gamble | 28. Compel |
| 14. Nonsense | 29. Affliction |
| 15. Prevent | 30. Obliterate |

16. Contagious

31. Imminent

17. Nuisance

32. Dilatory

Sub-test 8 is "Object Assembly." The materials are four sets of cardboard pieces, each in a separate box. The pieces in each box fit together to make an object, like the pieces of a jigsaw puzzle. The entire sub-test is given to all children. Each item has a time limit ranging from 120 seconds for Item 1 to 180 seconds for Item 4. The score ***845** for each item is based upon the number of pieces correctly joined plus time bonuses for completion within certain intervals less than the time allotted.

The test starts with a sample item, a five-piece assembly of an apple. The examiner shows the child how to assemble the apple and then proceeds to Item 1.

Item 1 is a figure of a white girl wearing a dress. The figure is in seven pieces. One piece is the head, a second piece is the upper torso, a third piece is the lower torso and portions of the legs and the other four pieces are the remaining portions of the legs and the two arms. The examiner arranges these seven pieces in a specified configuration, shown in the manual, before the child and says to the child, "If you put these pieces together the right way, they will make a girl. Go ahead and put them together." If the child does not complete the assembly perfectly within the allotted 120 seconds, the examiner shows him the correct assembly and says, "See, it goes like this." The test then proceeds to Item 2 and no further help is given.

Item 2 is a six-piece assembly of a horse. Item 3 is a seven-piece assembly of an automobile. It is an ordinary looking sedan. Item 4 is an eight-piece assembly of the face of a white male.

The ninth sub-test is entitled "Comprehension." It consists of seventeen questions. All children begin with Item 1 and continue until they have four consecutive failures. Each item is scored two, one or zero, depending upon the quality of the response. The maximum score is 34 points.

Each question is read slowly to the child and may be repeated if there is any indication the child does not understand it.

Item 1 is, "What is the thing to do when you cut your finger?" The general criteria for a two-point answer are "put a bandage on it, cleanse it, or medicate it." Specific examples of two-point answers are: "put a bandaid on it . . . fix it up with medicine (may mention a specific one, e. g., iodine) . . . wash it with soap and water . . . stop the bleeding . . . put pressure on it . . . wrap it . . . tape (tie, patch) it up . . . let it bleed (Q) it cleans the cut." Example of one-point answers are: "tell my mother (teacher) about it . . . treat it (Q) . . . go to the doctor (hospital) . . . get it stitched up . . . soak it (Q) . . . put it under water (Q)."

The "Q" means that the examiner should ask the child to elaborate because he may be able to give a two-point answer.

We will list the remaining items and the general criteria for a two-point response, without quoting in detail the sample responses given by the manual.

Item 2 is, "What are you supposed to do if you find someone's wallet or pocketbook in a store?" The general criteria are: "try to return it to the owner, either by looking for identification of the owner or by giving it to the store owner, policeman, etc."

Dr. Williams criticizes this question because " . . . given the context today of the negative emphasis on black crime, black children pulling wallets or snatching purses, it would be suicide for a child to say, 'I would pick it up and try to find the owner.' Black kids are afraid to say that. So that the correct answer is, 'try to find the owner' and that's not a good thing for them to do in a store. They would be accused of having snatched it."

We believe this criticism may well have merit.

Item 3 is, "What should you do if you see thick smoke coming from a window of your neighbor's house?" The general criteria are: "have the Fire Department or Police Department notified. Call the fire (police) department . . . get a policeman (fireman) . . . pull down the handle in a firebox . . . report it" and "perform some action until the firemen arrive."

Item 4 is, "What are some reasons why we need policemen?" The general criteria are: "to protect people or property (by preventing a possible criminal action), to enforce laws (when a crime is occurring or has occurred)," and "to help people with non-criminal problems."

***846** Dr. Williams testified that he has had responses from black children to the effect that "we don't need policemen, we can take care of our own community, because the policemen come in here and they treat us badly."

Item 5 is, "What is the thing to do if you lose a ball that belongs to one of your friends?" The general criterion is: "replace the loss." Samples of two-point responses are: "give him (her) one of mine . . . try to get it back or replace it . . . pay for it . . . buy her a new one . . . buy another one if I can't find it." Dr. Williams commented on this question by saying that when he asked this question several years ago, "a little black boy told me, he said, 'well, I take him to the hospital.' " Dr. Williams explained that the little boy was "coming from his own cultural reference point," which is that "balls" refers to a portion of the anatomy. Dr. Williams did not state how often he had encountered this confusion among black children or whether he had any reason to believe that the term could not have the same meaning to a white child.

Item 6 is, "What is the thing to do if a boy (girl) much smaller than yourself starts to fight with you?" The general criterion is "do not fight with him (her)." Examples of two-point responses are: "just walk away . . . don't hit him, find out what's the matter . . . let him be . . ."

This "fight" question in the WISC-R is undoubtedly the most famous item in the IQ controversy. It is the item most cited by critics of the tests as an example of serious racial bias. The reason it is biased, according to Dr. Williams and some of the other authors whose articles are in evidence, is that in black communities children are taught that if anyone hits them they should hit back. Defendants point out that in the only study which has

been made of the matter, it appears that this question is of the same relative difficulty in relation to the other items on the WISC-R for white children as it is for black children. This, however, does not answer the question of whether black children answer the item incorrectly more often than white children do. According to another study, black children do in fact fail this item with about twice the frequency of white children. We believe on this basis that there is strong reason to believe the item is racially biased.

Item 7 is, "In what ways is a house built of brick or stone better than one built of wood?" General criteria are: "more durable (more permanent, sturdier); "safer;" "better insulation;" "convenient (less upkeep . . . don't have to paint brick or stone . . .)." Two points are given for a response recognizing at least two of the four general criteria and one point for recognizing any one of them.

Item 8 is, "Why is it important for cars to have license plates?" The four general criteria are: "identification of the owner of the vehicle;" "identification of the vehicle itself;" "source of income for state government;" "statistical records." Two points are given for a response recognizing any two of the four general criteria and one point for recognizing one of them.

Item 9, another two-point question, is, "Why are criminals locked up?" Five general criteria are listed: "Protection for society;" "example to others;" "punishment or revenge;" "rehabilitation;" "segregation."

Item 10 asks, "Why do we have to put stamps on letters?" The general criterion is: "to pay for the mailing of the letter." A child receives two points for any response indicating "that the stamps are like money," and one point for any response which shows an awareness that the letter cannot be delivered without a stamp even though the child does not explain the purpose of the stamp.

Item 11 is, "Why is it important for the government to hire people to inspect the meat in meat packing plants?" The general criterion is: "to protect the consumer (to assure that certain standards are met, prevent widespread disease, etc.)."

Item 12 is, "Why is it usually better to give money to a well-known charity than to a street beggar?" The three general criteria listed are: "They investigate the merits ***847** of a case (public charities give assurance that the money goes to a really needy person);" "organized charity helps more than one person;" "a more orderly way of

contribution for the donor.” Two points are given for recognizing any two of the general criteria and one point for recognizing one of them.

Dr. Williams offered some cogent criticisms of this question. He pointed out that a poor child or a child on welfare would be less likely to give to an organized charity than to a blind or crippled person he saw on the street. One of the zero responses listed for this item is, “If you give it to a beggar, he is liable to keep it himself.” Dr. Williams suggested that this is exactly what you would want the beggar to do and that such a response is not inappropriate. We are persuaded by Dr. Williams’ analysis and regard this item as particularly inappropriate for poor children or children on welfare. It is conceded that a higher percentage of black children than white are poor, or on welfare, and the item therefore discriminates against blacks.

Item 13 is, “Why is it good to hold elections by secret ballot?” The general criterion is: “so a person can vote as he chooses without fearing what others will think or do about it (without fear of public pressure).”

Item 14 is, “In what ways are paperback books better than hard-covered books?” The general criteria are: “cheaper (also more expendable);” “lighter to carry;” “greater availability (i. e., for purchase);” “easier to store.” Two points are given for recognizing any two of these general ideas and one point for one idea.

Item 15 asks, “Why should a promise be kept?” The general criterion is: “basis of faith and mutual trust, has status of implied contract.” The child receives two points for expressing the idea of mutual trust or for expressing the idea of the mutual advantage to be gained from keeping promises. One point is given for a statement recognizing a particular advantage, usually to one person, such as “so you can be trusted.”

Item 16 is, “Why is cotton often used in making cloth?” The six general criteria are: “durability;” “washability;” “cheapness (or abundance);” “coolness;” “facility in weaving (or sewing);” “takes dyes well.”

Item 17 on the Comprehension sub-test is, “What are the advantages of having senators and congressmen?” The general criteria are: “to make laws (or any specific function such as to levy taxes, vote on bills, etc.);” “the population is too large to meet as a body;” “to ensure that all parts of the country are represented;” “to act as spokesmen for the people;” “checks and balances.”

Of the seventeen comprehension items, then, five were criticized by Dr. Williams as unfair to blacks. I accept his analysis as far as Items 2 (wallet), 6 (fight) and 12 (beggar) are concerned. I am not persuaded by his views of Item 4 (policeman) and regard his concern about Item 5 (ball) as farfetched.

Sub-test 10 on the WISC-R is “Coding.” It is divided into two parts. “Coding A” is for children under eight, and “Coding B” is for children eight and older.

“Coding A” is a printed worksheet with five figures at the top. They are a star, a circle, a triangle, a cross and a square, arranged horizontally across the paper. Each figure has a symbol inside it. Inside the star is a vertical line which resembles the figure “1.” The circle encloses two horizontal lines. There is one horizontal line inside the triangle. There is a circle in the center of the cross. The square encloses two vertical lines which resemble the figure “11.”

Below these figures on the worksheet are five rows of the same figures arranged in a random sequence. However, these figures have no symbols inside them. The object of the test is for the child, using a pencil, to put the appropriate symbol from the top row into each of the figures in the bottom rows. The first half of the top test row is a “sample,” consisting of one of each of the figures a circle, a star, a square, a cross and a triangle. These, like the others in the *848 test rows, are empty. The examiner explains the test to the child, “. . . I want you to fill in the things here with the same marks they have at the top” The child then works the sample exercise and is told whether he is doing it correctly. The examiner is cautioned not to begin the actual test until the child clearly understands the task.

The child is then told to fill in as many of the remaining test items as he can until he is told to stop, without skipping any. The examiner stops the child at the end of 120 seconds. The child receives one point for each item filled in correctly. A perfect score is 45. The child can achieve as many as five bonus points for completing the test in less than the time allotted.

“Coding B” is the same general kind of test, but more difficult. Instead of figures such as squares and triangles, there are numbers one through nine. There is a symbol corresponding to each number, and the symbols are not entirely dissimilar. The symbol for the number “1,” for instance, is a horizontal line with a dot above it. The

symbol for the number “8” is a horizontal line with a dot below it.

In the section below these figures on the worksheet are four rows of the numbers one through nine arranged in random order, with a box under each number. In each box the child enters the symbol corresponding to the number. Again, the examiner explains the test to the child and some sample boxes are provided. The examiner is instructed not to proceed with the test until the child clearly understands the task. The examiner stops the child after 120 seconds. One point is scored for each item filled in correctly. There is no penalty for poor drawing. The item is scored as correct if it is “clearly identifiable as the keyed figure, even if it is drawn imperfectly” There are 93 items in Coding “B”, and the maximum score is therefore 93.

Sub-test 11 of the WISC-R is entitled “Digit Span.” It tests the ability of the child to repeat numbers after the examiner. “I am going to say some numbers. Listen carefully, and when I am through say them right after me.” There are two parts to this sub-test, “Digits Forward” and “Digits Backward.” In the former, the child repeats the numbers in the order the examiner says them, and in the latter the child is told to repeat the numbers in reverse order.

Each item consists of two “trials.” For instance, the two trials of Item 1 in “Digits Forward” are “3-8-6” and “6-1-2.” There are seven items in each of the two tests, and the items become longer by one digit in each trial. Item 7 in the “Digits Backward” test, for instance, consists of these two trials: “6-9-1-6-3-2-5-8” and “3-1-7-9-5-4-8-2.”

The maximum score on each test is 14 points. The child receives two points if he passes both trials of the item and one point if he passes only one trial. The test is discontinued as soon as the child fails both trials of any item.

The twelfth and final sub-test on the WISC-R is “Mazes.” This consists of nine drawings contained in a worksheet (the same worksheet that contains the “Coding” sub-test). Each drawing consists of a series of concentric squares, with gaps in the lines, so that one can, with varying degrees of concentration, trace ones way from the center of the drawing to a point outside the drawing by finding the openings in the lines. At the center of each drawing is the figure of a boy or a girl, and the task is to lead the boy or girl out of the maze, using a red pencil. The child is

told to keep his pencil point on the paper and to avoid touching any lines. “You’re not allowed to go through a wall.”

There is a sample maze which the examiner uses to demonstrate the test to the child. “See this boy in the middle here? (point) He wants to get out to the street there (point). Let me show you how he could do it without getting stuck. Watch me. . . . No, not this way. You see, if he took this turn he would get stuck by the blocked road. He can’t go through a wall (point). He must go this way to get out.”

***849** The first maze is very simple, and, as is true of the items on all the sub-tests, they become progressively more difficult. Items 1 through 7 are reasonably easy; Items 8 and 9 are formidable.

Children ages 6 to 7 begin with Maze 1, children ages 8-16 begin with Maze 4. The scoring is somewhat complicated. There is a time limit for each maze. Numbers 1 through 4 are 30 seconds. Five is 45 seconds, 6 is 60 seconds, 7 and 8 are 120 seconds, and 9 is 150 seconds. The points awarded for each maze are a function of time and error. A perfect score, up to 5, is awarded for solving the maze within the allotted time with no errors. (An error is defined as entrance into a blind alley.) Partial credit is given for solving the maze within the allotted time, albeit making errors along the way.

The test is discontinued after the child has two consecutive failures that is, where he has failed to achieve any points at all on two consecutive mazes.

The maximum score is 30 points.

This completes the description of the sub-tests of the WISC-R. It will now be helpful to discuss briefly how the test was devised. The information about the test is derived from the WISC-R Manual, and the parties do not dispute the accuracy of that information.

As indicated by the “R,” this test is a revision of an earlier test, the simple “WISC.” It was an attempt to update the earlier test and to eliminate some deficiencies which the test authors acknowledged in their earlier efforts. We will discuss the WISC itself in some detail at a later point in this opinion, since it is still used by the defendants and is one of the tests challenged by the plaintiffs.

The WISC-R was administered in 1970 to 2,200 subjects, ages 6 1/2 to 16 1/2. There were eleven age groups, with

200 children in each group. Children of minority groups were included in the same proportion as minorities appeared in the 1970 census. Three hundred five black children were included in the sample of 2,200. Half of the children were male, half were female. The subjects were drawn from five categories, depending upon the education of the father or head of the household. The percentages were the same as were reflected in the 1970 census. The 2,200 subjects were drawn proportionately from the four geographical areas of the 1970 census, with the same distribution of urban and rural children as reflected in the census. Only normal children were used, except that children who were only suspected of mental deficiency were not excluded. Two hundred two different examiners administered the sample tests to the 2,200 subjects.

With this mix of children, deemed by the test authors to be representative of the general population, the test questions were administered for the purpose of determining the range of performance in each of the age groups. A statistical profile, resembling the familiar bell-shaped curve, was established for each age group. Mean scores and standard deviations were established on each of the sub-tests for each of the age groups. A child who takes the WISC-R achieves a score which is compared to the scores achieved by the other children in his age group. The total points achieved by the child on each of the sub-tests are added and then converted to a scaled score. This is the IQ score. An IQ of 100 is the mean. The standard deviation is 15. About two-thirds of all children achieve IQ scores between 85 and 115 that is, between one standard deviation below and one standard deviation above the mean. About 95 per cent score in the 70 to 130 range, two standard deviations above and below the mean.

The WISC-R has not totally displaced the earlier WISC test, at least as far as the Chicago School System is concerned. For reasons which are not entirely clear in the record, some children, without regard to race, are given the WISC rather than the WISC-R. The two tests combined are given to a total of 69 per cent of the children who are given intelligence tests, with the WISC-R being used in the majority of these instances. The revised Stanford-Binet, Form L-M, the third test under attack *850 in this case, is given to 19 per cent of the children tested.

WISC

WISC-R

There is a temptation to shorten this opinion by saying that the WISC is substantially similar to the WISC-R, so that whatever conclusions one draws as to the racial bias of one will apply to the other. The temptation is particularly strong because the parties themselves have made no effort to distinguish between the two tests. But as I have indicated, I believe the persuasiveness of judicial opinions in this area must depend upon exposition and analysis of the test materials, not upon bald pronouncements or conclusions. Where the WISC is identical to the WISC-R, I will so indicate. Where there are differences, I will describe the actual test items of the WISC.

THE WISC

The WISC test was published in 1949. It was standardized on 2,200 children, with 200 in each age group, in much the same manner as the WISC-R. The 1940 census was used as the statistical base and the sample was drawn so as to reflect the general population in terms of geographic areas, urban versus rural, and parental occupation. The manual notes, without comment, that "only white children were examined."

The WISC is divided into the same twelve sub-tests as the WISC-R. The general instructions to the examiner as to how to give the tests are in all significant respects of the same tenor as those in provided with the WISC-R. In describing the various sub-tests, it will generally be unnecessary to refer to the instructions concerning administration or scoring.

The first sub-test on the WISC is entitled "General Information." It consists of thirty items, as did the Information sub-test on the WISC-R. The following items on the two sub-tests are identical:

1

2

2

1

3

3

5

4

7

5

8

7

9

12

(Who discovered America?)

10

10

(Where does the sun set?)

11

11

13

14

14

13

(What does the stomach do?)

15

18

19

24

(How tall is the average American man?)

21

20

22 23

23 30

27 25

28 28

In addition, Item 20 on the WISC, “Where is Chile?,” was replaced by the less ambiguous, “In what continent is Chile?,” as Item 21 on the WISC-R. The remaining eleven items on the WISC were not carried over into the WISC-R. They are as follows:

Item 4 asks, “From what animal do we get milk?” The answer is: “cow(s); goat(s).” (On the WISC-R, Item 9 is, “From what animal do we get bacon?”).

Item 6 on the WISC is, “In what kind of a store do we buy sugar?” The answer is: “grocery store; food store; name of local store like A & P.”

Number 12 on the WISC is, “What is the color of rubies?” The answer is: “red; maroon.” Dr. Williams criticized this item as confusing to black children because “Ruby” can be a woman’s name. Dr. Williams testified that he had a little boy say, “Well, she’s black.” Dr. Williams commented he felt this was a “very creative response,” and that it was not appropriate to score that response a zero. It is doubtful that such an answer would be evaluated as indicating mental deficiency. The WISC manual (p. 18) cautions the examiner that the responses of the subject should be recorded exactly as they are given and states that “. . . a full recording of the subject’s answers permits a subsequent evaluation and a fuller consideration of them in comparison with responses obtained on other tests.” Thus, while the test item might be a poor one (it was not repeated on the WISC-R), I am not persuaded that a black child who misunderstands the

question would be penalized. If the more complete instructions in the WISC-R manual were applied to the WISC where appropriate, the chance of a mistaken *851 evaluation lessens. The WISC-R Manual (p. 60) indicates that the examiner should ask the child to elaborate upon an ambiguous answer and also to give full credit to any response that is “equal to or better than” the sample answers.

Notwithstanding these observations, we will not discount Dr. Williams’ suggestion that a black child could be penalized by Item 12. Item 16 on the information sub-test is, “Who wrote ‘Romeo and Juliet?’ ” The answer: “Shakespeare; Tchaikowsky.”

Item 17 is, “What is celebrated on the Fourth of July?” Answer: “American independence from England.”

Item 18 is, “What does C.O.D. mean?” The answer is given as “cash or collect on delivery (or a correct description of the process.)” Dr. Williams testified that, “. . . there are a lot of black kids who have not had the exposure and the experience to the word C.O.D., but it does not mean that black kids cannot codify abbreviations and symbolize abbreviations at a higher level.”

Item 20 on the WISC is, “Where is Chile?” The answer: “South America.” This same item appears on the WISC-R as No. 21. It may have been found that there was some ambiguity in the question, since the WISC-R item asks, “In what continent is Chile?” rather than simply, “Where is Chile?” Dr. Williams pointed out in connection with the

WISC item that a little boy told him, “Well, it’s at home on the shelf”, obviously thinking the question referred to the soup known as chile. This seems like the misunderstanding about “rubies.” When it is obvious that the child misunderstands the question, the competent examiner follows up with an additional question.

Item 24 asks, “How far is it from New York to Chicago?” Any answer from 800 to 1,000 miles is scored correct.

Item 25 is, “When is Labor Day?” The answer should indicate that it is the first Monday in September.

Item 26 is, “Who discovered the South Pole?” The answer: “Amundsen.”

Item 29 is, “Who was Genghis Khan?”. The answer: “a Mongol Conqueror of North China.” Dr. Williams criticized this item on the basis that “my knowledge of Genghis Khan has nothing to do with where I live my environment, so I would disqualify that as a question of intelligence.” We believe Dr. Williams interprets too literally Wechsler’s definition of intelligence as an ability to cope with one’s “environment.” The word “environment” in this context does not denote only one’s immediate physical surroundings. It means the milieu one is called upon to deal with, which, in the case of a school

child, includes the demands of the classroom as well as those of the playground and the home neighborhood. The ability to acquire and assimilate information certainly has something to do with one’s ability to cope in life. moreover, it should be noted that the question concerning Genghis Khan is the twenty-ninth of thirty items on the WISC Information sub-test. It is obviously an item considered by the test authors to be very difficult, one designed to test the upper reaches of the abilities the test purports to measure. The relevance of this particular item to the placement of children in classes for the mentally retarded seems nil.

Item 30 on the sub-test is, “What is a lien?” The answer given is “legal claim on property as a security for a debt or charge.”

We turn now to the second sub-test on the WISC, “General Comprehension.” This consists of fourteen items as opposed to the seventeen items on the WISC-R Comprehension sub-test. The following WISC items are identical to WISC-R items:

WISC

1

2 (friend’s ball)

4 (fight)

6

WISC-R

1

5

6

7

7	9
10 (beggar)	12
12	16
13	17
14	15

The following are the WISC items which are not repeated on the WISC-R:

***852** Item 3 asks, “What would you do if you were sent to buy a loaf of bread and the grocer said he did not have any more?” The general criterion is: “go to another store for it.” Dr. Williams indicated that “. . . the correct response is really culturally determined, because black kids used to tell me, ‘Well, I go back home, because my mama told me don’t be foolin’ around on the street, that if I go to the store, don’t get lost, don’t go any other place, because I’m going to beat you.’ ”

Dr. Williams was not cross-examined as to how frequently he encountered such a response, nor the particulars of any specific testing situation. Generally, this was true as to all of his testimony concerning the responses he claims to have obtained from black children on the specific test items. It would have been helpful to the court if plaintiffs had produced the actual scoring sheets used in tests given black children. These sheets would have shown the verbatim responses of the children. No problem of confidentiality prevented the use of actual test papers, since the names of the children involved would simply have been blocked out. As an example of how easily this could have been done, the defendants produced for plaintiffs’ use the records pertaining to 6,000 children in the Chicago school system who were tested for retardation. This was done pursuant to a

protective order which provided for blocking out the names of the children and otherwise preserving the confidentiality of the information. The production of that kind of evidence would have been far preferable to these almost casual recollections of Dr. Williams, the accuracy of which has to be taken on blind faith.

The other items appearing on the WISC “General Comprehension Sub-Test” which are not repeated in the WISC-R are Items 5, 8, 9 and 11. No witness commented about any of these items.

Item 5 asks, “What should you do if you see a train approaching a broken track.” The general criterion is: “Give appropriate warning to the approaching train.” Two-point answers are those which suggest signaling the train, such as waving a handkerchief or something bright. One-point answers involve something more indirect, such as “tell the man in the station, and he’d stop the train. . . .”

Item 8 asks, “Why should women and children be saved first in a shipwreck?” General criteria are: “women more necessary for the care of children; children have a longer life ahead than adults; women and children are not as strong as men.” Two-point responses recognize at least two of the above; one-point responses recognize at least one.

Item 9 is, “Why is it better to pay bills by check than by

cash?" General criteria are: "the returned check is a record of payment; it is safer; more convenient." Two-point responses recognize at least two of the above ideas; one-point responses recognize at least one.

It appears to me that Item 9 is subject to the same criticism as Item 18 of the General Information sub-test ("C.O.D.") as far as the experience of many black children is concerned.

Item 11 asks, "Why should most government positions be filled through examinations?" General criteria are: "To get better qualified and trained people; set standards; reduce political favoritism; prevent nepotism." Two-point responses recognize at least two of the above, one-point responses include at least one.

The third sub-test on the WISC is "Arithmetic." There are sixteen items, with one point for each. The test is discontinued after three consecutive failures. Items 1 through 13 are read to the subject. Items 14-16 are presented on separate cards for the subject to read. There is a time limit for each problem. The limit is 45 seconds for Items 1-3, 30 seconds for Items 4-11, 60 seconds for Item 12, 30 seconds for Item 13, 60 seconds for Item 14, and 120 seconds each for Items 15 and 16.

In Item 1, the examiner places nine blocks in a row before the child and says, "Count these blocks with your finger." Item 2 asks the child, "Now take away all *853 of the blocks except four. Leave four blocks for yourself."

Item 3 is similar to Item 2, except that the child is asked to take away all blocks except seven.

Item 4 is, "If I cut an apple in half, how many pieces will I have?"

Item 5 is, "John had four pennies and his mother gave him two more. How many pennies did he have altogether?"

Item 6 is, "James had 8 marbles and he bought 6 more. How many marbles did he have altogether?"

Item 7 is, "A boy had 12 newspapers and sold 5. How many did he have left?"

Item 8 is, "At 7 cents each, what will 3 cigars cost?"

Item 9 asks, "A milkman had 25 bottles and sold 11 of them. How many bottles did he have left?"

Item 10 is, "Four boys had 72 pennies. They divided them equally among themselves. How many pennies did each boy receive?"

Item 11 is, "A workman earned \$36.00; he was paid \$4.00 a day. How many days did he work?"

Item 12 reads, "If you buy 3 dozen oranges at 30 cents a dozen how much change should you get back from \$1.00?"

Item 13 is, "Thirty-six is two-thirds of what number?"

Item 14 is, "If 3 pencils cost 5 cents, what will be the cost of 24 pencils?"

Item 15 is, "If a taxi charges 20 cents for the first quarter mile and 5 cents for each quarter mile thereafter, what will be the fare for a two-mile trip?"

Finally, Item 16 asks: "Smith and Brown start a card game with \$27.00 each. They agree that at the end of each deal the loser shall pay the winner 1/3 of what he (the loser) then has in his possession. Smith wins the first three deals. How much does Brown have at the beginning of the fourth deal?"

No witness commented on the "Arithmetic" sub-test.

The next WISC sub-test is "Similarities." The first part of the test consists of four questions entitled "Analogies." These four questions are given to children under eight years or subjects suspected of mental deficiency. Each of the four items in "Analogies" is read to the subject and he is to complete the item:

1. "Lemons are sour but sugar is _____."
2. "You walk with your legs and throw with your _____."
3. "Boys grow up to be men and girls grow up to be _____."
4. "A knife and a piece of glass both _____."

If a subject passes two of the four "Analogies" items, he proceeds with Similarities, which are Items 5 through 16 of the test. He continues until he has three consecutive zero responses. Responses are scored two, one or zero, depending upon the degree and quality of the generalization. The examiner presents each item by asking, "In what way are a _____ and a _____

alike?”

Item 5 is “plum-peach.” Two points are given for a response stating that they are both fruits. One point is given for a response saying they are both food, both round, both have a skin or some other specific characteristic.

Item 6 is “cat-mouse.” Two points are given for a response stating they are both animals, mammals or creatures. One point is given for a response indicating they both have four legs, eyes, that they both eat, or any other specific common characteristic.

The remaining items in the sub-test are:

7. Beer-wine
8. Piano-violin
9. Paper-coal
10. Pound-yard
11. Scissors-copper pan
12. Mountain-lake
13. Salt-water

4. Letter

6. Cushion

9. Fur

12. Spade

13. Sword

17. Hero

19. Nitroglycerine

14. Liberty-justice

15. First-last

16. The numbers forty-nine and twenty-one.

***854** The maximum score on the sub-test is 28 points. As will be noted, some of the WISC items were repeated on the WISC-R.

No witness referred to this sub-test on either the WISC-R or the WISC.

The next sub-test is “Vocabulary.” It is administered in the same manner as the vocabulary test on the WISC-R. It contains forty words rather than the thirty-two words on the WISC-R. Twenty of the WISC vocabulary words are repeated in the WISC-R. They are: Bicycle, knife, hat, umbrella, nail, donkey, diamond, join, nuisance, brave, nonsense, gamble, fable, belfry, espionage, stanza, seclude, affliction, mantis and dilatory.

The remaining twenty WISC words that were not used in WISC-R (with WISC item number indicated) are:

28. Hara-kari

29. Recede

31. Ballast

32. Catacomb

33. Imminent

35. Vesper

36. Aseptic

20. Microscope

21. Shilling

27. Spangle
37. Chattel

39. Flout

40. Traduce

There was no reference in the testimony to any of the WISC vocabulary items other than “umbrella.”

The next sub-test is “Digit-Span,” consisting, as in the WISC-R, of “Digits Forward” and “Digits Backward.” The fourteen items on this sub-test are identical to the fourteen items on the WISC-R “Digit Span” sub-test. The only difference is in the method of scoring. On the WISC-R, the subject is given either two points, one point or zero points for an item, depending upon whether he correctly repeats both trials, one trial or neither trial of the item. On the WISC, the subject is given a number of points for each item equal to the number of digits he repeats correctly on either trial of the item. This

WISC

- 1 (comb)
- 3 (fox)
- 5 (cat)
- 6 (door)

difference in scoring does not seem material to the question of racial bias.

The next sub-test on the WISC is “Picture Completion.” There are some differences between this sub-test and its WISC-R counterpart. The WISC contains twenty items, the WISC-R twenty-six items. The child is given 15 seconds to respond to each item on the WISC, 20 seconds on the WISC-R. Some of the items on the two are different, some are the same. The following items are identical:

WISC-R

- 1
- 3
- 5
- 13

9 (scissors)	18
12 (screw)	20
16 (thermometer)	22
18 (umbrella)	26
19 (cow)	21
20 (house)	23

Four other items on the two tests are substantially identical. Item 4 on the WISC shows a white woman missing a mouth whereas Item 2 on the WISC-R shows a black woman missing a mouth. Item 8 on the WISC shows a seven of diamonds missing one of the diamonds, whereas Item 14 on the WISC-R shows a five of diamonds missing one of the diamonds. Item 10 on the WISC shows a 1940's style jacket missing its buttonholes, whereas Item 16 on the WISC-R shows a modern jacket missing the buttonholes. Item 15 on the WISC, showing a male profile without an eyebrow, is repeated as Item 25 on the WISC-R, but the male has a modern hair style rather than the shorter hair of the 1940's.

Picture completion Items on the WISC which were not repeated on the WISC-R are the following: Item 2 shows a table having only three legs. Dr. Williams testified that this item is racially biased because

Now, in some places there are so many broken things around, lying around, that folks just don't sensitize themselves to the fact that a leg is missing, because of

the economic conditions.

So, if a child says nothing is missing, he may be just reflecting what he's come out of.

He went on to explain that a three-legged table can be used "innovatively" by placing *855 it in a corner against a wall, so that the third leg is not really necessary and is forgotten.

Item 11 is a fish missing its dorsal fin; Item 13 is a fly missing its antennae; Item 17 is a fedora hat without a hatband. There was no testimony directed to these items.

"Picture Arrangement" is the next sub-test on the WISC. It is similar and in some respects identical to its counterpart on the WISC-R. The items which are identical (or substantially so, with only changes in such things as dress or hair style, reflecting the passage of time) are:

WISC

WISC-R

D (scale)

sample

Demonstration (Fight)

1

1 (fire)

3

2 (burglar)

5

4 (picnic)

2

5 (sleeper)

6

6 (gardener)

10

7 (rain)

12

There are four arrangements in WISC which are not in WISC-R. Three of them are sample arrangements. The first of these, the first item administered to any subject, consists of three cards which, when placed side-by-side in the correct order, show a picture of a dog. One card shows the head and one of the front legs; another card shows the mid-section and two more legs; the third card shows the rest of the dog, including the tail.

The second sample item, entitled "Mother," consists of three cards which, when placed in proper sequence side-by-side, show a woman and a girl playing with a toy train.

The third sample item, "Train," consists of four cards which when correctly placed show a locomotive pulling two cars.

The scoring system in the WISC is similar to the WISC-R, with bonus points being given for completion of the items in less than the allotted time.

As indicated in connection with our discussion of the WISC-R picture arrangement sub-test, Dr. Williams criticized two of the WISC items which appear on the WISC-R. These are the items entitled "Fight," and "Gardener." This was Dr. Williams testimony about these two items:

Now, the first three or four, I don't think that they're in any particular problem with those. But in the one on fight, f-i-g-h-t, it depends on the logic of what the child is using. And as someone said, you cannot have no there is no truth, and can one have a half truth to any of these particular items, and in the fight one, a child can indicate, "Someone else started the fight," rather than the way it actually turns out.

And if he has a creative story, he's not given credit. In other words, there is a way, a particular way for this individual to have it to have his own creative way might not even be counted.

And the same is true for the gardener, No. 6. The people have their own creative ways, and I'm saying that on a test like this where you have a fixed way of doing it, creativity is discouraged; there is no range, or there is no freedom, here, for having your own individual story.

And we know that if you come out of a different society, you make up your stories differently.

For example, on the Thematic Apperception Test, I might look at a picture and see one thing and someone else might see something else; my creativity should be given the same measure, but they do not permit that on this particular instrument.

Defendants did not cross-examine Dr. Williams on these or any other test items he testified about. At the time the testimony was given, the picture arrangement cards were not available in court. Plaintiff had offered in evidence copies of the WISC and WISC-R manuals, which contain the test instructions and many of the test items along with the answers. However, some things such as the picture arrangements and the object assemblies are contained in separate boxes or booklets or on separate cards. Plaintiffs had no intention of offering these materials in evidence nor did the defendants, and it was necessary for me to

*856 ask that they be produced. Plaintiffs' counsel did not have them available, and I am satisfied that neither they nor defendants' counsel had ever seen them. Defendants managed to round up complete kits containing all of the materials for the three IQ tests in question, and they were finally offered in evidence long after Dr. Williams had testified. My first opportunity to study these test materials has been during the period since the close of the evidence.

Comparing the actual picture cards in these two items to the analysis made by Dr. Williams, I am unable to follow him. The "fight" arrangement consists of three cards which, to me, have no possible logical sequence other than the one prescribed by the test authors. One card shows two unmarked boxers, one in black trunks and one in white trunks, sparring in the ring. Another card shows the boxer in black trunks unmistakably getting the worst of it, being pummeled by his opponent in the white trunks and sagging to his knees. Another card shows the boxer in the black trunks being carried away from the ring while, in the background, the boxer in the white trunks stands in the ring with his arms upraised in triumph.

Dr. Williams believes that "... there is no truth, and one can have a half truth to any of these particular items, and in the fight one, a child can indicate, 'someone else started the fight,' rather than the way it actually turns out." He goes on to say that if a child has "a creative story" at odds with the view of the test publishers, it "might not even be counted."

Clearly, these cards do not present any question as to who started the fight. I am inclined to believe that Dr. Williams simply did not recall what these cards were about. Moreover, if a child could indeed arrange these cards in a sequence other than the prescribed one and still tell "a creative story," it should be remembered that the test instructions admonish the examiner to give credit for answers which are "equal to or better than" the prescribed ones.

The "gardener" story is told on six cards. In the WISC version, one card shows a woman and a boy standing in the garden. The woman is pointing to the ground, indicating that the boy, who holds a hoe in his hand, should begin work. Another card shows the boy working in the garden. Another card shows the boy unearthing a worm with the hoe. Another card shows the boy with a can and a fishing pole, leaving the garden. Another card shows the boy walking down a lane, happily carrying his fishing pole and worm can. Another card shows the boy on the bank of a creek, fishing. The test publishers in the

WISC version give full credit if the subject arranges the cards either in the order just described or if the order of the second and third cards is switched. The WISC-R version involves only five cards. Card No. 2 of the WISC (showing the boy simply working in the garden) was dropped. Card No. 2 of the WISC-R shows the boy unearthing the worm, and Card No. 3 shows him stealing away with the fishing pole and the can. Three points, plus any time bonus points earned, are given for the correct arrangement in the WISC-R, but two points, with no time bonus, are given for a version which reverses Cards 2 and 3.

It would have been helpful had Dr. Williams spelled out the other “creative ways” these cards could be arranged in the gardener story. His oblique implication that there are such creative alternatives to the sequences prescribed by the test publishers is not persuasive.

The next WISC sub-test is “Block Design.” It differs from the WISC-R in that it has one less item (ten instead of 11). The blocks in the WISC are 4-color (red, white, blue and yellow) whereas in the WISC-R they are 2-color (red and white). Nine of the ten WISC items are repeated on the WISC-R. One hundred fifty seconds was allowed for each item on the WISC, and that was reduced to 120 seconds on the WISC-R. The WISC-R also changes some of the rules for bonus points.

Next on the WISC is the “Object Assembly” sub-test, consisting of four items. They are similar to four of the five items on the WISC-R sub-test.

***857** Item 1 is the figure of a girl in five pieces. (Item 1 of the WISC-R was a girl in seven pieces.)

Item 2 is the identical six-piece assembly of a horse that is Item 2 on the WISC-R.

Item 3 on the WISC is identical to Item 4 on the WISC-R, an eight piece assembly of a male face.

Item 4 on the WISC is a seven-piece assembly of what looks like a 1939 automobile, whereas the counterpart on the WISC-R, Item 3 is a seven-piece assembly of a later model.

The “Coding” sub-test on the WISC is identical to the Coding sub-test on the WISC-R. All items were simply carried over into the later test.

The final test on the WISC, “Mazes,” is almost the same

as the WISC-R. The WISC consists of eight items. All eight were carried over to the WISC-R, and a new one was added for a total of nine on the later test.

This completes the description of the Wechsler Intelligence tests. We will now turn to a description of the third test in question, the Stanford-Binet.

THE STANFORD-BINET

This test attempts to ascertain a child’s “mental age” compared to his chronological age. The sub-tests have been standardized for different age groups. A child who performs below the mean for his age group is considered to have a “mental age” below his chronological age, and one who exceeds the mean performance of his age group has a mental age above his chronological age. Performance equal to the mean for the age group is given the numerical IQ score of 100. An 8-year old child who scores at the 10-year old average level would have an IQ of 125, whereas an 8-year old who performed at the 6-year level would have an IQ of 75.

The test items have been revised from time to time, most recently in 1960. The 1960 version, the one currently in use, is based upon a 1937 standardization, brought up-to-date by a sampling of tests administered from 1950 to 1954.

The 1937 standardization group consisted of 3,184 native born white subjects. There were approximately 100 subjects at each half year interval from age 11/2 to 51/2 years, 200 at each age level from 6 to 14, and 100 at each age from 15 to 18. Subjects were limited to those who were within one month of a birthday or the half year point. Each age group was half male and half female. The subjects were taken from 17 urban, suburban and rural communities in eleven widely separated states.

In 1972, the scoring norms of the tests were revised based on samples which for the first time included black and Hispanic children.

It is generally agreed that there is a high correlation between the results on the WISC tests and those on the Stanford-Binet. A child’s success on one will be substantially the same as his success on the other.

The Stanford-Binet manual instructs the examiner to record the child's responses as nearly verbatim as possible. The test should be commenced at a point where the child is likely to succeed, but with some effort. For children who appear to be normal, it is recommended that they begin with the test one age level below their chronological age.

A child will typically have successes and failures over a wide range of sub-tests. The examiner determines the child's "basal" and "ceiling" levels. The former is the age level just below that at which he first fails an item. The ceiling is the age level at which the child fails all items. The mental age of the child is computed by taking his basal age and adding to it all credits earned for items beyond the basal level.

The test begins with the sub-test for 2-year olds. Item 1 is "Three-Hole Form Board." This consists of a small green board and three red blocks. One block is round, one is square and the third is triangular. They fit into holes of corresponding shapes in the board. The child is shown the board with the blocks in place. The examiner ***858** says, "Watch what I do," and removes the blocks, placing them on the table near the board. The examiner then says, "Now put them back into their holes." The child is allowed two trials. The child scores one point if he correctly replaces the blocks on either of the two trials.

Item 2 is "Delayed Responses." The materials are three small boxes and a small toy cat. The boxes are placed in a row about 2 inches apart. The examiner says, "Look, I'm going to hide the kitty and then see if you can find it again." While the child watches, the examiner places the cat under one of the boxes, waits 10 seconds and then has the child find the kitty. This is done three times, once for each box. The child scores two points if he finds the kitty on all three trials.

Item 3 is "Identifying Parts of the Body." The child is shown a large paper doll and is asked to point to the doll's hair, mouth, feet, ear, nose, hands and eyes. The item is worth four points.

Item 4 is "Block Building: Tower." The materials are twelve small wooden cubes. The examiner stacks four of the cubes on top of each other and says, "You make one like this." The examiner's tower is left standing as an example while the child builds his own.

It is appropriate to inquire how the examiner motivates a

2-year old to participate in the test and how one distinguishes between lack of ability and simple lack of interest or cooperation. The test publishers recognize these problems and emphasize the necessity for building rapport with the child before the test begins. A number of specific suggestions are offered as to how to win the child's cooperation and avoid distractions which would interfere with his concentration.

Item 5 for 2-year olds is "Picture Vocabulary." These are eighteen small cards with pictures of common objects. The child is shown the cards one at a time by the examiner who says, "What's this? What do you call it?" The objects are: An airplane, a telephone, a man's hat, a baseball, a tree, a key, a horse, a table knife, a jacket, a ship, an umbrella, a human foot, a flag pole, a walking cane, a man's arm, hand and shoulder, a jackknife with blades extended, a pitcher and a tree leaf.

The manual provides examples of acceptable responses. A "plus" response for the airplane, for instance, would include "plane, flying 'chine, airship, jet." A plus response for the umbrella would include " 'brella. Parasol. Sunshade." (Manual, pp. 127, 128). In connection with all of the tests, the examiner is instructed that he ". . . must free himself from purist prejudices regarding ungrammatical speech and be as willing to credit a correct response awkwardly expressed as one that is given in the best literary form. It is necessary to be very alert in order to judge whether the poorly formulated answer really carries the correct meaning." (p. 55)

Item 6 is "Word Combinations." The examiner listens for any spontaneous two-word expressions by the child at any time during the course of the interview. A plus response is any appropriate combination of two words, such as "see kitty" or "all gone."

As I understand the instructions in the manual, the scoring for these items is all or nothing. On the "Picture Vocabulary" item, for instance, the child must correctly identify all eighteen objects or he receives no points. I have some difficulty understanding the rationale for denying any credit at all to a child who can name, say, sixteen of the eighteen items. It is not explained in the manual. It may be that any unfairness in this approach would penalize blacks no more than whites. On the other hand, if the item is unfair to begin with, so that black children have less chance than white children of naming all eighteen objects correctly, the all or nothing system would heighten the discrimination against blacks. This discussion is somewhat academic, however, because it

appears to me that each of the eighteen items is culturally fair. The only one questioned by any witness is the umbrella. Dr. Williams testified that this is commonly called a parasol in the black community. What he neglected to point out is that “parasol” is one of the acceptable responses listed in the manual.

***859** The next sub-test is for children aged 2 years and 6 months.

Item 1 is “Identifying Objects by Use.” It consists of a card to which are attached a U.S. penny and five miniature items a cup, a shoe, a knife, an automobile, and an iron. The child is asked, “Show me what . . . we drink out of . . . goes on our feet . . . we can buy candy with . . . we can cut with . . . we ride in . . . we use to iron clothes.” The child designates the object by pointing. Item 1 is worth three points.

Item 2 is the same as Item 3 for 2-year olds identifying parts of the body. Item 2 is worth six points.

Item 3 is “Naming Objects.” This consists of more miniature items handed to the child one at a time, with the question, “What is this? What do you call it?” The items are a chair, an automobile, a box, a key, a fork, and a flag. Item 3 is worth five points.

Item 4 is the same as Item 5 on the 2-year test “Picture Vocabulary.” The score here is eight points.

Item 5 is “Repeating 2 Digits.” The examiner says a two-digit number and the child repeats it after him. There are three trials. The digits are 4-7, 6-3 and 5-8. The item is worth one point.

The final item on the 2 1/2 year old test is “Obeying Simple Commands.” The materials are a block, a button, a toy dog, a small box and a pair of scissors. These objects are placed on the table in front of the child and the child is asked to perform the following tasks:

- A. “Give me the dog.”
- B. “Put the button in the box.”
- C. “Put the scissors (shears) beside the block.”

The item is scored two points.

The next sub-test is for 3-year olds. Item 1 is “Stringing Beads.” The materials are a box of 48 blue beads and two

shoestrings. Sixteen of the beads are square, 16 are round and 16 are cylindrical. The examiner begins to string the beads, asking the child to do the same thing. The child gets credit for the number of beads he strings in the time limit of two minutes. There is no requirement that he string beads of a particular shape.

Item 2 is “Picture Vocabulary” the same as Item 5 on the 2-year old test.

Item 3 is “Block Building: Bridge.” The materials are twelve one inch cubes which are placed before the child. The examiner then takes three of the cubes and constructs a “bridge” consisting of two blocks on the bottom and one on top. The examiner then says, “See if you can make one like this.” The child is given credit for any similar structure which stands.

Item 4 is “Picture Memories.” The materials are four cards with animal pictures. The first card shows a picture of a cow. The child is shown this card and asked what it is. If he does not name it correctly, he is told what it is. Then this first card is removed and the child is told, “Now we are going to find it.” He is then shown a second card containing pictures of a variety of objects, including the cow. The cow is in the center of the card. The other objects are a dog, a flat iron, an umbrella, a chest of drawers, a teapot, a pair of scissors, a squirrel, a comb, a cat, a key and a clock. The child is asked, “Where is it?”

The next part of the item is to show the child a card containing three figures, a small bird, a rooster and a duck. He is asked what they are, and if he does not know, he is told. Then the card is removed and he is told, “Now we are going to find them.” He is then shown the final card, which, in addition to the bird, rooster and duck, contains the following figures: a half moon, a bicycle, a bucket, a spoon, an open book, an American flag, an ear of corn, a drum and a wagon.

The child receives one point for a correct response to all parts of the item.

Item 5 for 3-year olds is “Copying a Circle.” The child is shown a circle printed on a card and is asked to “make one just like this. Make it right here.” He uses a pencil and makes his effort in a space provided on the score sheet used for the test. He is ***860** given three trials, with the directions being repeated each time. Success results in one point. Any figure which shows a rotary movement and is “approximately” round is acceptable. The circle need not be closed.

Item 6 on the 3-year old test is "Drawing a Vertical Line." The examiner draws a vertical line and says to the child, "You make one like this. Make it here." The child has only one trial. Any mark which is "approximately" vertical will do.

The next sub-test is for 3 1/2 year olds. Item 1 is "Comparison of Balls." The child is shown a card with two printed circles on it. One is approximately 1 1/2 times the size of the other. The child is asked, "Which ball is bigger? Put your finger on the big one." The child is asked to do this three times, with the card in a different position each time. If the child fails on one of the first three trials, he is given three additional trials. He is given credit if he responds correctly on three trials or in five of six trials.

Item 2 is "Patience: Pictures." The materials are two sets of cards. The first set of cards is a white background with a black ball in the middle; the card has been vertically cut in half. The two halves are then placed before the child with the white sections touching that is, with the halves of the ball to the outside. The child is told, "Put these two pieces together and make a ball."

The other two cards are a vertically bisected drawing of a pig. The child is told, "Put these two pieces together and make a pig." Credit is given if the arrangement of cards indicates that the child is trying to make the correct picture and the cards are fairly well in alignment. The cards need not touch.

Item 3 on the 3 year old test is "Discrimination of Animal Pictures." The materials are two cards with pictures of animals. The same animals are shown on each card, but in a different order. At the bottom of card "A" there is a large rectangular hole. The examiner places this hole over each of the animals on card "B" and asks the child to find that animal on card "A." "Find me another one just like this up here." The animals shown on the cards are a cat, a pig, an elephant, a rat, a cow, a squirrel, a camel, a goat, a rabbit, a dog, a bear and what appears to be an antelope.

The child receives four points for correct responses.

Item 4 is "Response to Pictures: Level 1." These are three pictures entitled "Grandmother's Story, Birthday Party, Wash Day." The child is presented with the pictures in that order and told, "Look at this picture and tell me all about it."

The "Grandmother" picture shows a lady sitting in a rocking chair, with her back to the kitchen stove, talking with three children who are seated on the floor near her, apparently listening to what she is saying. On the stove there is a pot boiling over, with the contents of the pot running down the stove onto the floor.

The "Birthday" picture shows three children arriving at a house. One child is shown carrying a package with a ribbon and bow on it. Another child is ringing the doorbell. In the window of the house is shown a cake with five glowing candles.

The third picture, "Wash Day," shows a clothesline hung with wash in the backyard. A dog, clutching a shirt in its mouth is racing away from the clothesline. The dog is being pursued by a woman.

At this age level, any of the following responses are acceptable: naming at least three objects in the picture; describing any one or more elements of the picture; interpreting the picture, even if inadequately or incorrectly, so long as it is not "purely fanciful or bizzare." Examples of "plus" answers for the "Grandmother" story are "lady, man, baby, chair," "lady sitting down," and "boy fall down."

Item 5 on the sub-test is "Sorting Buttons." There are ten white buttons, ten black buttons and two small boxes. The child is given two minutes to put all the white buttons in one box and all the black buttons in the other. No errors are allowed.

***861** The sixth item on the 3 1/2 year old test is "Comprehension 1." The item consists of two questions: "What must you do when you are thirsty?" and "Why do we have stoves?" Examples of "plus" responses are "Drink. Water. Go to sink. Have juice," and "That cook dinner. Cook on. For make our food. Warm the house. Heat things."

Next is the 4 year old sub-test. Item 1 is the same "Picture Vocabulary" we have seen before. The scoring is fourteen points.

Item 2 is "Naming Objects from Memory." A miniature automobile, dog and shoe are placed in a row before the child. His attention is called to each object and he is asked to name it. The examiner accepts whatever name the child gives. The child is told, "Now shut your eyes tight so that you can't see them." The examiner then covers one of the objects with a box. The child is told, "Open your eyes."

Look! Which one did I hide?" The child gives the name of the hidden object. The exercise is repeated for each of the three objects. The child must designate the hidden object either by its correct name or by the name he used when it was first shown to him.

Item 3 is "Opposite Analogies 1." The child is to complete each of the following sentences: (correct responses are in parenthesis)

- A. "Brother is a boy; sister is a ____." (girl, girly, little girl)
- B. "In daytime it is light; at night it is ____." (dark)
- C. "Father is a man; mother is a ____." (woman, lady, girl)
- D. "The snail is slow; the rabbit is ____." (swift, quick, rapid, speedy, fast, faster)
- E. "The sun shines during the day; the moon at ____." (night, nights, nighttime, during the night, in the night.)

This item is worth two points.

Item 4 on the 4-year test is "Pictorial Identification." This consists of a card with pictures of nine items a cat, an apple, a book, a clock, a bird, a cow, a rabbit, a stove, a fish, an umbrella, a moon, and a nest with eggs. The child is asked a series of questions: "Show me the one that . . .

- A. ". . . we cook on."
- B. ". . . we carry when it is raining."
- C. ". . . gives us milk."
- D. ". . . has the longest ears."
- E. ". . . shines in the sky at night."
- F. ". . . catches mice."

The child must point to the object on the card. He need not name it, nor is naming it sufficient. This item scores three points.

Item 5 is "Discrimination of Forms." This consists of a card with ten printed geometrical forms a square, a circle, a triangle, a rectangle and various irregular shapes. At the

bottom of the card is an "X." There is also an envelope containing cutouts of each of the ten geometrical figures. The examiner places the circle on the "X" and says, "Find me another one just like this," tracing his finger around the circumference of the circle. The same procedure is repeated with each of the other figures. Scoring is eight points.

The last item on the sub-test is "Comprehension II." The questions are: "Why do we have houses?" and "Why do we have books?" Acceptable answers are: "To go in. To go home. To cook in. To make us warm. To stay in for people . . ." and: "We read. We write'em. To color in. To go to school with. To see the pictures. To play with them . . ."

The next sub-test is for 4 1/2 year olds. Item 1 is "Aesthetic Comparison." This consists of three cards, each showing two human heads. The first card shows a pretty woman and an ugly woman. The same is true of the second card. The third card shows a handsome male and an ugly one. All of the persons depicted are white. The child scores three points for the correct answers. Dr. Williams questioned whether these comparisons really test intelligence and stated that in his opinion it is not a fair item for a child.

***862** What is "pretty" is so subjective a matter that I join Dr. Williams in questioning this item as a measure of "intelligence." Fortunately, the item appears only on this one sub-test. Although the sub-test is for 4 1/2 year olds, it would be used in the testing of school aged children whose "mental ages" are in this range.

The second item in the 4 1/2 year old sub-test is "Opposite Analogies," the same item which appears as Item 3 on the 4 year old test.

Item 3 is "Pictorial Similarities and Differences 1." The materials are six cards printed with pictures. All objects shown on a card are the same, except for one. The child is told, "Put your finger on the one that is not the same as the others." One card, used to demonstrate the test to the child, has four plus signs and a minus sign. The cards used for the actual test contain the following: three squares and a circle; three white balls and a black ball; three dogs and a cat; three chairs and a table; three saucers and a cup.

Item 4 is "Materials." The child is asked, "What is a house made of? What is a window made of? What is a book made of?" Examples of "plus" answers for the

house are wood, boards, cement, shingles, tile, lumber, blocks, and rocks. "Plus" answers for the window are glass, wood and glass, glass and steel. "Plus" answers for book are paper, cloth, leather, cardboard, plastic.

Item 5 is "Three Commissions." The examiner arranges the room in advance, making sure that a chair is available, that the door is open (or shut) and that a box is in plain view. The examiner then tells the child that he would like him to do something. He hands the child a pencil and says, "Here's a pencil. I want you to put it on the chair; then I want you to shut (open) the door; and then bring me the box which you see over there." The examiner points to each of the objects and repeats the instructions, making sure the child understands he is to perform the tasks in the sequence mentioned.

The child gets credit for the item if he performs the three tasks in the order mentioned.

Item 6 is "Comprehension III." This consists of two questions: "What do we do with our eyes?" and "What to we do with our ears?" Examples of "plus" answers are: "see . . . keep them open . . . read . . . wink'em . . . watch TV" and "hear . . . listen . . . to hear the radio . . ."

Next is the 5 year old sub-test. Item 1 is "Picture Completion: Man." The examiner shows the child a printed figure in the test booklet. It is an incomplete drawing of a man, consisting of a head, a torso, and one leg. The child is told it is a man if he does not recognize it, then told "See, he has only one leg. You finish him. Make all the rest of him." The child is given a pencil. If the child stops after making only the missing leg, the examiner says, "Make all the rest of him."

The scoring instructions caution that artistic qualities of the drawing are not taken into account. "The significant thing is the presence or absence of arms, legs, eyes, nose and mouth." The child receives a point for adding the other leg, a point for adding either both arms or one arm and a hand, and a point for attempting to fill in any additional facial features. (The face of the printed figure contains only the eyes.)

Item 2 on the five year old test is "Paper Folding: Triangle." The examiner folds a square sheet of paper once along the diagonal to make a triangle. He then folds this triangle once through the middle to make a smaller triangle. The examiner then gives the child another square sheet of paper and says, "Now you do it. Make one just like this." The examiner's triangle remains on the table.

The child is given credit if there are two folds and the resulting figure is approximately triangular in shape. The edges may be irregular.

Item 3 is "Definitions." The questions are: "What is a ball?," "What is a hat?," and "What is a stove?" Examples of "plus" *863 answers for ball are: "to play with . . . to roll . . . throw . . . catch the ball . . . round . . . kick . . .". A similar variety of definitions will suffice for the other two words.

Item 4 is "Copying a Square." The child is shown a printed square in the test booklet and told to "make one just like this. Make it right here." The scoring criteria are that the angles be preserved and that the lines not be broken.

Item 5 is "Pictorial Similarities and Differences II." This is a series of cards, each containing two objects. On some of the cards, the objects are identical, on some they are different. The child is asked to tell which objects are alike and which are different. There are two demonstration cards to familiarize the child with the test. One of these shows two identical trees. The examiner says, "See these two trees? They are just alike, aren't they? Just the same." The examiner shows the child the second card, which contains a circle and a square. He says, "But these two aren't alike, one is round and one is square." Then the examiner proceeds with the test cards, which show the following: Two pigs; a bug and a bird; a toy top and a football; two birds; a banana and an apple; two rabbits; two dolls; a chair and a table; two houses; a locomotive and a wagon.

Item 6 is "Patience: Rectangles." The materials are two small rectangular cards. One has been cut in half diagonally to make two triangles. The uncut card is placed on the table and the two halves of the divided rectangle are placed next to it. The two halves are separated and turned so that their relationship is not obvious. The child is then told, "One of my cards has been cut in two; you put these two pieces together to make a whole one just like this." The child is given three trials and there is no time limit.

The next sub-test is for 6-year olds. Item 1 is "Vocabulary." The examiner reads the words off a card and the child is asked what the word means. "Just tell me in your own words; say it any way you please." If the child can read, he is shown the vocabulary card as the examiner reads the words.

The test kit contains just one vocabulary card for all ages. The words on the card are listed in order of the difficulty experienced by subjects tested in 1950-54. The 45 words on the card are as follows:

- | | | |
|-------------|----------------------|------------------|
| 1. orange | 16. haste | 31. frustrate |
| 2. envelope | 17. peculiarity | 32. flaunt |
| 3. straw | 18. priceless | 33. incrustation |
| 4. puddle | 19. regard | 34. retroactive |
| 5. tap | 20. tolerate | 35. philanthropy |
| 6. gown | 21. disproportionate | 36. piscatorial |
| 7. roar | 22. lotus | 37. milksop |
| 8. eyelash | 23. shrewd | 38. harpy |
| 9. Mars | 24. mosaic | 39. depredation |
| 10. juggler | 25. stave | 40. perfunctory |
| 11. scorch | 26. bewail | 41. achromatic |
| 12. lecture | 27. ochre | 42. casuistry |
| 13. skill | 28. repose | 43. homunculus |

14. brunette

29. ambergris

44. sudorific

15. muzzle

30. limpet

45. parterre

A “plus” score for a 6 year old is six words defined correctly. They can be any six on the card, although the test publishers note that there is very little likelihood of success beyond the point where the child misses six consecutive words. Again, the examiner is told that “awkwardness of expression is disregarded.”

Item 2 on the 6 year old test is “Differences.” The child is asked, “What is the difference between . . .

- A. A bird and a dog.
- B. A slipper and a boot.
- C. Wood and glass.

Examples of “plus” responses for the bird and dog are: “A bird flies and a dog runs . . . a bird can’t bite . . . bird says ‘tweet’ . . . a bird flies . . . a bird got wings and a dog got ears . . . they’re different shaped . . . dogs have fur . . .”

Item 3 is “Mutilated Pictures.” The child is shown a card containing a number of items and is asked to tell, as to each item, “What part is gone?” or “What isn’t there?” The card shows the following: a wagon missing one wheel; a rabbit missing one ear; a pair of shoes with one shoelace missing; a teapot with no handle; a glove with a finger missing. The child must name or describe the missing part. Pointing is not sufficient.

Item 4 is “Number Concepts.” This test involves the twelve wooden cubes again. *864 The examiner places a piece of white paper on the table next to the blocks and asks the child to place specified numbers of blocks on the paper. The examiner starts by saying, “Give me three blocks. Put them here.” Then the test is repeated with ten, six, nine and seven blocks.

Item 5 is “Opposite Analogies II.” The child completes

the statement started by the examiner. The statements are:

- A. A table is made of wood; a window of _____.
- B. A bird flies; a fish _____.
- C. The point of a cane is blunt; the point of a knife is _____.
- D. An inch is short; a mile is _____.

Item 6 is “Maze Tracing.” The child is shown three diagrams which are identical except for the placement of the “schoolhouse” (indicated by a small drawing of a house) and “the little boy” (indicated by a figure). The schoolhouse is at the same point on each diagram, but the little boy is at a different place on the perimeter of the diagram. There is a “sidewalk” drawn around the perimeter of the diagram, and the idea is for the child to draw a path along the sidewalk which is the shortest way for the boy to go to school. “This little boy lives here, and here is the schoolhouse. The little boy wants to go to school the shortest way without getting off the sidewalk. Here is the sidewalk. Show me the shortest way. Mark it with your pencil, but don’t go off the sidewalk. Start here and take the little boy to school the shortest way.” The exercise is repeated with each of the three mazes. The child is given credit if he chooses the shortest way each time and the marking “is more inside than outside the boundaries of the path.”

The 7 year old sub-test starts with “Picture Absurdities I.” The child is shown five printed pictures and asked, “What’s funny about that picture?” The pictures are: a man walking in the rain, carrying an open umbrella in a manner that does not shield him from the rain; a man sawing a piece of wood, with the teeth of the saw pointing upward; a dog and a rabbit running at right angles to each other (instead of the dog chasing the rabbit); a man and a woman sitting outside their house in the rain, without

bothering to go inside; a cat lying still while mice play on and around him.

Item 2 on the seven year old test is "Similarities: Two Things." The child is asked to tell how each of the following are alike:

- A. Wood and coal.
- B. Apple and peach.
- C. Ship and automobile.
- D. Iron and silver.

Item 3 is "Copying a Diamond." The child is shown a printed diamond figure in the test booklet and is asked to make one just like it. The scoring criteria are that the angles should be preserved, that the pairs of angles be approximately opposite each other and that the figure be more diamond shaped than square.

Item 4 is "Comprehension IV." This consists of the following six questions:

- (a) "What should you do if you found on the streets of a city a three-year old baby that was lost from its parents?"
- (b) "What's the thing for you to do when you have broken something that belongs to someone else?"
- (c) "What's the thing for you to do when you are on your way to school and see that you are in danger of being late?"
- (d) "What makes a sailboat move?"
- (e) "What's the thing for you to do if another boy (or girl, depending on the sex of the subject; or another person, for adults) hits you without meaning to do it?"
- (f) "What should you say when you are in a strange city and someone asks you how to find a certain address?"

"Plus" answers for the question about finding the baby are: "Take it to a lost and found place . . . take her home and ask her where she lives . . . ask it its name . . . find the parents . . . bring him home until he is found . . . advertise in the paper . . . take it to the police station . . . try to take her home."

*865 Dr. Williams criticized this question because ". . .

finding a three-year old baby on the street is somewhat of an unlikely event. You know, it seems to me that questions which would plug into the child's lifestyle would be something which would be much closer to him." Dr. Williams thought, therefore, that the item is not a good measure of intelligence "if intelligence is measuring one's ability to cope with one's environment."

Item 5 is "Opposite Analogies III." The child completes the following statements read to him by the examiner:

- (a) "Rabbits' ears are long; the rats' ears are _____."
- (b) "Snow is white; coal is _____."
- (c) "The dog has hair; the bird has _____."
- (d) "Wolves are wild; dogs are _____."

Item 6 is "Repeating 5 Digits." The examiner reads a set of digits and the child repeats the set after him. The set must be repeated in correct order after only one reading. The three sets of digits are: 3-1-8-5-9, 4-8-3-7-2, and 9-6-1-8-3.

The eight year old sub-test starts with the vocabulary card. A passing score is any eight words.

Item 2 is "Memory for Stories. The Wet Fall." The examiner reads the child a story from a printed card. "Here is a story about 'The Wet Fall.' Listen carefully while I read it because I shall ask you questions about it." The child is allowed to read a copy of the story along with the examiner. Here is the story:

Once there was a little girl named Betty. She lived on a farm with her brother Dick. One day their father gave them a Shetland pony. They had lots of fun with it. One day, when Dick was riding on it, the pony became frightened and ran away. Poor Dick fell into a ditch. How Betty laughed when she saw him! He was covered with mud from head to foot."

The examiner then takes away the child's copy of the

story and asks the following questions:

- (a) "What is the name of this story?"
- (b) "What was Betty's brother's name?"
- (c) "Where did they live?"
- (d) "Who gave the pony to them?"
- (e) "What did the pony do?"
- (f) "What happened?"

A passing score is correct answers on five of the six questions. The scoring criteria allow some flexibility for the answers.

Item 3 is "Verbal Absurdities I." The examiner reads the child four statements and after each one asks, "Why is that foolish?" The statements are:

- (a) "The man had flu twice. The first time it killed him, but the second time he got well quickly."
- (b) "Walter has to write with his left hand because two years ago he lost both his arms in an accident."
- (c) "A man said, 'I know a road from my house to the city which is downhill all the way to the city and down hill all the way back home.' "
- (d) "An old gentleman complained that he could no longer walk around the park as he used to; he said they could now go only half way around the back again."

Item 4 is "Similarities and Differences." The child is told, "I am going to name two things and I want you to tell me how they are alike and how they are different." The things are:

- (a) Baseball and orange
- (b) Airplane and kite.
- (c) Ocean and river.
- (d) Penny and quarter.

Item 5 is "Comprehension IV." This is identical to Item 4 on the seven year old sub-test.

Item 6 is "Naming the Days of the Week." The child is asked to name the days of the week and then asked to tell

what day of the week comes before (a) Tuesday, (b) Thursday, and (c) Friday.

The nine year old test begins with an item called "Paper Cutting." The examiner has two 6 squares of paper and scissors. He says to the child, "Watch carefully what *866 I do. See, I fold the paper this way." He folds the paper in half, making a rectangle. "Now I will cut out a piece right here." The examiner makes a square cut at the center of the creased edge. The examiner then points to a printed square in the test booklet and asks the child to "make a drawing here to show how this paper would look if it were unfolded. Draw lines to show where the paper would be creased and show how and where it would be cut."

A similar but more difficult exercise is done with the second sheet of paper. In each instance, the child's drawing must show where the cut and folds would be if the examiner's sheet were unfolded.

The scoring instructions show various samples which are acceptable and unacceptable. The drawing may be rough, the important thing being the relative locations of the folds and cuts.

Item 2 is "Verbal Absurdities II." The question in each instance is, "Why is that foolish?" The statements are:

- (a) "Bill Jones' feet are so big that he has to pull his trousers on over his head."
- (b) "A man went one day to the post office and asked if there was a letter waiting for him. 'What is your name?' asked the postmaster. 'Why,' said the man, 'you will find my name on the envelope.' "
- (c) "The fireman hurried to the burning house, got his fire hose ready, and after smoking a cigar, put out the fire."
- (d) "In an old graveyard in Spain they have discovered a small skull which they believe to be that of Christopher Columbus when he was about 10 years old."
- (e) "One day we saw several icebergs that had been entirely melted by the warmth of the Gulf Stream."

Item 3 is "Memory for Designs I." This consists of a card printed with two designs. One is a series of straight lines joined together in a succession of right angles. The other is a box-like structure. The child is told that he is able to look at the drawings for ten seconds before the card will

be taken away and he will be asked to draw both designs from memory. The child is given full credit, half credit or no credit for each drawing, depending upon the completeness of detail.

Item 4 is "Rhymes: New Form." The examiner says to the child, "You know what a rhyme is, of course. A rhyme is a word that sounds like another word. Two words rhyme if they end in the same sound, like 'hat' and 'sat.' Now I want you to:

- (a) "Tell me the name of a color that rhymes with head."
- (b) "Tell me a number that rhymes with tree."
- (c) "Tell me the name of an animal that rhymes with fair."
- (d) "Tell me the name of a flower that rhymes with nose."

Item 5 is "Making Change." The examiner asks the child, "If I were to buy 4 cents worth of candy and should give the store keeper 10 cents, how much money would I get back?" The same question is asked for 12-15 and 4-25 cents.

Item 6 is "Repeating Four Digits Reversed." The examiner says a series of four numbers and the child repeats them backwards. The three sets of numbers are: 8-5-2-6, 4-9-3-7 and 3-6-2-9.

Next is the sub-test for 10-year olds. Item 1 is "Vocabulary." A passing score is eleven words.

Item 2 is "Block Counting." This consists of a printed card showing a number of pictures of stacks of blocks. The object is to count the number of blocks contained in each stack. There are three demonstration pictures for the child to practice on, with the help of the examiner, before he starts the test. Then the child proceeds to the fourteen pictures which constitute the test. They become progressively more difficult as only portions of blocks are shown or the presence of blocks cannot be seen but has to be inferred. A "plus" score is earned with eight out of fourteen correct.

Item 3 is "Abstract Words I." The child is asked to define (a) pity, (b) curiosity, (c) grief, and (d) surprise.

Item 4 is "Finding Reasons I." The child is asked two

questions:

*867 (a) "Give two reasons why children should not be too noisy in school.

(b) "Give two reasons why most people would rather have an automobile than a bicycle."

The reasons need not be good ones nor need they be greatly dissimilar to count as two. Examples of "plus" responses on Item (a) are: " 'cause they'll get a lickin . . . they'll have to sit in dunce chair . . . so they won't disturb them working, and won't disturb them reading . . . because the other pupils couldn't hear and if they couldn't hear what the teacher was saying then they couldn't do the problems and things"

Dr. Williams was critical of the question about noise in school. "First of all, there is not enough information in that particular question. I've had children tell me, 'Do you mean at recess?' What am I supposed to say? I tend to say, 'No, in school.' To me, that question is rather ambiguous."

As for the other item, why most people would rather have an automobile than a bicycle, Dr. Williams commented, "Well, that's absurd. There are people who do want I have no data on why people would prefer an automobile to a bicycle, especially in countries where a bicycle is the common modality. You know, there are arguments for bicycles just as there are arguments for cars. And some children are very, very bright and have creative responses on them. But there is only, again, a set of answers that has been predetermined."

The manual provides a number of examples of "plus" answers on the bicycle question, including: "because an automobile can go faster than a bicycle . . . because you fall down on a bicycle . . . because an auto can go faster . . . and a bicycle is too slow . . . you have to pump a lot"

Item 5 is "Word Naming." The child is told to say as many different words as he can in one minute. "Just any words will do, like 'clouds, dog, chair, happy'. When I say 'ready' you begin and just say the words as fast as you can and I will count them. Ready; go ahead."

A child receives a plus score for saying twenty-eight different words in one minute.

The final item on the sub-test is "Repeating Six Digits."

The three sets of six digits are 4-7-3-8-5-9, 5-2-9-7-4-6, and 7-2-8-3-9-4.

The 11 year old sub-test starts with “Memory for Design I” and the item is the same as Item 3 on the 9 year old test.

Item 2 is “Verbal Absurdities IV.” Again, the child is asked to say “what is foolish” about each of the following statements:

- (a) “The judge said to the prisoner, ‘you are to be hanged, and I hope it will be a warning to you.’ ”
- (b) “A well-known railroad had its last accident five years ago and since that time it has killed only one person in a collision.”
- (c) “When there is a collision, the last car of the train is usually damaged most. So they have decided that it will be best if the last car is always taken off before the train starts.”

Item 3 is “Abstract Words II.” The child is asked to give a definition for each of the following: (a) connection, (b) compare, (c) conquer, (d) obedience, (e) revenge. Examples of “plus” definitions of “connection” are: “connect two things together . . . like you add another part to a train . . . if somebody had a broken rope they’d tie the rope together, connected . . .”

Item 4 is “Memory for Sentences II.” The examiner reads two sentences which the child is asked to repeat. Before giving each sentence, the examiner says, “Now listen, and be sure to say exactly what I say.” The sentences are:

- (a) “At the summer camp the children get up early in the morning to go swimming.”
- (b) “Yesterday we went for a ride in our car along the road that crosses the bridge.”

A “plus” score permits no errors. Errors include omissions, additions or any changes in the order of the words.

***868** Item 5 is “Problem Situation II.” The examiner says, “Listen, and see if you can understand what I read.” The examiner then reads:

Donald went walking in the woods.

He saw a pretty little animal that he tried to take home for a pet. It got away from him, but when he got home, his family immediately burned all his clothes. Why?

Examples of “plus” answers are: “It was a skunk . . . it was a bad smelling animal . . . because it stunk . . .” Dr. Williams testified that this item is inappropriate as a measure of intelligence. “Now, how many kids would think of a skunk, especially in the urban area?”

Item 6 is “Similarities: Three Things.” The child is asked to state how the following things are alike:

- (a) snake, cow, sparrow
- (b) rose, potato, tree
- (c) wool, cotton, leather
- (d) knife-blade, penny, piece of wire
- (e) book, teacher, newspaper

“Plus” answers for Item (a) are: “all are animals . . . all move . . . they can all make a noise . . . they all stay out in the pasture . . . they have tails (eyes, tongue, skin, etc.) . . .”

Dr. Williams criticized the “snake, cow, sparrow” item on the basis that the word “sparrow” is uncommon.

“Now, it’s been a long time since I’ve heard the word ‘sparrow.’ It would seem to me that, first of all, that is a test that is a vocabulary test, that one has to know what those three phenomena are before one can really get the similarity. If I don’t know what a sparrow is, then I’m penalized from the beginning, and there are many kids who don’t know what a sparrow is. They know what a bird is, but not a sparrow.”

There may well be many younger children who do not

know what a sparrow is, but the item does not appear on the test until the 11 year old level. It would have been helpful to have some documentation of the fact that substantial numbers of 11-year olds are unfamiliar with the word. A high fail rate on Item (a) would be reflected in a sampling of test records.

The next sub-test is for 12 year olds. Item 1 is "Vocabulary" and a "plus" score is fifteen words.

Item 2 is "Verbal Absurdities II," and it is identical to Item 2 of the 11 year old test.

Item 3 is "Picture Absurdities II." It consists of one picture showing a man walking into the sunset and casting his shadow in front of him. The child is asked to state what is foolish about the picture. Examples of "plus" answers are: "the shadow is going the wrong way . . . the man should be between the sun and his shadow . . ."

Item 4 is "Repeating Five Digits Reversed." The numbers are 8-1-3-7-9, 6-9-5-8-2, and 9-2-5-1-8.

Item 5 is "Abstract Words I," and is identical to Item 3 on the 10 year old test.

Item 6 is "Minkus Completion I." The child is to insert the missing word in each of the following sentences printed in the record sheet. If the child is unable to read, the examiner may read the words for him. There is a five-minute time limit. The sentences are:

- (a) "We like to pop corn to roast chestnuts over the fire."
- (b) "One cannot always be a hero, one can always be a man."
- (c) "The streams are dry there has been little rain."
- (d) "Lincoln aroused no jealousy he was not selfish."

"Plus" answers for Item (a) are either "and" or "or." There is a variety of suggested "plus" answers for each of the other sentences.

The 13 year old sub-test starts with Item 1, "Plan of Search." The child is shown a diamond-shaped diagram in the test booklet. There is an opening at one of the corners. The examiner says, "Let's suppose that your purse with a lot of money in it has been lost in this big field. Take this pencil and start here (pointing) at the gate, and show me

where you would go to hunt for the purse so as to be sure not to miss *869 it." The child then marks the path he would take to conduct a search of the field. A "plus" answer is any series of lines showing a systematic plan that will cover the entire field. One example is concentric diamonds.

Item 2 is "Abstract Words II," which is the same as Item 3 on the 11 year old test.

Item 3 is "Memory for Sentences III." The child repeats these sentences after the examiner:

- (a) "The airplane made a careful landing in the space which had been prepared for it."
- (b) "Tom Brown's dog ran quickly down the road with a huge bone in his (sic) mouth."

Item 4 is "Problems of Fact." The examiner reads the child each of the following problems:

- (a) "A man who was walking in the woods near a city stopped suddenly, very much frightened, and then ran to the nearest policeman, saying that he had just seen hanging from the limb of a tree a a what?" (If the reply is "man," say, "Tell me what you mean; explain it.")
- (b) "My neighbor has been having queer visitors. First a doctor came to his house, then a lawyer, then a minister (preacher, priest, or rabbi). What do you think happened there?"
- (c) "An Indian who had come to town for the first time in his life saw a boy riding along the street. As the boy rode by, the Indian said, 'The white boy is lazy; he walks sitting down.' What was the boy riding on that caused the Indian to say 'he walks sitting down?' "

"Plus" answers for (a) include: "a dead man . . . a man who had been hung . . . a body of a person" "Plus" answers for (b) include: "somebody died . . . lawyer came to see about the will" "Plus" answers for (c) include: "bike . . . bicycle . . . kitty car"

Item 5 is "Dissected Sentences." The materials are three large cards on which disarranged words are printed. Before the examiner gives the child the first card, he says, "Here is a sentence that has the words all mixed up so that they don't make any sense. If the words were changed around in the right order they would make a good sentence. Look carefully, and tell me how the sentence

ought to read.” The child is given one minute for each card. The cards read as follows:

(a) “FOR THE STARTED AN WE COUNTRY
EARLY AT HOUR.”

(b) “TO ASKED PAPER MY TEACHER CORRECT I
MY.”

(c) “A DEFENDS DOG GOOD HIS BRAVELY
MASTER.”

For sentence (a) there are three “plus” versions and two versions which carry half-credit. For (b) there is only one “plus” response (I asked my teacher to correct my paper), but two other possibilities (My teacher I asked to correct my paper; To correct my paper I asked my teacher) rate half-credit. For (c) there are four “plus” possibilities and four half-credit possibilities.

Item 6 is “Copying a Bead Chain from Memory.” The materials are the same beads and shoestrings involved in Item 1 of the 3 year old test. The examiner makes a nine-bead chain, holding the string so the child can see it. “Watch carefully what I am making because I am going to take this one away and see if you can make one just like it.” The examiner strings beads on the chain in the following order: two round, one square, one round, one cylindrical, one round, one square, two round. When the chain is complete, the examiner lets the child look at it for five seconds and then removes it. The child has two minutes to make his own chain. No error is allowed.

The next sub-test is for 14-year olds.

Item 1 is “Vocabulary,” and seventeen words are required for a plus score.

Item 2 is “Induction.” The materials are six sheets of 8 1/2 X 11 paper. The examiner takes the first sheet, folds it in half and cuts out a small notch in the center of the folded edge. He unfolds the paper on the table and displays the one hole that has *870 been made in the paper. The examiner then takes the second sheet and folds it twice, the second fold being at right angles to the first. He cuts another notch from the folded edge and then unfolds the paper, displaying two holes.

The examiner then continues with sheets three, four and five, each time adding an additional fold and displaying to the child the resulting increase in the number of holes. Finally, when the examiner gets to the sixth sheet, he asks

the child how many holes there will be if he adds one more fold. He then says, “Give me a rule so that I can know each time how many holes there are going to be.” The child receives a plus score if he has grasped the rule by the time the sixth sheet is reached.

Item 3 is “Reasoning I.” The child is given the following problem printed on a card:

My house was burglarized last Saturday. I was at home all of the morning but out during the afternoon until 5 o'clock. My father left the house at 3 o'clock and my brother was there until 4. At what time did the burglary occur?”

A “plus” answer is any time between 4 and 5.

Dr. Williams was critical of this item:

Now, questions of this nature, again, get at reasoning to specific things, here, and is it a question of intelligence if the child does not come up with the correct answer? Why not you know, there are other kinds of reasoning problems, which, I think we could come up with which would be much more satisfactory. In other words, what does this tell? There is nowhere in the manual that would explain in theoretical terms how this is related to a theory of intelligence if the person can or cannot complete that particular item. And that's at the upper levels.

Item 4 on the 14 year old test is “Ingenuity I.” It consists of three problems which are read to the child. The problem may be repeated if necessary. No pencil or paper is allowed. The child is given three minutes to solve each problem. A “plus” score requires a correct solution of all three problems, so if the child misses the first one it is not

necessary to go on to the others. The problems are:

(a) "A mother sent her boy to the river to bring back exactly 2 pints of water. She gave him a 5-pint can and a 3-pint can. Show me how the boy can measure out exactly 2 pints of water using nothing but these two cans and not guessing at the amount. You should begin by filling the 5-pint can first. Remember, you have a 5-pint can and a 3-pint can and you must bring back exactly 2 pints of water."

(b) "This time he has to bring back exactly 13 pints of water. He has a 9-pint can and a 5-pint can. Show me how he can measure out exactly 13 pints of water using nothing but these two cans and not guessing at the amount. You should begin by filling the 9-pint can first. Remember, you have a 9-pint can and a 5-pint can and you must bring back exactly 13 pints."

(c) "And this time he has to bring back exactly 1 pint of water. He has a 3-pint can and an 8-pint can. Show me how he can measure out exactly 1 pint of water using nothing but these two cans and not guessing at the amount. You should begin by filling the 3-pint can first. Remember, you have a 3-pint can and an 8-pint can and you must bring back exactly 1 pint of water."

Item 5 is "Orientation: Direction I." The examiner reads the following problems to the child, emphasizing the critical words. It is permissible to repeat the questions if the child cannot remember.

(a) "Which direction would you have to face so that your left hand would be toward the east?"

(b) "Suppose you are going west, then turn to your right ; in what direction are you going now?"

(c) "Suppose you are going north, then turn to your left, then turn right ; in what direction are you going now?"

*871 (d) "Suppose you are going south, then turn left, then turn right, then turn left again; in what direction are you going now?"

(e) "Suppose you are going north, then turn left, then turn left again, then right, and then right again; in what direction are you going now?"

The final item (No. 6) on the 14 year old test is "Reconciliation of Opposites." The child is asked to tell in what way the following things are alike:

(a) Winter and Summer.

(b) Happy and sad.

(c) Loud and soft.

(d) Much and little.

(e) Beginning and end.

If the child says that these things are opposite, the examiner should repeat the question with emphasis on alike.

The final Stanford-Binet sub-test we will discuss is that for the "Average Adult." It is doubtful that a child suspected of mental retardation would advance this far in the test, but we will include a discussion of it just to be sure we have covered all items which could possibly be relevant.

Item 1 is "Vocabulary." A "plus" score is twenty words.

Item 2 is "Ingenuity I," which is identical to Item 4 on the 14 year old test.

Item 3 is "Differences Between Abstract Words." The examiner asks the subject to state the difference between

(a) "Laziness and idleness?"

(b) "Poverty and misery?"

(c) "Character and reputation?"

Item 4 is "Arithmetical Reasoning." The examiner shows the subject three cards, saying "Read this out loud and give me the answer." The examiner begins timing as soon as the subject has completed his reading of the card. No pencil or paper is allowed. The time limit is one minute for each card. The three problems are:

(a) "If a man's salary is \$20 a week and he spends \$14 a week, how long will it take him to save \$300?"

(b) If 2 pencils cost 5 cents, how many pencils can you buy for 50 cents?

(c) At 15 cents a yard, how much will 7 feet of cloth cost?

Item 5 is "Proverbs I." The examiner says, "Here is a proverb, and you are supposed to tell what it means. For

example, this proverb, 'Large oaks from little acorns grow,' means that great things may have small beginnings. What does this one mean?" The three proverbs are:

- (a) "We only know the worth of water when the well is dry."
- (b) "No wind can do him good who steers for no port."
- (c) "Don't judge a book by its cover."

Dr. Williams testified that he believes black children quickly understand proverbs from Africa but are rarely able to understand the proverbs in the Stanford-Binet test. He cited the proverbs, "The mouse that has but one hole is easily taken" and "You must not throw pearls before swine" as "ones they are rarely able to get." Because we did not have the test materials before us while Dr. Williams was testifying, we did not realize that these two proverbs are on the "Superior Adult II" sub-test of the Stanford-Binet that is, two sub-tests beyond the "Average Adult" sub-test where a proverb item first appears. Since these two proverbs have been standardized on adults of "superior" intelligence, it may well be that children, black or white, would have difficulty with them. But clearly the item is irrelevant to the question of whether young children in the Chicago school system are being unfairly tested, since the item would never be used on them.

The sixth item on the sub-test is "Orientation: Direction II." The examiner reads the following questions aloud, emphasizing the critical words:

- (a) "Which direction would you have to face so your right hand would be toward the north?"
- (b) "Suppose you are going east, then turn to your right ; in what direction are you going now?"
- (c) "Suppose you are going south, then turn to your left, then turn to your right ; in what direction are you going now?"
- *872** (d) "Suppose you are going north, then turn right, then turn right again, then turn left ; in what direction are you going now?"
- (e) "Suppose you are going west, then turn right, then turn right again, then turn right again, and then left ; in what direction are you going now?"

Item 7 on the sub-test is "Essential Differences." The

subject is asked to state the "principal difference between" the following things:

- (a) Work and play.
- (b) Ability and achievement.
- (c) Optimist and pessimist.

The 8th and final item on the "Average Adult" sub-test is "Abstract Words III." The subject is asked to state what is meant by the following words: generosity, independent, envy, authority and justice.

Three more sub-tests complete the Stanford-Binet Intelligence Scale. They are "Superior Adult I," "Superior Adult II," and "Superior Adult III." As the titles imply, they are more difficult, and increasingly so, than the "Average Adult" sub-test. It may come as some consolation to know that even on "Superior Adult III" the subject is not expected to know all 45 vocabulary words. A plus score is 30 words.

Because these superior adult sub-tests would not be used to test a school child suspected of mental retardation, it will not be necessary to describe them here.

The Evidence Concerning Bias

Now that we know what is on the tests, it will be useful to look again, this time more closely, at the testimony of the experts. The focus is on the question of whether the WISC-R, the WISC and the Stanford-Binet are culturally biased against black children so that it is unfair to use these tests in the determination of whether a black child is mentally retarded. We do not address the broader questions of whether these IQ tests are generally valid as measures of intelligence, whether individual items are appropriate for that purpose, or whether the tests could be improved. Those questions are not involved in this case. As plaintiffs make clear, "This case does not involve the general use of standardized IQ tests by Chicago defendants plaintiffs challenge only their utilization as part of the classification procedure for placing black children in EMH classes. . . ." (post-trial brief, p. 1)

Dr. Kamin, the psychologist from Princeton University, testified about the views of Goddard, Yerkes and Terman

that blacks and other ethnic groups are genetically inferior. This discredited theory, and the tenacity with which these pioneers of the IQ tests espoused it, undeniably prompts skepticism as to the validity of any work they did in the area of measuring intelligence. One would not be surprised by evidence that their theories of racial inferiority infected not simply their interpretation of the test results but the actual structuring of the tests. However, Dr. Kamin did not make that charge. Neither he nor any other witness attempted to demonstrate any bias in the test items traceable to the racist notions of Goddard, Yerkes, Terman and their followers. If evidence concerning their racial attitudes was offered only to show why they misinterpreted the test results, the evidence bears upon something which is not an issue in this case. Defendants agree with plaintiffs that there is no evidence to support a hypothesis that blacks have less innate mental capacities than whites.

Dr. Kamin's second major point was that the difference between white and black IQ scores is to be explained entirely on the basis that the test items call for "information" to which blacks, by reason of their different culture, have had less exposure than whites. He admitted, however, that he had not gone through the tests item by item to determine what these differences in exposure are. He testified that blacks have had different experiences than whites, which account for the differences in their performance on the test, but said that he does not know what these experiences are. Dr. George Albee also expressed the view that a poor performance by a black child on the IQ tests indicates his lack of exposure *873 to information peculiar to white culture. Like Dr. Kamin, however, he did not point out which items on the tests reflect white rather than black culture.

Dr. Williams described the characteristics of black culture which he believes are pertinent to this case. He stated that black culture is rooted in African philosophy, whereas "Anglo-Saxon culture" is rooted in European philosophy. He stated that these two cultures are "diametrically opposed." In European society, the emphasis is on the nuclear family, whereas in the African culture emphasis is on the extended family, which includes persons living outside the household. "Eurocentric" culture is individualistic, whereas in black culture there is emphasis upon the group. Eurocentric culture is highly competitive, with emphasis on the Protestant work ethic. "Afrocentric" culture emphasizes cooperation and collective responsibility, people trying to work together. Blacks also share a common language, different from the language used by whites. This is sometimes referred to as

non-standard English. Dr. Williams gave two examples of "standard" English which he says black children have difficulty understanding in test situations. One example was, "Point to the toy that is behind the sofa," which the black children missed because they had trouble with the words "behind," and "sofa." The question was changed to, "Point to the toy that's in back of the couch," and all the children understood it. Dr. Williams also gave the example of the black children who failed to understand, "Point to the squirrel that is beginning to climb the tree," but who did understand, "Point to the squirrel that is fixing to climb the tree."

Neither of these examples is taken from any of the IQ tests in issue in this case. It is necessary to follow Dr. Williams quite closely in order to know when he is talking about the WISC and Stanford-Binet and when he is changing the subject. It is also necessary to know what is on the tests. Someone not familiar with the specific test items would assume that Dr. Williams is giving examples of WISC and Stanford-Binet items when in fact he is not.

In the case of the WISC and WISC-R, the only language items Dr. Williams accused of racial bias were "rubies" (which might be mistaken for "Ruby") and "C.O.D." In the case of the Stanford-Binet, he did not identify any language items he considers racially biased.

Dr. Williams did not explain how he relates the other characteristics of black culture to performance on the tests. It is not clear, for instance, how the extended family as opposed to the nuclear family would pertain to performance on the tests. Like Dr. Kamin's description of the racist attitudes of Goddard, Yerkes and Terman, Dr. Williams' description of black culture has not been connected to the specific issue in this case.

As was noted during the description of the various test items, Dr. Williams addressed himself to very few of them. Whole sub-tests on the WISC and WISC-R, such as arithmetic, digit span, block design, object assembly, coding and mazes were not even mentioned. It is apparent that plaintiffs are not able to point out how these items are more typical of white culture than they are of black culture. The fact is, they are typical of neither. Many children, white and black, would never have worked a maze or a coding-type problem before confronting it on the IQ test.

Some test items, on the other hand, are quite similar to the material both black and white children are exposed to in the classroom before they would be asked to take an IQ

test. This is true of the arithmetic items. Generally, black and white children are exposed to such material in school to the same degree, and there is no evidence that arithmetic plays a bigger role in the nonschool environment of a white child than of a black child.

Dr. Kamin's argument that the black child does not obtain the same "information," and Dr. Albee's argument that the black child does not share in the dominant white culture, seem inapplicable to most items on all three of the tests in question. As already noted, many of the categories of *874 test items have no precise counterpart in the experience of any children, of whatever race. Others have almost precise counterparts in the everyday experience of American children of all races. Any number of test items could be cited to illustrate this point. However, it is impossible to draw valid conclusions about the fairness of the tests from examples of test items. The examples may not be representative of the test as a whole. Dr. Williams testified about certain test items which he believes are unfair. One who had not examined the tests in detail would be left with the impression that Dr. Williams had made a representative random sampling that typifies a cultural unfairness permeating the tests. In fact, the items were carefully selected and were not representative of the tests as a whole. They appear to be the only items Dr. Williams could find among the hundreds contained in the three tests which he thought would illustrate his charge of racial bias.

The evidence does not establish how the use of non-standard English would interfere with performance on the Wechsler and Stanford-Binet tests. In addition to the examples of the tree and the sofa, which are not on the tests, Dr. Williams testified that a black child might say, "John go to town" instead of "John is going to town," or "John book" instead of "John's book." These are not test items either, but that is not the principal point. What is unclear is how the use of such non-standard English would handicap a child either in understanding the test items or in responding to them. The fact that the child might say "John book" does not indicate that he would not understand the phrase "John's book." Moreover, responding to a test item in non-standard English should not affect a child's score on the item, since the examiners are specifically instructed by the test manuals to disregard the form of the answer so long as the substance is correct.

The vocabulary items on the tests should have drawn more fire from plaintiffs' witnesses if "non-standard English" presents a real problem. Yet, the only plaintiffs' witness who testified about any test items, Dr. Williams,

did not refer to a single one of the vocabulary items on any of the three tests. There are a total of 51 different vocabulary words on the two WISC tests and 45 on the Stanford-Binet. No effort was made to show that any of these 96 words are peculiar to white culture as opposed to black culture. Many of the words, such as "homunulus," "sudorific" and "parterre," are difficult, to be sure. But this is because they are not part of the common usage of anyone, black or white. The difficult words are at the upper reaches of the vocabulary scales, and even the brightest school children are not expected to know them. At the levels where a school child suspected of retardation would be tested, which the parties agree would be fairly early in the child's school experience, the vocabulary items are words of ordinary, common usage.

Dr. Williams testified that it is possible to devise a "black" vocabulary consisting of words which white persons would generally fail to understand. Dr. Williams has in fact prepared a test consisting of such items, which he calls the "BITCH" test (an acronym for Black Intelligence Test of Cultural Homogeneity), and which he says demonstrates the existence of a black language not shared by whites. Dr. Williams claims that the unfairness of asking whites to define such terms as "alley apple," "black draught," and "boogie jugie," three of the items on the "BITCH" test (Pl.Ex. 47, p. 124), is the same as expecting black children to deal with the vocabulary items on the standard IQ tests. However, the comparison is not apt. Dr. Williams concedes that his test includes words which are peculiar to black culture perhaps even to a sub-culture and would not be familiar to whites; indeed, that is his point. But there are no vocabulary items on the IQ tests, so far as I can tell, which are peculiar to white culture.

It would be possible to devise countless esoteric tests which would be failed by persons unfamiliar with particular subject matter. Every ethnic group, every business, trade or profession has its own vocabulary, its own universe of information *875 which is not generally shared by others. The fact that it would be possible to prepare an unfair test does not prove that the Wechsler and Stanford-Binet tests are unfair.

Dr. Williams' criticism of many test items appears unrelated to the question of racial bias. In fact, of the relatively few items he did discuss, most of them were criticized as inappropriate tests of any child's intelligence, not simply a black child's intelligence. Recall his criticism of the WISC and WISC-R questions about the height of the average American man, Genghis Kahn,

Charles Darwin, hieroglyphics, the location of Chile, and who discovered America and the Stanford-Binet questions about what to do if you discover a body, why one should not be noisy in school, why an automobile is better than a bicycle, the time of the burglary and why the boy's clothes smelled bad.

Dr. Williams did criticize some specific items on the ground that they were culturally biased against black children. I believe there is a substantial basis for some of these criticisms, and I have indicated in the preceding section of this opinion some of those items as to which I agree. On the WISC and WISC-R, I believe the following items are either racially biased or so subject to suspicion of bias that they should not be used:

1. "What is the color of rubies?"
2. "What does C.O.D. mean?"
3. "Why is it better to pay bills by check than by cash?"
4. "What would you do if you were sent to buy a loaf of bread and the grocer said that he did not have any more?"
5. "What does a stomach do?"
6. "Why is it generally better to give money to an organized charity than to a street beggar?"
7. "What are you supposed to do if you find someone's wallet or pocket book in a store?"
8. "What is the thing to do if a boy (girl) much smaller than yourself starts to fight with you?"

On the Stanford-Binet, I believe the one item which is racially inappropriate is the "aesthetic comparison" on the 4 1/2 year old sub-test, where the child is asked to tell which of two persons is "prettier."

It is important to know the extent to which these items are likely to figure in the EMH placement of a black child. Dr. Williams' view is that each item could be crucial. Testifying about the WISC item, "How tall is the average American man?," for which the child receives credit only for an answer between 5 ft. 7 in. and 5 ft. 11 in., Dr. Williams remarked that "... missing by one inch on a test might mean that one misses in life, educationally." Although I do not consider this particular item to be racially biased, Dr. Williams' point still deserves attention. The importance of missing a particular

item is, of course, magnified if indeed the difference between being sent to a class for the mentally retarded and not being sent there could rest on so slim a basis. However, there are factors which tend to protect against such an occurrence. First, as far as the WISC and WISC-R tests are concerned, the importance of an individual item is lessened by the fact that the child continues with the sub-test until he has a certain number of consecutive misses. Thus, the height item is No. 19 of 30 on the "General Information" sub-test of the WISC. Missing this item does not disqualify the child from continuing with the test unless he has missed the previous four items. No one is expected to answer all items correctly, and the fact that a child misses one item does not prevent him from accumulating points on others. Second, as far as this particular item is concerned, it occurs at a level of difficulty which a child in his early school years the time when EMH placement is generally made is not expected to reach anyway. This is clearly true of the identical item on the WISC-R, which appears as Item 24 in a sequence of 30.

This item about height does not appear on the Stanford-Binet, but the same concern for the possible disproportionate consequences of missing one question is appropriate. *876 It is even more appropriate in the case of the Stanford-Binet, because, unlike the Wechsler tests, it contains a number of multiple-part items for which no credit at all is given unless all parts are answered correctly. However, since I find that there is only one racially biased item on the Stanford-Binet, the problem seems negligible.

The third factor which mitigates the impact of missing any particular question is the fact that the IQ score is not the sole determinant of whether a child is placed in an EMH class. First, the score itself is evaluated by the psychologist who administers the test. The child's responses are recorded verbatim, and the significance of his numerical score is a matter involving judgment and interpretation. It is difficult to believe that a child who missed the height of the American male by one inch would be regarded by a trained psychologist as having thereby given evidence of mental retardation. Dr. Terrence Hines, who holds a masters degree in psychology from Howard University, testified that as a school psychologist in the Chicago system he finds that clinical judgment plays a large role in the interpretation of IQ test results. Regarding the item about the height of the average American male, for instance, he testified that while he scores the item in the manner required by the manual, he also makes a notation as to what the child's

response was and evaluates that response in light of the overall impression the child makes. He also gave examples of the clinical approach to other test items. If a child states that a stomach growls, he will ask that child what else it does. If a child says that he would go back home if the store was out of bread, Dr. Hines follows up with a question something like, "Then what?," to which the child might reply that his mother would then send him to another store.

In regard to the "fight" question what to do if a much smaller child hits you Dr. Hines pointed out the voice inflection of the examiner is important. The understanding of the child can be enhanced if the examiner emphasizes the fact that the other child is much smaller. Similarly, the item about what a child should do if he finds a wallet in a store can be made fairer for the black child if the examiner in asking the question puts the emphasis on the word should. In short, Dr. Hines testified that the examiner who knows the milieu of the child can correct for cultural bias by asking the questions in a sensitive and intelligent way.

Psychologists in the Chicago public school system must have at least a masters degree, with a minimum of 54 hours in psychology or educational psychology. Defendants also require applicants for employment to pass a written examination in school psychology as well as a practical examination involving the assessment of an individual child. It is relevant to note that 44 of the 193 school psychologists in the Chicago system are black and that defendants' witnesses testified without contradiction that the likelihood of a black child being placed in an EMH class without at least one black professional having participated in the evaluation is very slight.

Finally, the IQ test and the psychologist's evaluation of the child in the light of that test is only one component of several which forms the basis for an EMH referral. We will deal with this aspect of the case in a later section of this opinion.

I conclude that the possibility of the few biased items on these tests causing an EMH placement that would not otherwise occur is practically nonexistent.

The parties have each offered a number of articles from various psychological and educational journals. Plaintiffs also offered a book written by Dr. Kamin (Pl. Ex. 1) and one edited and partially written by Dr. Williams (Pl. Ex. 47). I have read these works and have derived very little help from them. Despite the prodigious volume of test

papers which has been accumulated over the past half century, there has been no extensive study undertaken to determine in specific terms just how blacks and whites compare to each other on all test items. Among the few researchers who have examined the subject at all, there is agreement that the mean black score is lower than the mean white score on all items, *877 across the board. It also seems to be agreed that the difference in mean score is greater on some sub-tests than others. I discern no agreement as to which sub-tests blacks find easier than others.

Plaintiffs argue that the racial bias of the IQ tests is shown circumstantially by the fact that blacks, although possessing the same innate mental ability as whites, do not score as well as whites on the tests. Plaintiffs say this of itself shows the tests must be measuring the amount of culture-specific information acquired by whites and not by blacks. There is no dispute in this case about the equality of innate intellectual capacity. Defendants assert no less strongly than plaintiffs that there are no genetic differences in mental capacity. However, the rest of plaintiffs' argument sidesteps their inability to point to any actual racial bias in the test items. All but a few of the items on their face appear racially neutral. It is not valid to draw an inference of unfairness if that can be done only by ignoring direct evidence of fairness. A preferable analysis would seek to account for all the data.

The IQ tests do not purport to measure innate intelligence. The test authors expressly disavow any such purpose. Both sides agree that no test has been or probably can be devised which will do that. What the tests appear to measure is the extent to which one utilizes his innate abilities in the performance of certain general categories of learned intellectual tasks. Performance on the tests reflects, for instance, the extent to which one has learned to observe, to see similarities and differences, to notice causal relationships, to remember, to draw inferences, to generalize, and to concentrate. The tests also indicate how well one has used those abilities to acquire certain specific knowledge, such as language and arithmetical concepts.

The acquisition and development of these mental skills is, according to defendants' witnesses Elmer Smith, a school psychologist, and Dr. Alice Zimmerman, Director of the Mentally Handicapped Development Program of the Chicago public school system, greatly affected by the child's early experiences. Early intellectual stimulation is essential. If the child does not receive it, or if he receives it to an insufficient degree, his intellectual development

his ability to use his innate capacity to deal with intellectual problems will be delayed. Lack of opportunity for cognitive stimulation is, according to these witnesses and some of the literature received in evidence, often due to factors associated with economic poverty in the home. Defendants offered a collection of census data (Def. Ex. 8(A)) documenting the fact that poverty and reduced socio-economic status is far more often a condition of black families than of white families in Chicago. According to Dr. Zimmerman, the areas from which most of the children in EMH classes in Chicago are drawn are the poverty pockets of the inner city. A high percentage of the population in these areas is black. Dr. Zimmerman explained how the Headstart program, a massively funded federal program designed to enhance the intellectual functioning of pre-school children in preparation for the school experience, operates on the premise that children in poverty areas are often deprived of adequate cognitive stimulation in the home environment.

Plaintiffs totally reject this suggestion that the performance of black children on IQ tests can be explained by deficiencies in their cognitive environment. Dr. Williams and Dr. Gloria Powell, the Director of Child Psychiatry in the Neuropsychiatric Institute at the UCLA School of Medicine, strongly contended that the mental stimulation received by poor black children in the inner city is adequate and in all respects equal to the intellectual stimulation received by middle class white children. They emphasized the richness of black culture and denied that any socio-economic disadvantages can explain the difference between black and white IQ scores. Dr. Powell's explanation for the disproportionate placement of black children in EMH classes is the "(f)ailure of the school to adequately assess the intellectual function of children who are culturally different." This failure, in the view of Drs. Powell and Williams, is caused by use of the culturally biased IQ tests.

***878** Dr. Powell's explanation does not satisfy me. She had no explanation for the presence of the Headstart program in the inner city and did not answer a question specifically directed to her about it. She seemed ill-prepared to discuss the question of cognitive stimulation. Most of her testimony was directed to whether health factors in the prenatal and postnatal environment, such as malnutrition and disease, can affect the intellectual functioning of the child. She strongly contends they do not. While defendants have submitted published material supporting the position that such physical health factors do play a significant role in mental development, I would be satisfied to rely upon the studies

cited by Dr. Powell. But as to whether factors associated with poverty in the home often inhibit an environment that will stimulate cognitive functioning in the child, I believe the defendants have presented the more persuasive argument.

Defendants' explanation of the IQ difference, that it is caused by socio-economic factors which interfere with the development of intellectual skills, is consistent with other circumstances not accounted for by plaintiffs' theory of cultural bias. It is uncontradicted that most of the children in the EMH classes do in fact come from the poverty pockets of the city. This tends to suggest that what is involved is not simply race but something associated with poverty. It is also significant that many black children who take the tests score at levels high enough to preclude EMH placement. Plaintiffs have not explained why the alleged cultural bias of the tests did not result in EMH level scores for these children. Plaintiffs' theory of cultural bias simply ignores the fact that black children perform differently from each other on the tests. It also fails to explain the fact that some black children perform better than most whites. Nationally, 15 to 20 per cent of the blacks who take the tests score above the white mean of 100. (Def. Ex. 36, p. 1).

I conclude that plaintiffs' have failed to prove their contention that the Wechsler and Stanford-Binet IQ tests are culturally unfair to black children, resulting in discriminatory placement of black children in classes for the educable mentally handicapped. Plaintiffs, however, claim that it is not their burden to show the tests are culturally biased against black children. Rather, they claim that defendants must prove the tests are culturally fair to black children. They base this argument on a provision of the Education of the Handicapped Act, 20 U.S.C. s 1412(5)(C), requiring, as a qualification for federal funding of education for the handicapped, that a state demonstrate it has established

" . . . procedures to assure that testing and evaluation materials and procedures utilized for the purposes of evaluation and placement of handicapped children will be selected and administered so as not to be racially or culturally discriminatory. Such materials or procedures shall be provided and administered in the child's native

language or mode of communication, unless it clearly is not feasible to do so, and no single procedure shall be the sole criterion for determining an appropriate educational program for a child.

Federal funds do provide a portion of the financing for defendants' EMH program, and the Department of Health, Education and Welfare has periodically reviewed defendants' assessment policies and procedures. The Department has not taken a position as to whether the standard IQ tests violate the statute or any regulation it has promulgated under the statute.

I do not read the statute as relieving plaintiffs of the burden of proof. The requirement that "materials and procedures" used for assessment be non-discriminatory, and that no single procedure be the sole criterion for assessment, seems to me to contemplate that the process as a whole be non-discriminatory. It does not require that any single procedure, standing alone, be affirmatively shown to be free of bias. The very requirement of multiple procedures implies recognition that one procedure, standing alone, could well result in bias and that a system of cross-checking is necessary.

***879** *The Assessment Process*

Defendants' system for the identification and placement of mentally handicapped children, which is spelled out in manuals and printed regulations, involves several levels of investigation. It is important to understand that an IQ test is not the first level, nor is an IQ score the catalyst for the assessment process. The first level of investigation is the classroom. Unless the child is having difficulty with his studies in the classroom, the question of EMH placement will never arise and there is no occasion for an IQ test. Individually administered IQ tests of the kind involved in this case have never been given routinely in the Chicago school system, and the former practice of giving group-administered general intelligence tests to all students was discontinued some years ago.

If the classroom teacher has reason to believe the child has an educational handicap, the matter is taken up with

the school principal. The teacher prepares a written report concerning the child, using a prescribed form. The principal then convenes a screening conference. The participants are the principal, the classroom teacher, a parent of the child and any other appropriate persons.

The screening committee makes a recommendation to the principal as to whether a case study should be requested for the child. If the principal determines on the basis of this recommendation that a case study is warranted, the matter is referred to the Special Education Bureau for the development of an appropriate case study program.

Various professional personnel then determine what areas of evaluation are appropriate for the child. On the basis of this determination, the child is examined by persons in the appropriate disciplines. This may involve a medical examination, a psychiatric examination, a psychological assessment or reference to a social worker or a speech therapist.

When the case study evaluation has been completed, the principal convenes a multidisciplinary staff conference. The members of this staff include a representative of the special education program, all of the professionals who evaluated the child, the school principal, and the parents of the child. The purpose of this multidisciplinary staff meeting is to determine whether the child should be placed in a special education program, and if so, what program it should be. The report of the staff conference is in writing. Each participant must sign it and indicate whether he or she concurs in the recommendation.

No child can be placed in an EMH class unless the placement is recommended by a psychologist who has evaluated the child. While the conference can decline an EMH placement recommended by the psychologist, it cannot make such a placement without the psychologist's recommendation.

If either the child or his parents on the one hand or the school officials on the other are dissatisfied with the decision, they may request further hearings on the matter. Placement of the child is stayed pending the hearing, which is conducted by an impartial hearing officer assigned from another school district by the State Department of Education. Ultimately, the matter is subject to administrative review in the courts of Illinois.

The evaluation and placement process is not carried out hastily. There are more children in need of placement than there are available seats in the EMH classrooms.

Sometimes the decision is against placement even though the parent desires it. A motive for unnecessary placement is nonexistent, since the cost to the local system of administering the program far exceeds the state and federal aid received for it. The total cost of the Chicago special education program exceeds by 50 million dollars per year the state and federal funds received to support it.

Plaintiffs claim that, despite the various steps involved in EMH placement, the placement decision is really made primarily on the basis of the child's IQ score. They argue that the IQ score has a "hypnotic effect" on the participants in the multidisciplinary staffing, so that a child with an IQ ⁸⁸⁰ of less than 80 stands a high chance of being put in an EMH class on that basis alone. Several of plaintiffs' witnesses so testified. Plaintiffs complain that social workers are not used often enough and that there is insufficient investigation of the child's family situation and his adaptive behavior outside the school environment.

To prove their contention that IQ scores are the dominant factor in EMH placement, plaintiffs called Richard Berk, a professor of sociology at the University of California. Dr. Berk specializes in statistical analysis of sociological data. He did a computer study of the records of 6,000 students examined by school psychologists in the Chicago system during the period 1973 to 1976. This was a random 10 per cent sample of all students examined by the school psychologists during that period of time.

Dr. Berk concluded that the most constant characteristic of all children who were placed in EMH classes was the fact that they had low IQ scores. He concluded from this that there is a strong relationship between low IQ and placement in an EMH class.

I fail to see how this testimony proves plaintiffs' contention that IQ scores are given undue weight in the placement decision. One thing I would expect of a mentally retarded child is that he would have difficulty with an IQ test. If children with high IQ's were being referred to classes for the mentally retarded, there would be great cause for concern. The fact that not all children referred to EMH classes have other characteristics, such as dental problems or various other traits considered by Dr. Berk in his computer study, is not surprising. While there may be reasons other than mental retardation that would account for a low IQ score, it is difficult to see how a high IQ could be reconciled with a finding of mental retardation.

One exhibit prepared by Dr. Berk (Pl. Ex. 75) shows the percentages of the sample who were recommended for EMH placement and actually placed in EMH classes at each IQ level. This exhibit does not support plaintiffs' contention that a low IQ score unduly influences either the recommendation or the placement decision. In the IQ range of 55-59, the lowest range shown on the exhibit, 81 per cent of the children were recommended for placement, but only 54 per cent were placed. (Apparently the recommendation referred to is that of the psychologist, although this is not clear from the exhibit.) This means that 46 per cent of the children whose IQ's were between 55 and 59 were not placed in EMH classes even though they were considered for EMH placement. In the IQ range of 60-64, 88 per cent were recommended and 62 per cent were placed. For IQ's of 65-69, 88 per cent were recommended and 56 per cent were placed. In the 70-74 range, 78 per cent were recommended and 48 per cent were placed. In the 75-79 range, 46 per cent were recommended and 32 per cent placed in EMH classes.

In the circumstances of this case, where defendants have shown that IQ scores are only one factor which enters into the EMH assessment and that a low IQ score frequently does not result in such placement, I believe the burden of showing an absence of racial bias in the tests does not rest on the defendants.

It is apparently plaintiffs' position that the tests in question could be vindicated only by a showing that blacks do as well on them as whites. Plaintiffs regard any differential performance between the races as evidence of cultural bias. Plaintiffs' witnesses emphasized what they felt was the unfairness of the standardization process which discarded items on which females did not do as well as males but did not discard items on which blacks did less well than whites. These witnesses did not explain the nature of the items on which there was a sex difference. One of the published articles in evidence (Pl. Ex. 29, p. 181) refers to the study by McNemar in 1942 which identified the items as to which boys and girls scored differently on the Stanford-Binet:

The differences between the sexes were peculiar to just a small number of items, about thirty altogether (McNemar, 1942: 50): about half the differences went in ⁸⁸¹ one direction and about half in the other. The great bulk of items showed no differences.⁴

(Emphasis in original). In the case of the racial differences, we are not talking about just thirty sub-items or even items. We are talking about all items on the tests. If all items on which the mean black score is lower than the mean white score were to be eliminated, this would mean that the entire tests would be eliminated.

Plaintiffs believe that their theory of the case accounts for the fact that the mean black score on each and every item is lower than the mean white score. They contend that the difference is entirely due to cultural bias. The implications of the argument are striking. Plaintiffs' hypothesis implies, for instance, that of the 328 items on the WISC-R, spread across twelve sub-tests of different kinds of subject matter and standardized on a sample that included representative numbers of blacks, there is not a single item which is not culturally biased against blacks. That such a thing could happen by chance, or because of simple inadvertence on the part of the psychologists who devised the test, is difficult to believe. No statistical evidence was presented on the question, but it seems highly unlikely that if mere inadvertence were involved, at least a few culturally fair items would not have found their way onto the test simply by chance. The strain on the laws of probability does not, of course, end with the WISC. The plaintiffs' hypothesis also includes the Stanford-Binet. This is a separate test, devised at a different time by different people according to a different format than the WISC tests. The Stanford-Binet contains 104 items spread over the seventeen sub-tests ranging from age 2 years through average adult. Each one of them according to plaintiffs' theory, is culturally biased against blacks.

It is unfortunately true that, despite what I believe are sincere efforts on the part of the defendants to avoid erroneous placements, some children are placed in EMH classes who should not be there. Small but significant numbers of EMH students are constantly being moved out of EMH classes back into the regular curriculum. Some of these transfers are due to the fact that the child has progressed in the EMH class and is ready for a greater challenge. In other instances, however, the child is transferred because it is belatedly discovered that he should not have been there in the first place.

These erroneous placements have not been shown to be due to racial bias in the IQ tests. The situations of the two named plaintiffs illustrate this failure of proof. These two black children, Barbara B. and Angela J., were each evaluated as being mentally retarded and were transferred out of their regular classes to EMH classes. Each child

was evaluated by a school psychologist and achieved a low score on one of the WISC tests or the Stanford-Binet.

Later reevaluation of these children disclosed that they are not mentally retarded. They have normal intelligence but suffer from learning disabilities which make it difficult for them to perform well in certain kinds of learning situations. One of plaintiffs' witnesses was Robert E. Stoner, a clinical psychologist who examined and evaluated each of the named plaintiffs in connection with this case. He gave each of them a variety of tests, including the WISC. He found that each of these children had perceptual problems which interfered with their visual and auditory discrimination between different shapes and sounds. This accounted for their difficulty with such items as picture completion and picture arrangement. Incorrect answers to some of the verbal items may have been caused by the fact that the child did not correctly perceive the words spoken by the examiner.

Mr. Stoner testified that Barbara and Angela should have been placed in special classes for the learning disabled rather than *882 classes for the mentally retarded. He pointed out that an educational handicap should have been suspected by reason of the profiles these two children show on their IQ tests. If there is a significant difference in the level of performance on the various sub-tests, this suggests a learning disability rather than mental retardation. Such differences did appear on the WISC tests Mr. Stoner administered to Barbara and Angela, and his interpretation of this data was a primary basis for his conclusion that these children are learning disabled.

While Mr. Stoner testified that he believes the standard IQ tests are culturally biased against black children, he did not indicate any particular items he believes are biased. The most remarkable thing about Mr. Stoner's testimony in the context of the present discussion is that he did not ascribe the misassessment of Barbara and Angela to any racial bias in the tests. He did not say that they missed any items because of cultural bias. The problems these two children had with the tests were caused by their learning disabilities, not by any bias in the test items. On the sub-tests where their perceptual problems did not inhibit performance, these children scored within the normal range. Had their test results been properly interpreted in the first instance they would not have been assigned to EMH classes.

Plaintiffs seem not to realize that their own evidence shows the two class representatives, Barbara and Angela, do not have claims which are typical of the class they

purport to represent.⁵

on the tests. It is *883 clear that this was not undertaken in the Larry P. case.

The Larry P. Case

This is not a case of first impression. The exact issue of racial bias in the WISC, WISC-R and Stanford-Binet tests has been decided by Judge Robert F. Peckham of the United States District Court for the Northern District of California in the case of Larry P., by his Guardian ad Litem, Lucille P., et al. v. Wilson Riles, Superintendent of Public Instruction for the State of California, et al., 495 F.Supp. 926 (1979). Plaintiffs rely upon that decision heavily, since Judge Peckham held that the tests are culturally biased against black children. Judge Peckham heard a number of the same witnesses who testified here, including Professors Kamin, Albee and Williams and Dr. Gloria Powell. He found their testimony persuasive. Judge Peckham's lengthy and scholarly opinion is largely devoted to the question of what legal consequences flow from a finding of racial bias in the tests. There is relatively little analysis of the threshold question of whether test bias in fact exists, and Judge Peckham even remarked that the cultural bias of the tests ". . . is hardly disputed in this litigation. . . ." (p. 959; see also n.69). I find reference to specific test items on only one page (p. 958) of the opinion. Judge Peckham mentions the WISC "fight" item, finds that it is culturally biased against blacks and then remarks, "Similarly, it may be that such questions as who wrote Romeo and Juliet, who discovered America, and who invented the lightbulb, are culturally biased." Finally, Judge Peckham noted that ". . . such skills as 'picture arrangement' may be tested in a biased fashion if the pictures, which generally are of caucasian persons, relate to situations more typical of white, middle class, life than the experiences of many black children." (p. 958).

As is by now obvious, the witnesses and the arguments which persuaded Judge Peckham have not persuaded me. Moreover, I believe the issue in the case cannot properly be analyzed without a detailed examination of the items

Conclusion

I have found one item on the Stanford-Binet and a total of eight items on the WISC and WISC-R to be culturally biased against black children, or at least sufficiently suspect that their use is in my view inappropriate. These few items do not render the tests unfair and would not significantly affect the score of an individual taking the test. The evidence fails to show that any additional test items are racially or culturally unfair or suspect.

I believe and today hold that the WISC, WISC-R and Stanford-Binet tests, when used in conjunction with the statutorily mandated ("other criteria) for determining an appropriate educational program for a child" (20 U.S.C. s 1412(2)(D)(5), do not discriminate against black children in the Chicago public schools. Defendants are complying with that statutory mandate.

Intelligent administration of the IQ tests by qualified psychologists, followed by the evaluation procedures defendants use, should rarely result in the misassessment of a child of normal intelligence as one who is mentally retarded. There is no evidence in this record that such misassessments as do occur are the result of racial bias in test items or in any other aspect of the assessment process currently in use in the Chicago public school system.

I find the issues in favor of the defendants and against the plaintiffs. The Clerk is directed to enter judgment for the defendants.

All Citations

506 F.Supp. 831

Footnotes

¹ The statutes upon which plaintiffs rely are s 504 of the Rehabilitation Act of 1973, 29 U.S.C. s 794, which prohibits discrimination against handicapped persons in certain programs receiving federal funding; Title VI of the Civil Rights

Act of 1964, 42 U.S.C. s 2000d et seq., prohibiting racial discrimination in the administration of federally funded educational programs; The Equal Educational Opportunities Act of 1974, 20 U.S.C. s 1703, which prohibits the denial of equal educational opportunity on the basis of race; and The Education of the Handicapped Act, 20 U.S.C. s 1412(5)(C), requiring as a condition of federal funding that states not use testing and evaluation procedures which are racially or culturally discriminatory.

- ² I will also save discussion of the legal issues for that section of the opinion. As will appear, I believe this case is primarily a factual controversy and that most of the legal questions raised by the parties need not be reached. Plaintiffs assert that if the IQ tests are racially discriminatory they are entitled to relief under the Constitution and all of the federal statutes mentioned in n. 1, supra, as well as certain regulations issued by the United States Department of Health, Education and Welfare. Defendants, while denying that the tests are discriminatory, also contend that plaintiffs are not entitled to relief under the Constitution unless discriminatory intent is shown; they further argue that none of the federal statutes relied upon by plaintiffs creates a private right of action.

If there is racial bias in the tests and the plaintiffs are being placed in classes for the mentally retarded because of that bias, I believe plaintiffs would have a right to relief on at least some of their theories. However, in view of the way I decide the facts, it will not be necessary to discuss these legal issues.

- ³ Rule 702 of the Federal Rules of Evidence provides that expert witnesses may testify to their opinions when their specialized knowledge "will assist the trier of fact to understand the evidence or to determine a fact in issue. . . ." The rule does not mean that the trier of fact must rely upon expert testimony which is unsatisfactory or that the trier of fact is precluded from making an independent determination of the facts, regardless of how complicated or "specialized" the subject matter might be. "Expert testimony . . . is not conclusive upon the trier of fact . . . even though unimpeached and uncontradicted, since the trier may apply his own experience or knowledge in determining how far to follow the expressed opinion. . . ." *Security First National Bank of L. A. v. Lutz*, 322 F.2d 348, 355 (9th Cir. 1963). See, also *Mopkins v. St. Louis Die Casting Corp.*, 569 F.2d 454, 455 (8th Cir. 1978), where the court, commenting on the testimony of a psychologist to the effect that racial discrimination was involved in certain ethnic expressions, noted that ". . . (o)pinions testified to by an expert are certainly not binding on the court." See also *Pittman v. Gilmore*, 556 F.2d 1259, 1261 (5th Cir. 1977) (expert medical testimony); *Magno v. Corros*, 439 F.Supp. 592, 603 (D.C.S.C.1977): "However, it is well established that if men of ordinary knowledge are as capable of comprehending primary facts, and of drawing correct conclusions therefrom, as are 'experts,' such expert testimony may be properly disregarded." I have not disregarded the expert testimony in this case, but neither do I feel bound or limited by it. The factual determinations to be made are well within the capability of any competent trier of fact.

- ⁴ The reference to "items" in the above quotation probably means sub-items since there are only 104 "items" on the test.

- ⁵ This fact has no independent impact upon the rights of the class, since I have found that the class has itself sustained no injury of the kind alleged in the complaint. The evidence not only fails to show that the named plaintiffs were injured by racial bias in the tests, it fails to show that there is racial bias in the tests sufficient to cause injury or pose a threat of injury to any member of the class.

The individual plaintiffs have not sought monetary relief for their misplacement in the EMH program. Therefore,

there is no occasion to discuss the legal issues which would be raised by such a claim.
