Problems of Adapting Intelligence Scales from One Culture to Another

PABLO ROCA

Director, Technical Division, Department of Education, Commonwealth of Puerto Rico, Hato Rey, P. R.

This article deals with the problems encountered by psychologists of the Division of Research and Statistics of the Department of Education of Puerto Rico in the translating and adapting for use in the schools of Puerto Rico of the following intelligence tests: Wechsler Intelligence Scale for Children, Stanford-Binet Scale, Revised, Form L, and the Goodenough Intelligence Test. These adaptations were necessary because of the lack of measuring instruments for Puerto Rican children in particular and for Spanish-speaking children in general.

The Wechsler Scale

The first instrument with which we worked was the Wechsler Intelligence Scale for Children. This scale consists of two parts: one verbal and one performance. The verbal scale comprises five subtests: general information, general comprehension, arithmetic, similarities, and vocabulary. The performance scale consists of the following subtests: digit span, picture completion, picture arrangement, block design, object assembly, coding, and mazes.

Three experiments were conducted in the process of adaptation. The first included three children in each grade: one average, one above average, and one below average. The range of IQ’s found was from 70 to 114 with a median of 84.5 in the verbal scale and a range of 47 to 120 in the performance scale with a median of 85.5. The total scale ranged from 59 to 106 with a median of 80.75.

Correlated with the Goodenough test we obtained a correlation of .672 ± .09. Correlated with teachers’ judgment a contingency correlation of .32 was obtained.

As a result of this first experiment, the following changes were effected in the scale: General information, question number 16, Who wrote Romeo and Juliet? to Who wrote Don Quixote? Question number 19, How tall is the average American man? to How tall is the average Puerto Rican man? It happens that whereas the average height of the American man is 67 inches, the Puerto Rican is...
65. Question number 24, How far is it from New York to Chicago? became How far is it from New York to Puerto Rico?

In the arithmetic test we changed question number 16, which reads: Smith and Brown start a card game ... to John and Peter start a marbles game ...

In the Analogies test we changed question number 5, In what way are a peach and a plum alike, to ... an orange and a banana.

In the vocabulary test we discovered that certain words such as spade, mantis, spangle, and belfry were either too easy when translated into Spanish, or else had different meanings. We had to study the frequency lists in both English and Spanish in order to select a Spanish word on the same difficulty level. One problem here was the lack of Spanish frequency counts at the time.²

In the performance scale no changes were effected in this first experiment.

In the second experiment we selected middle-class socio-economic status children from three schools. The number of students was 69, three from each grade: one average, one above average, and one below average. This time all questions in each subtest in the verbal scale were given, because we wanted to find the difficulty level of each question. In order to avoid boredom one verbal subtest was followed by one performance subtest.

As a result of this second experiment the following changes were effected:

General information. Questions were arranged in a new order of difficulty. It is interesting to discover which questions became either easier or more difficult when translated into Spanish. For example, question number 2, What do you call this finger? (Thumb) became question number 11. The reason is that Pulgar which is the Spanish word for thumb is much more difficult in Spanish than in English. Of course, the common name of thumb is dedo grande which simply means big finger.

On the other hand question number 7 now became question number 4; that is, it is easier for Spanish-speaking children to know how many pennies make a nickel. All other questions changed positions, but not so much as did numbers 2 and 7. The only questions keeping their places were numbers 1, 12, and 30. Question one is How many ears have you?, question 12 is What is the color of rubies, and question 30 is What is a lien?

² At a later date the University of Puerto Rico published a ten million word count of the Spanish language: Consejo Superior de Enseñanza, Recuento de Vocabulario Español, Vol. I, Editado por OEA y UNESCO. Rio Piedras: Universidad de Puerto Rico 1952.
Comprehension. Questions 1 and 2 remained in their original position. Question number 6 became number 3 in our adaptation. This question refers to brick houses, which are non-existent in Puerto Rico. Rather, we utilized concrete houses, which are common here. Question number 8 contained the word *shipwreck* in English. When translated into Spanish, *naufragio* became so difficult that children could not answer the question. We had to describe the word by substituting the phrase *when a ship sinks* to enable the children to understand the problem involved.

Question 11 asks *Why should most government positions be filled through examinations?* Pupils related examinations to medical examinations and it was not until we added *civil service examinations* that they understood the idea.

Arithmetic. Question 12 in the original had to do with buying oranges. This question was easier for our children, who are accustomed to buying oranges, and therefore, it became question number 10 in our adaptation.

Similarities. Only changes in position were effected in this subtest. *Liberty and Justice* became the most difficult question rather than 49 and 121 as in the original.

Vocabulary. The lowest rho correlation between ranks in the original and ranks in our adaptation was obtained for this subtest. Some words like spade and spangle became easier when translated into Spanish. Even though the word *fur* had the same count in Eaton’s list (2a) as its Spanish equivalent *piel* (2a), we found the Spanish word more difficult for our children. We must keep in mind that we have summer time all throughout the year. The same thing happened with the word *diamond* although in this case the English count was 2a versus a Spanish count of 4a in Eaton’s list.

We could not find a word in Spanish that would translate the word *mantis* and we had to substitute for it *beetle*, which of course resulted in an easier word in Spanish than *mantis* in English. We could not translate adequately *belfry* which means *campanario* in Spanish, a word too easy in a predominantly Catholic country. We chose *espadaña* which became a more difficult word than *belfry*.

Digit Span—No alterations were made in this subtest.

Picture completion. No great changes were effected in this subtest, except changes in rank order of difficulty. Some noticeable changes in this respect were as follows: The fish with the dorsal fin missing became easier in Puerto Rico where there is abundant fish-
ing all year around. The rooster with the spur missing also became easier. I should explain that cock fighting is legal in our country. On the other hand, the card with the center spade missing became more difficult. Card games are not so popular in Puerto Rico.

Picture arrangement. No alterations were made in this subtest, but the percentage passing dropped considerably from number 3 to number 4 and from number 4 to number 5. Question number 3 is titled farmer, question number 4 picnic, and question number 5 sleeper. Both 4 and 5 are rare events in our children’s lives.

Block Design. No changes were effected here. However, the number passing was lower than in the picture arrangement test.

Object assembly. In this subtest the automobile was found to be easier for our children than the face.

Coding and mazes. No changes made.

All questions were tested for validity in accordance with Kelley’s upper and lower groups. Only 12 out of 30 questions in the general information test were found with satisfactory phi coefficients. In comprehension 4 out of 14, in arithmetic 4 out of 16 and in analogies 4 out of 16 had satisfactory phi coefficients. In vocabulary, however, 37 out of 40 were found with satisfactory phi coefficients. The only subtest where all questions were found valid was the object composition test. The block design test had 8 out of 10 questions valid.

A third experiment was conducted with 40 subjects, 20 boys and 20 girls, selected in the same manner from the elementary grades in three schools. In analyzing the test questions, all subjects of the three experiments were combined, thus making a total of 128 subjects. The distribution of IQ’s found ranged from 46 to 138, with a median of 87.94. The mean was 88.01 and the standard deviation 21.60.

The conclusion reached indicated that the average Puerto Rican child scores 12 IQ points less than the average American child; in other words, IQ’s of 77 to 99 as given by the Wechsler Intelligence Scale for Children Adaptation for Puerto Rico seem to us normal.

A second group of children who were tested independently of the one used in the three experiments described above confirmed the average IQ found. This confirmation showed us that the adaptation was consistent in its results. Why our children scored lower than the American children could be explained in terms of cultural differences.
The Binet Scale

Three translations of this scale had been made in Puerto Rico and one in Spain before we started our experiments. None of these had developed norms for the scale nor were they adaptations based on experimental evidence. They were simply translations.

The first problem we met in our attempt to translate and adapt the scale occurred with Test XI, No. 5, Problem Situation, which referred to the skunk, an unknown animal in Puerto Rico. Naturally, we had to replace the skunk problem with something else. Four problems were selected and tested; one had to do with a wasp, one with a crab, and two were taken from Form M of the same scale, Test VIII, No. 5 and Test XIV, No. 1.

A second problem had to do with the codes in Average Adult, No. 2. This was not an easy problem, inasmuch as the Spanish alphabet has two double letters following the single letter, which each constitute a separate, independent and different letter. This is the case of the double l (ll) and the double r (rr).

Three experiments were conducted with the Binet scale, following the same procedure for the selection of students that was described for the Wechsler adaptation. In the first, 34 children were used; in the second 138, and in the third over 700.

The first experiment yielded IQ's ranging from 50 to 126 with a median of 79.33. As a result of this first experiment the third of the four Problem Situation tests was selected to replace the skunk problem. The selected problem corresponds to Form M, VIII, 1. As to the codes, the second one was selected.

Some other changes were effected as result of this first experiment. In VII, 2, Similarities: Two Things, the order was changed. In XI, 3, Abstract Words I, conexión which is the Spanish word for connection was changed to conectar, which is the verb form, because children understood the word to be colección, which means collection.

In XII, 5, Abstract Words II, courage (valor in Spanish) had to be changed to valentía (gallantry) because valor also means value and children responded to this easier meaning.

In Average Adult, No. 3, Differences Between Abstract Words, Poverty and misery (Spanish pobreza y miseria) had to be changed to pobreza y martirio (poverty and martyrdom) because the Spanish word miseria means extreme poverty.

After these changes were effected, the second experiment with 138 children was carried out. The IQ's of this group ranged from
62 to 145 with an average of 97.33 and a median of 95.97. The standard deviation was 15.75. A large number of questions were found to be in an incorrect order of difficulty and had to be rearranged. Tests VII and XII were found to be too difficult for our children. Many other questions did not seem clear to our children, thus becoming ambiguous and therefore difficult.

The following changes were effected in order to make the ambiguous questions clearer to our children. In V, 5, Memory for Sentences II, the original translation had too many words beginning with the K sound, so that it became a tongue twister rather than a regular sentence with a clear meaning to test memory. So the words were changed and the sentence, while retaining the same general idea, read like this: Maria quiere hacer un castillo grande en la arena (Mary wants to make a big castle in the sand). The original in English is as follows: Jane wants to build a big castle in her playhouse.

With reference to VII, 1a, Man with an umbrella, our children could not see the absurdity intended to be seen; that is, that the man in the picture has the umbrella in the wrong direction and is getting wet. Our children saw an absurdity all right, but not this one. Instead they would say, "Look, a man with an umbrella," and they would laugh. To them is was absurd for a man to carry an umbrella. That is something you do not see there; men do not use umbrellas. As a matter of fact, it is very rare to see a man wearing a hat. In revising this, we changed the picture to a woman with an umbrella.

In VII, 2b, Similarities: Two Things, we had to change the pair ship and automobile for airplane and automobile for many of our children have never seen a ship but have seen lots of planes and automobiles.

In VII, 5c, Opposite Analogies I, A bird flies, a fish. . . ., our children would complete the sentence by saying a fish does not fly. We revised it by adding: In the air a bird flies; in the water a fish. . . .

In VIII, 2a, Memory for Stories, The Wet Fall, we had to change the title, La Caída en el Charco, for La Caida Mojada. However, this second title didn't work either, and as a result of our third experiment, we had to ask the children themselves to select a title for the story. The title having the greatest frequency was finally selected. It reads as follows: El Caballito de Juan (John's Little Horse).
XII, 2c, Verbal Absurdities, refers to icebergs in the Gulf Stream; we had to change this to ice blocks that had been entirely melted by the heat. Almost all subjects missed the absurdity as originally presented. Puerto Rico is a land of perpetual summer where cold, icebergs, and snow represent foreign concepts.

In Average Adult, 3b, where we had changed misery for martyrdom we found that the question was then too easy. Four new pairs were selected for trial: admiration and envy, debt and gratitude, ambition and envy, and humility and poverty. In our third experiment the first pair proved to be the best question.

In Superior Adult III, 3b, Opposite Analogies II, the question *The pine tree is evergreen; the poplar is*... had to be changed, because in our land where there is no winter, trees are green all year around. We had substituted for the poplar the flamboyant, a well-known tree in our country, but even so no one could answer the question. We changed the analogy thus: *One side of a medal is called the obverse, the other the*... .

In addition to the changes mentioned previously, other minor modifications and changes in the order of presentation were effected.

One departure from the recommended method of presentation of the memory for digits test was made. All digits were organized in a series for trial in our third experiment. Fradsen had found no difference in presentation, and we confirmed his conclusion.

In our third experiment 717 subjects were used. They covered children from 5 to 17 years of age and from kindergarten to twelfth grades. Only the subtests to be tried were given to the appropriate groups. Nearly all of the revised questions gave satisfactory results, and in many instances in which we had changed the positions of questions, we found that they could be returned to their original positions.

In conclusion, we should say that a total of 889 subjects were used in the three experiments with the Stanford-Binet Scale. After all changes had been effected the scale was tried with 104 new subjects, and norms were established for Puerto Rico. These children were selected from different cities and towns. The IQ's ranged from 51 to 134 with a mean of 95.65 and a median of 95.83. The standard deviation was 18.16, and the probable error 12.24.

The probable error was used in determining seven categories which we classified as follows: The central category (one PE above and one below the mean) was denoted average. One PE above

---

and one below this were respectively denominated high average and low average. One PE above the high average classification was denominated superior and one below low average, inferior. The category above superior was called very superior and the one below inferior very inferior. In this way we avoided the classical nomenclature of genius, moron, borderline, etc. with their attendant emotional charge. In this we followed Wechsler’s point of view.

**The Goodenough Test**

The third instrument for measuring intelligence which I want to discuss here is the Goodenough Intelligence Test. In the process of adaptation we went through the same process used by the author in developing the scale. We followed the directions for scoring given in the original manual. This scale has to do with the drawing of a man, and points are given for each detail that is recognizable in the drawing. After all directions had been carefully checked, norms were established. A total of 2,100 children from 4 to 16 years of age were used; however, the number was reduced to 1,767 when the under-age and over-age children were not considered. These 1,767 children were all of normal age for the grade in which they were enrolled. Norms were established for ages 6 to 10.

Comparing these norms with the norms for American children, we found that Puerto Rican children scored lower than American children.

**Summary**

Our studies involving the translation and adaptation of three intelligence tests for use in the schools of Puerto Rico indicate that in general the Puerto Rican child will score lower than the American child. With the Wechsler Intelligence Scale for Children, the average IQ for Puerto Rican children was found to be 87.94, and with the Stanford-Binet, Revised, Form L, it was 95.65. In the case of the Goodenough test the norms for the different ages were also found to be lower with the exception of ages 5 and 6, which were selected from private schools. There is no doubt that no matter how well an intelligence scale is adapted from one culture to another, there are cultural differences, which will make the children for the second culture score lower than those from the first. The proper interpretation of these facts is to consider whatever average is obtained as equivalent to an IQ of 100.