Cross-Cultural Research

Cross-Cultural Testing of Intelligence

Nico Bleichrodt, René A. C. Hoksbergen and Usha Khire

The online version of this article can be found at:

Published by:

$SAGE

http://www.sagepublications.com

On behalf of:

Society for Cross-Cultural Research

Additional services and information for **Cross-Cultural Research** can be found at:

Email Alerts: http://ccr.sagepub.com/cgi/alerts

Subscriptions: http://ccr.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations: http://ccr.sagepub.com/content/33/1/3.refs.html

Cross-Cultural Testing of Intelligence

Nico Bleichrodt
Free University, Amsterdam

René A. C. Hoksbergen
University of Utrecht

Usha Khire
Institute of Psychology, Pune

Insight is sought into the usefulness and comparability of an intelligence test for schoolchildren living in India and in the Netherlands. A Dutch test, the Revisie Amsterdamse Kinder Intelligentie Test (RAKIT), was adapted to the Indian situation (Indian Child Intelligence Test [ICIT]). Using the ICIT and the RAKIT, 612 Indian children and 1,007 Dutch children were tested, respectively. The average test scores for the ICIT group were lower in a number of subtests; however, coefficients of internal consistency and stability for the ICIT and RAKIT were satisfactory, varying from 0.84 to 0.94. Results of a factor analysis showed that both test batteries have a clear psychometric equivalence, making comparable test interpretations possible. Predictive validities for ICIT and RAKIT were high (0.56 and 0.48, respectively), with reading ability as a criterion.

There are countless theoretical questions that give rise to the need for an appropriate intelligence measure: questions concerning the

3

connection between intelligence and learning, the question of the structure of intelligence in different developmental phases, the question of the relationship between heredity and environmental influences on the one hand and intelligence (or various identifiable facets of intelligence) on the other, and the question of the constancy and modifiability of intelligence at a younger age. Yet, other questions arise within the context of cross-cultural research, such as: What differences exist in the level and structure of intelligence or other mental capacities among children from various groups, for example, different linguistic groups, cultural groups, and social-economic classes? How are the different intelligence tests related to each other and to other variables such as personality factors, child-rearing practices, environmental factors, and so forth? And especially, do systematic differences with respect to such interrelationships exist between cultures or between other groups under study?

Practical questions for which an intelligence test is required also are legion, such as the determining of mental deficiency and organic brain damage, the readiness for school concerning admission to various types of schools, psychotherapeutic and orthopedic help, and pedagogical advice to parents. An important issue in intercountry or foreign adoption is how children will adapt to a new culture. Several studies (cf. Hoksbergen, 1995; Verhulst, Althaus, & Versluis-den Bieman, 1992) have examined later adjustment, finding that early negative experiences are of great influence on development in later life.

In all these cases, intelligence measurement is at least one of the first possibilities for investigation.

Present theories on the structure of intelligence can be classified into one of two main schools: hierarchical theories of intelligence in which a single aspect—general intelligence—dominates over other, specific aspects of intelligence (Burt, 1949; Cattell, 1971; Vernon, 1960) and overlapping, nonhierarchical or multifactorial theories in which intelligence is thought to consist of a number of dimensions that are not entirely independent (Guilford,

1967). However, a large body of research has shown that a combination of both types of theory is probably the most realistic approach (cf. Carroll, 1993; Sternberg, 1982, 1988), although in recent studies, the hierarchical model of intelligence is given more and more preference (Anastasi, 1994; Carroll 1993; Gustafsson, 1984, 1989; Marshalek, Lohman, & Snow, 1983). The fact that there is such a thing as a general intelligence factor may be deduced from the correlation between scores on a wide variety of tasks in an intelligence test. However, as the overlap is not complete, clearly there are also more specific factors. Methodologically, it has been shown that the multifactor solution and the hierarchical solution are mathematically equivalent and transposable from one to the other (Harman, 1976).

The American psychologist Thurstone (1938, 1947) is a representative of this middle-of-the-road standpoint. A direct outgrowth of Thurstone's theory of cognitive abilities—but also of the factor models proposed by Spearman (1927), Vernon (1960), and Cattell (1971)—is the three-stratum theory of cognitive abilities developed by Carroll (1993), in the sense that "it relies on successive factorizations of correlation matrices of higher orders" (p. 637). The abilities are classified on three strata (narrow, broad, and general) according to the generality over the total domain. Thurstone (1938, 1947) assumes that intelligence comprises a number of distinct abilities occurring in ever-differing combinations in complex tasks, with a general, broader underlying factor. On the basis of outcomes of factor-analytic research, he has identified seven primary mental abilities that together comprise the domain of intelligence—that is, verbal, word fluency, number, spatial, memory, perceptual speed, and reasoning. Many intelligence tests are based on Thurstone's factors or on elaborations of these factors.

STUDY GOAL

A plan was drawn up in cooperation with authorities and professionals from institutes in India and the Netherlands[1] involved in child welfare. These institutes primarily were interested in an instrument that could be used to compare the performances of different groups of children, especially children adopted by Indian and

by Dutch couples, children growing up in their biological families, and children living in a residential setting. One of the relevant aspects in this comparison is the level of cognitive ability.

The main goal of this study was to develop an intelligence test that could be used in the two different cultures. Then, at a later stage, this test could be used to gather data on the cognitive abilities of the specific groups of children just mentioned.

SELECTION OF THE TEST

The first step was to examine the IQ tests used in India to adapt one for use in the Netherlands. After an extensive inventory (Khire, Bleichrodt, Hoksbergen, & Bharat, 1992), it was found that in India, insufficient research had been done on psychometric aspects and on the suitability of such tests for Indian children. Most tests proved to be translations of Western instruments with occasional adaptations to Indian conditions. Other tests were appropriate for one specific group but not for the group of children we wanted to examine, namely, schoolchildren between the ages of 6 and 12. Finally, a number of other tests dating from the 1950s and 1960s were identified, but these required major updating.

An important problem with most of the tests was of a practical nature: Many lacked a useful instruction manual, or the available manual did not include the psychometric information required to verify the validity and reliability. In short, it soon became clear that either a new IQ instrument appropriate to school-age children would have to be developed or an existing, foreign test—preferably a Dutch one—would need to be adapted for Indian children. The latter was the more attractive option for a variety of reasons—that is, cost, speed of completion, and readiness for use in Dutch culture.

It was decided to adapt a Dutch test, the Revised Amsterdam Child Intelligence Test (RAKIT) (Bleichrodt, Drenth, Zaal, & Resing, 1984) to the study's requirements. Sufficient experience has been built up during the past decade in different cultures (e.g., Indonesia, Kenya, Uganda, Tanzania, Surinam, Spain, Malawi, and Gambia) to allow its utility to be assessed (Bleichrodt, Escobar del Rey, Morreale de Escobar, Garcia, & Rubio, 1989; Drenth, Bleichrodt, Setiono, & Poespadibrata, 1975; Drenth et al., 1980; Resing, Bleichrodt, & Drenth, 1986; Sijtsma & Resing, 1991). The

RAKIT is an intelligence test that can be used both for determining the general intelligence level and for assessing more specific intelligence factors (Bleichrodt, Resing, Drenth, & Zaal, 1987). This view of intelligence as a complex, multisided, and structural concept is more meaningful and relevant than is the idea that intelligence is a one-dimensional quality. To obtain information on specific intelligence factors, the various parts of the test series must cover a broad spectrum of intellectual capabilities.

The RAKIT concept relates closely to Thurstone's (1938, 1947) primary factor theory supplemented with associative memory and a number of performance tasks. The RAKIT is an individual test suitable for children between the ages of 4 and 12 and consists of 12 subtests. The appendix gives an overview of the subtests of the Indian Child Intelligence Test (ICIT) and the RAKIT and of the number of items in each subtest.

METHOD

To carry out a pilot study, each subtest was studied and then self-administered and instructions for administration were translated into Marathi, the main language of Maharashtra, a state more than 300,000 km^2 in area and with approximately 70 million inhabitants, including the city of Pune with its 2.5 million inhabitants. Open-ended tryouts of all subtests were administered to a group of 40 children in the age range of 6 to 12 years. The children were asked to think aloud while deciding the answers. Their perception of items, experience, and familiarity were revealed by these tryouts. The primary purpose was to judge the suitability of the test content for Indian children. As a result of this pilot study, three subtests (Analogies, Idea Production, and Story Telling) were not included in the Indian version. The main reason for this was that several of the items in these tests described typical Dutch objects, products, and situations. Although these more verbal tests were omitted, the verbal factor is represented in the ICIT by the Verbal Meaning subtest. For two subtests—Closure (50 items) and Verbal Meaning (60 items)—a number of items (10 and 24, respectively) were adapted for Indian children (see Figure 1).

Before finalizing the test, these changes were tried out and tested in a second pilot study involving a group of 20 children. The
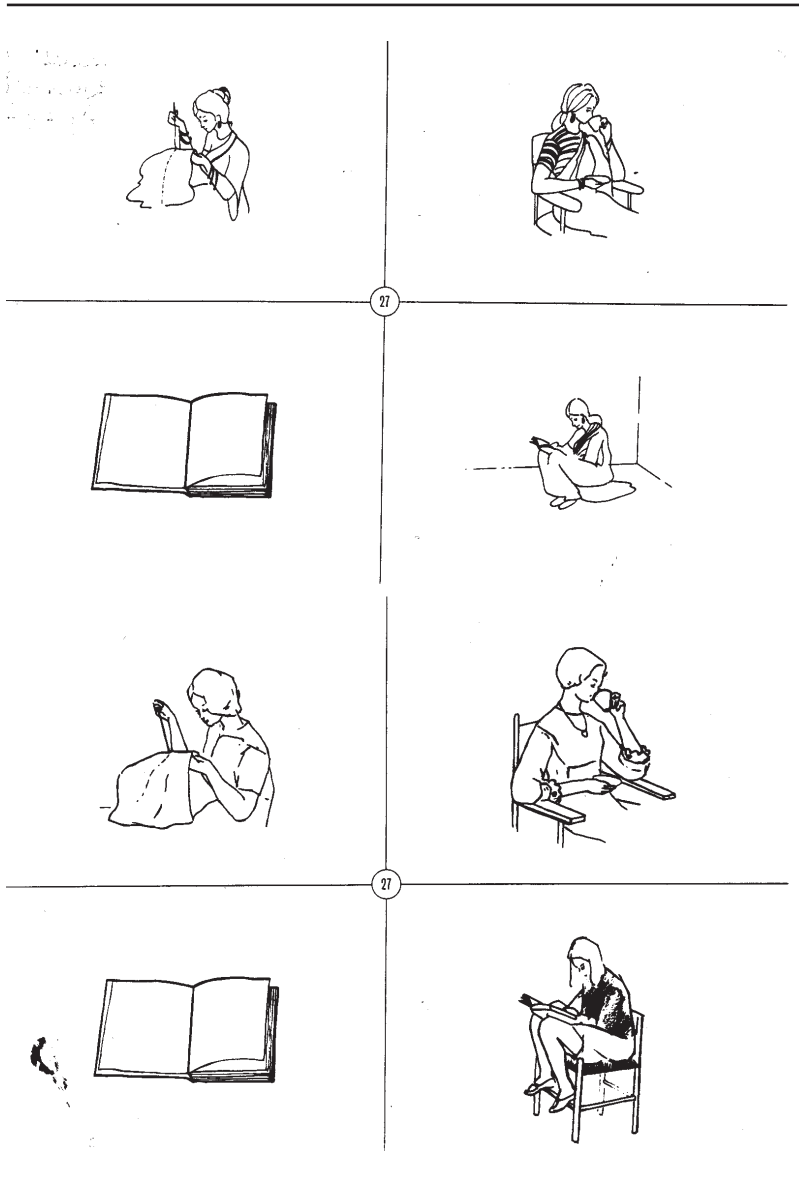
Figure 1:   Item of the Verbal Meaning Subtest: "Which Picture Belongs to the Word *Read*?"
NOTE: Indian Child Intelligence Test (ICIT) is shown in the top four squares, and the Revisie Amsterdamse Kinder Intelligentie Test (RAKIT) is shown in the bottom four.

TABLE 1
Number of Boys and Girls in ICIT and RAKIT Sample

Group	ICIT			RAKIT		
	Boys	Girls	Total	Boys	Girls	Total
6 and 7 years	101	104	205	194	198	392
8 to 11 years	206	201	407	308	307	615
Total	307	305	612	502	505	1,007

NOTE: RAKIT = Revisie Amsterdamse Kinder Intelligentie Test; ICIT = Indian Child Intelligence Test.

definitive version of the ICIT then was administered to a sample of 612 children (see Table 1). The original RAKIT sample consisted of 1,007 children in the age range of 6 to 12 years—the range on which the norms for the RAKIT are based. To make the sample as representative as possible and to reduce unwanted sample fluctuations, the following stratification criteria were chosen: region (all parts of the Netherlands), urbanization level (urban and rural areas), size of schools, age, and gender. A random sample of approximately 100 schools was chosen, taking into account the size of the school. From each school, 10 children again were selected randomly (1 boy and 1 girl from each school's five classes).

A comparable procedure was used for the ICIT sample. It was not easy to find a representative sample, however, as the composition of the school population in the Pune area is decidedly diverse. Therefore, the most random selection possible was made across the entire region. For this purpose, the Pune region was divided into 11 areas: central, north, east, west, southwest, south, and 5 surrounding suburbs. A total of 19 schools were selected on the basis of the representativeness of their locations and their willingness to participate in the project. Larger schools provided more students; consequently, a larger number of students came from central Pune (Khire et al., 1992).

For both the RAKIT and ICIT samples, the number of boys and girls per age category was approximately the same (see Table 1). Care also was taken to represent pupils of all socioeconomic classes in the samples. However, there is one important difference between the two samples: In the Netherlands, education is compulsory and all children are sent to school, which means that the RAKIT sample is representative of the population of children

between the ages 6 and 12 years as a whole. The situation in India is different: Although education is compulsory, a large number of children—mostly those from lower socioeconomic backgrounds—either do not attend school or fail to complete primary school.

RESULTS OF THE COMPARATIVE
STUDY OF ICIT AND RAKIT

Six Indian psychologists who had been trained to use the ICIT administered the test at the schools in the random sample. On the basis of the ICIT test results of the 612 Indian children and of the RAKIT test results of 1,007 Dutch children, an analysis of the comparability of the two tests was undertaken. This included an examination of the level of difficulty of the ICIT and RAKIT with a subsequent review of the reliability and validity of data.

DIFFICULTY

The items for the subtests are arranged according to level of difficulty. Because the test is applicable to a wide age range, items are arranged in two overlapping sets: The first part of the subtest (e.g., Items 1 to 45) is intended for 6- to 7-year-olds, and the whole test (e.g., Items 1 to 50) for 8- to 12-year-olds (see appendix).

The average scores and standard deviations of the two age groups (6 to 7 and 8 to 12) are presented in Table 2. Multivariate analysis showed that for both age groups, significant differences ($F < 0.001$) existed between the subtest scores of Dutch and Indian children. For the younger age group, the univariate F tests were significant ($F < 0.001$), with the exception of the Memory Span subtest. For the older age group, F tests for three subtests (Exclusion, Memory Span, and Verbal Meaning) were not significant.

Because the number of children on which the statistical data were based was very large, significant differences in test scores turned out to be rather small. A commonly used measure to indicate the size or strength of a difference is the d value, where d is the difference between the two group means divided by the (mean of the) standard deviation. Cohen (1988) classified d values into small effect size ($d = 0.2$), medium effect size ($d = 0.5$), and large

TABLE 2
Means (raw scores), Standard Deviations,
and *d* Values for the Indian and Dutch Samples

Tests	6 to 7 Years					8 to 12 Years				
	ICIT N = 205		RAKIT N = 392		d	ICIT N = 407		RAKIT N = 615		d
Closure	21.3	8.1	30.0	5.6*	1.3	31.7	6.7	37.2	5.3*	0.9
Exclusion	32.8	6.2	30.6	6.1*	0.4	39.0	5.9	38.5	5.9	0.1
Memory Span	6.6	2.1	6.9	2.0	0.1	8.5	2.4	8.7	2.1	0.1
Verbal Meaning	33.7	6.3	36.7	4.6*	0.6	46.6	6.4	46.1	4.8	0.1
Mazes	61.9	9.8	66.1	8.3*	0.5	73.6	8.7	75.4	7.1*	0.2
Quantity	41.1	8.2	43.0	8.0*	0.2	52.6	6.5	55.6	5.5*	0.5
Discs[a]	493	149	285	98*	1.7	335	123	201	75*	1.4
Learning Names	9.9	4.6	11.5	3.6*	0.4	15.8	4.2	16.6	3.9*	0.2
Hidden Figures	26.4	5.0	23.7	6.5*	0.5	34.1	5.3	32.3	6.1*	0.3

NOTE: RAKIT = Revisie Amsterdamse Kinder Intelligentie Test; ICIT = Indian Child Intelligence Test.
a. Scoring is based on the time of completion. Lower scores are better.
*The differences between ICIT and RAKIT means are significant, $p < .01$.

effect size ($d = 0.8$). For nearly all subtests, the differences between the scores of Dutch and Indian children were small or medium. Sometimes, the average scores of the Indian children were higher than were those of the Dutch children, and sometimes the reverse was the case. For two subtests (Closure and Discs), both age groups showed a large effect size ($d > 0.8$).

In both age groups, the standard deviations of the subtests were larger in the ICIT sample, with the exception of the Hidden Figures subtest.

RELIABILITY

The reliability of the tests for both age groups taken together was checked by examining the internal consistency and stability of the tests (see Table 3). The coefficients of internal consistency, calculated with the (odd-even) split-half method, were high for both the ICIT and the RAKIT (0.92 and 0.94, respectively). The reliability coefficient for nearly all of the individual subtests was above

TABLE 3
Coefficients of Internal Consistency, Coefficients
of Stability, and Group Differences (Indian vs. Dutch)
on the Coefficients (z)

Tests	Internal Consistency (6-12 years)				Stability (6-12 years)		
	ICIT N = 612	RAKIT N = 1,007	z		ICIT N = 168	RAKIT N = 149	z
Closure	.93	.84	11.43 [a]		.82	.82	0.05
Exclusion	.87	.87	−0.48		.65	.62	0.44
Memory Span	.84	.79	3.72 [a]		.62	.59	0.42
Verbal Meaning	.88	.80	7.04 [a]		.58	.59	−0.13
Mazes	.83	.84	−0.80		.53	.61	−1.05
Quantity	.87	.87	0.57		.66	.59	1.01
Discs	.76	.70	3.01 [a]		.63	.65	−0.30
Learning Names	.75	.82	−4.36 [a]		.75	.78	−0.64
Hidden Figures	.83	.86	−2.61 [a]		.39	.57	−2.07
Total test	.92	.94	−2.90 [a]		.84	.87	−0.98

NOTE: RAKIT = Revisie Amsterdamse Kinder Intelligentie Test; ICIT = Indian Child Intelligence Test.
a. Significant at 1% level (two-sided).

0.80. The lowest correlations were found for the Discs subtest. The differences between the coefficients of internal consistency for the two countries were tested by the procedure described in Hakstian and Whalen (1976). For six subtests, the coefficients differed significantly at the 1% level; some of the RAKIT subtests were more reliable than were the ICIT subtests, and for some other subtests, the reverse was the case.

To check the stability of the tests, 168 Indian and 149 Dutch children were retested. These children were selected randomly from three age groups (6-, 8-, and 10-year-olds) from the total ICIT and RAKIT sample (see Table 4). The interval between the first and second test was 3 to 4 months.

The coefficients of stability for ICIT and RAKIT were 0.84 and 0.87, respectively. With one exception (Hidden Figures in ICIT), the coefficients for each subtest were higher than 0.50. The differences between the coefficients of stability were tested by comparing the Fisher's transforms of the correlations. No significant differences were found between any of the paired correlations.

TABLE 4
Sample for the Test-Retest Study

	6 Years	8 Years	10 Years	Total
ICIT	50	60	58	168
RAKIT	49	51	49	149

VALIDITY

In cross-cultural validation research, it is crucial to know to what extent the meaning of the test battery is comparable for the different cultures. A comparison of the factorial structure of both test batteries was made; this was followed by an analysis of the links between test scores and a criterion variable. Moreover, the influence of two personal characteristics (age and sex) was analyzed.

For ICIT and RAKIT, a positive relationship was found between the age of the children studied and the test results: The older groups scored higher, which was why age norms were established for both tests. Given the significance of the correlation between age and test scores, it was decided that 4-month norms would be used for the RAKIT and half-yearly norms for the ICIT. This was because of the lower correlations between age and test scores for the ICIT sample. Table 5 shows that within the norm groups, there was no significant correlation between age and test scores and so there was no point in developing even more specific norm groups. The raw scores were converted into standard scores with a mean of 15 and a standard deviation of 5.

In general, the differences in test scores between boys and girls were not large. The most important difference for both test series lay in the Mazes subtest, which had a strong accent on spatial reasoning. Here, boys scored higher than girls (see Table 5; $p < 0.01$, two-sided).

The differences between the correlations with age and sex across countries were tested by comparing the Fisher's transforms of correlations. Only for the sex variable did the correlations of two subtests, Closure and Mazes, differ significantly ($p < 0.01$, two-sided). The differences between boys and girls for both subtests were larger for the Indian group than for the Dutch group.

TABLE 5
Correlations of Age and Sex With ICIT and
RAKIT Scores and Group Differences (Indian vs. Dutch)
on the Correlation Coefficients (z)

	Age			Sex		
	ICIT N = 612	RAKIT N = 1,007		ICIT N = 612	RAKIT N = 1,641	
Tests	(6-12 years)	(6-12 years)	z	(6-12 years)	(4-12 years)	z
Closure	−.07	.04	−2.14	.18[a]	−.04	4.68 [a]
Exclusion	.03	.01	0.38	−.05	−.03	−0.42
Memory Span	−.04	.03	−1.36	−.01	−.09	1.69
Verbal Meaning	−.01	.04	−0.97	.09[a]	.10 [a]	−0.21
Mazes	−.02	.07	−1.75	.29[a]	.14[a]	3.32 [a]
Quantity	−.01	.01	−0.39	.07	.03	0.85
Discs	−.01	.02	−0.58	−.04	.01	−1.05
Learning Names	.00	−.02	0.39	−.00	−.03	−0.63
Hidden Figures	−.03	.07	−1.95	.01	−.08 [a]	1.90

NOTE: RAKIT = Revisie Amsterdamse Kinder Intelligentie Test; ICIT = Indian Child Intelligence Test. Boys = 1; girls = 0.
a. Significant at 1% level (two-sided).

To examine the factor structure of both test batteries, it was decided to carry out a principal component analysis with varimax and Procrustes rotation rather than a confirmatory factor analysis such as LISREL. Bookstein (1986) states that "LISREL will frequently, but erratically, fail to produce meaningful results at all" (p. 228). McCrae, Zonderman, Costa, Bond, and Paunonen (1996) conclude that there are no theoretical reasons why intelligence dimensions should not have meaningful loadings on more than one factor. When using a confirmatory factor analytic technique, one needs to build in both secondary loadings and oblique rotations. However, even then, the models usually have poor fits unless several difficult-to-interpret and very specific factors are built into the model. Moreover, large samples more often lead to rejection of the postulated model. Therefore, McCrae et al. (1996) postulated that orthogonal Procrustes rotation "offers a powerful technique for hypotheses-guided rotation" (p. 559).

On the basis of the intercorrelations of the subtests, factor analyses for the total ICIT and RAKIT groups were carried out. A principal component analysis method with orthogonal varimax and Procrustes rotation was used. The results of these factor

TABLE 6
Factor Matrices After Orthogonal
Varimax Rotation (loadings > .35)

| | Orthogonal Varimax Rotation | | | | | | Procrustes Rotation | | |
| | RAKIT | | | ICIT | | | ICIT | | |
	I	II	III	I	II	III	I	II	III
Closure		.7	.4		.7			.6	
Exclusion	.6	.4		.4	.6		.5	.5	
Memory Span	.8			.8			.7		
Verbal Meaning			.7	.4	.4	.7			.9
Mazes		.7			.7			.7	
Quantity	.6			.4	.6		.5	.5	
Discs	.5	.6		.4	.5		.7	.4	
Learning Names			.8	.8			.6		.5
Hidden Figures	.4	.4			.7			.6	

NOTE: RAKIT = Revisie Amsterdamse Kinder Intelligentie Test; ICIT = Indian Child Intelligence Test.

analyses are set out in Table 6. The factor matrix of the RAKIT was used as the target matrix. On the basis of the curve of the eigenvalues on one hand and the comparability and interpretability of both factor structures on the other, it was decided to work with three rotated factors. In general, the factor matrices of ICIT and RAKIT appeared to show comparable structures, with notable factors being (a) General Reasoning and Memory, (b) Perceptual and Spatial Performance, and (c) Verbal Learning Ability.

For measuring the degree of invariance of the different factors, Tucker's phi coefficients were calculated. If the coefficient is higher than 0.85, the two factors concerned are considered virtually equal (Ten Berge, 1977). The comparability of the factor matrices of ICIT and RAKIT was high, with phi coefficients for Factors I, II, and III of 0.87, 0.90, and 0.89, respectively.

To gain an impression of the predictive validity of both test series, the correlation coefficients were calculated between test scores and scores determined for reading ability (see Table 7). In India and the Netherlands, part of the group of children from the norm sample was tested for reading skills. In both countries, the correlations for all subtests were significant ($p < 0.01$) with the exception of Closure and Quantity from ICIT. The correlations matched closely for the Dutch and Indian groups. No significant

TABLE 7
Predictive Validity: Correlations of Test Scores
With Reading Ability and Differences Between the
Correlations (*z*) and Regression Analysis for the Test Series
(β weights) and Differences Between β Weights (*t*)

Tests	Test Scores-Reading Ability Correlations			Regression Analysis (β weights)		
	ICIT (N = 156)	*RAKIT* (N = 1,007)	z	*ICIT* (N = 156)	*RAKIT* (N = 1,007)	t
Closure	.10	.19[a]	−1.069	−.149	.017	−1.972
Exclusion	.42[a]	.28[a]	1.859	.266	.081	2.002
Memory Span	.28[a]	.26[a]	0.251	.140	.100	0.493
Verbal Meaning	.35[a]	.32[a]	0.393	.234	.130	1.214
Mazes	.26[a]	.07[a]	2.277	.123	−.071	2.348
Quantity	.17	.31[a]	−1.730	−.186	.115	−3.255[a]
Discs	.25[a]	.19[a]	0.733	.123	−.111	1.581
Learning Names	.29[a]	.33[a]	−0.514	.069	.175	−1.209
Hidden Figures	.33[a]	.33[a]	0.000	.141	.143	−0.014

NOTE: RAKIT = Revisie Amsterdamse Kinder Intelligentie Test; ICIT = Indian Child Intelligence Test.
a. Significant at 1% level (two-sided).

difference ($p < 0.01$, two-sided) was found between any of the paired correlations.

In general, the results of the study of the predictability of the tests for younger children show rather low values. The correlations for the ICIT subtests varied from 0.10 to 0.42 and for the RAKIT subtests from 0.07 to 0.33, with a median for both series of 0.28.

To determine whether the β weights of the predictors (subtests) of reading ability were different for ICIT and RAKIT, a direct regression analysis was conducted. In Table 7, the β weights for the different subtests of ICIT and RAKIT are presented. The overall test of significance of the differences between the β weights of each of the paired subtests gave a chi-square of 24.72 ($df = 9$), which was significant at the 1% level. Results of the univariate analyses showed a significant difference ($p < 0.01$) for the Quantity subtest. The differences in β weights for the other subtests were not significant. The multiple correlations for the Indian and Dutch groups of children were 0.56 and 0.48, respectively.

DISCUSSION

One of the traditional problems confronting cross-cultural psychology concerns the comparability of test results of children and adults with different ethnic backgrounds.

In this study, attention was paid to the usefulness and comparability of an intelligence test for schoolchildren living in India and in the Netherlands.

From the research of Hofstede (1991), it appears that there are large differences between the two cultures in the following dimensions, which he distinguished: power distance, uncertainty avoidance, individualism, and masculinity. Compared with Dutch culture, Indian culture scores higher values on the power distance index (e.g., greater dependence on subordinates), the masculinity index (sexual roles clearly separated), and the uncertainty avoidance index (feelings of threat brought about by unsure and unknown situations) and a lower value on the individuality index (Indian society is oriented more toward collectivity than is Dutch society). In these quite different cultures, it was investigated whether it is possible to apply comparable tests (i.e., ICIT and RAKIT) to determine in a reliable manner the intellectual levels of children in India and Holland, respectively. Based on the results of the research, an attempt could be made to answer the question of whether "certain forms of intelligence do not arise in cultures in which these forms are not valued or in which they are unintelligible in terms of indigenous cultural understandings" (Miller, 1997, p. 291). A Dutch test (RAKIT) was adapted for the Indian situation (ICIT). Six subtests in ICIT and RAKIT were identical. Items of three other subtests had to be adapted. Some of the pictures used in the Verbal Meaning and Closure tests were adapted to the Indian culture. These adaptations concerned matters of personal appearance and clothing, artifacts, family situations, street scenes, and architecture. Naturally, the names employed in the Learning Names subtest were different in the ICIT and the RAKIT.

In two subtests, Closure and Discs, large differences ($d > 0.8$) in test means for both age groups were found. Closure—a test consisting of incomplete (concrete) figures—required major adjustments. A number of pictures were adjusted to the Indian situation. On the whole, these new, incomplete figures were more difficult than were the corresponding RAKIT items. However, the difficulty indexes

for those items that remained the same scarcely differed, which probably means that Closure test tasks are not much harder for children in India. A possible explanation for the significantly lower scores for the Discs subtest might be Indian children's reduced familiarity with this kind of material. Younger children in the Netherlands often play at home with toys (putting discs on a pin) that are similar to the Disc test tasks.

Although the average test scores for the ICIT group were lower in a number of subtests, in these cases, the effect sizes could be categorized as either medium or small. These lower scores for the group of Indian children may have been caused by the fact that Indian children start school rather later than do their Dutch counterparts. In the Netherlands, parents legally are obliged to send their children to school from the age of 5, and 95% of Dutch children actually have started school by the age of 4. This means that the 7-year-olds in the Indian sample had, on average, less school education than did the 7-year-olds in the Dutch sample. Unfortunately, reliable figures on this matter are not available. In the older age group (8-12 years), the differences in average test score were smaller; that is, the older groups of Dutch and Indian children gained more closely comparable test scores. We suggest the following explanation for this fact: First, it is reasonable to suppose that the longer a child is at school, the smaller the influence exerted by any difference in the total duration of school attendance. Disadvantaged students will have had more time in which to do some catching up. This also can be distinguished from the data collected for the stability study (see Table 4). A number of children took the same test twice after an interval of 3 to 4 months. In both the ICIT and the RAKIT groups, the average scores for the second test were higher (apart from the Quantity subtest for the youngest RAKIT group). Table 8 shows the d values: the difference between the two mean scores of the first and second test administration divided by the standard deviation of the scores for the first administration. In the youngest group, these d values are larger for the ICIT, and only in the Verbal Meaning subtest are the d values equal. This means that the learning and habituation effects are stronger for the Indian children than for the Dutch children. For the older group of children (8-12 years), this is not the case; the differences in d value are negligible. In other words, the rise in test scores is comparable for both groups, and there is no longer a catch-up effect in the ICIT group.

TABLE 8
Difference in mean scores for test and retest divided by
standard deviation of the test scores: _d_ values

Tests	6 - 8 Years		8 - 12 Years	
	ICIT	_RAKIT_	_ICIT_	_RAKIT_
Closure	.67	.57	.52	.73
Exclusion	.36	.08	.23	.29
Memory Span	.65	.29	.48	.27
Verbal Meaning	.23	.23	.03	.18
Mazes	.99	.62	.83	.88
Quantity	.31	−.07	.22	.13
Discs	.53	.30	.58	.50
Learning Names	1.19	.95	.79	.81
Hidden Figures	.27	.11	.39	.35

NOTE: RAKIT = Revisie Amsterdamse Kinder Intelligentie Test; ICIT = Indian Child Intelligence Test.

A second explanation could lie in the fact that part of the group of older children in India leaves school earlier in order to work; this applies especially to less academically successful children from lower socioeconomic groups. In other words, the higher classes will contain a more selective—and therefore more homogenous—group of children.

For two subtests—Exclusion and Hidden Figures—the Indian children gained higher average scores; the differences, however, were quite small. These tests are a measure of children's perceptual reasoning, and the test material cannot be said to be very culturally sensitive, as it comprises abstract or simple images.

The coefficients of internal consistency for ICIT and RAKIT were high: 0.92 and 0.94, respectively. The differences in subtest reliabilities between both groups generally are associated with differences in standard deviations. Closure, Memory Span, Verbal Meaning, and Discs subtests from the ICIT and the Hidden Figures subtest from the RAKIT had larger standard deviations as well as higher coefficients of internal consistency. Only the Learning Names subtest showed the reverse.

No significant difference in stability coefficient was found for any of the ICIT or RAKIT subtests. For the older group of children,

the stability coefficients for the ICIT and RAKIT subtests were practically identical, with averages of 0.69 and 0.67, respectively. For the younger group of children, the stability coefficient was a little lower for the ICIT than for the RAKIT, namely, 0.54 as compared to 0.67. This could be a consequence of the above-mentioned learning and habituation effect, which is not equally strong for all children.

Taken as a whole, boys performed better than girls on the ICIT. This could be because in India, boys in particular are stimulated to do well at school (Desai, 1991; Krishnaraj & Chanana, 1989). In three subtests—Closure, Verbal Meaning, and Mazes—Indian boys scored significantly higher than did girls. For the last two of these three subtests, the same applied to the Dutch children, whereas Dutch girls scored better than Dutch boys on the Hidden Figures subtest.

These results largely concur with the findings of earlier studies of the significance of sex-related differences in intelligence scores. On the basis of the most important of these studies (Born, Bleichrodt, & Van der Flier, 1987; Guilford, 1967; Hakstian & Cattell, 1975; Hyde, 1981; Maccoby & Jacklin, 1974; Sherman, 1978), it can be concluded that males do indeed score higher on Spatial (the Mazes subtest) and Closure (the Closure subtest) intelligence factors, and females score higher on Memory (the Memory Span subtest) and Perceptual Speed (the Hidden Figures subtest). Guilford (1967), Sherman (1978), and Hyde (1981) also found that females scored significantly higher on the verbal intelligence factor. A study by Hakstian and Cattell (1975) found no difference in verbal skills between men and women, and a study by Maccoby and Jacklin (1974) found no such differences between boys and girls of primary school age. Nevertheless, in our research, the girls did better in the Verbal Meaning subtest both in India and in the Netherlands.

Factor analyses were carried out to obtain an insight into the similarities between the results of the ICIT and RAKIT tests. The findings of these factor analyses, after varimax and Procrustes rotation, respectively, gave three factors for both test batteries: General Reasoning and Memory, Perceptual and Spatial Performance, and Verbal Learning Ability. There is strong agreement in factor structure and therefore clear psychometric equivalence between ICIT and RAKIT. Only the Learning Names subtest would appear to have a slightly different meaning in each of the two cultures: In

the Netherlands, the test has a more verbal emphasis, whereas in India, the test also makes more of a demand on memory.

The predictive validity also will depend on the choice of criterion and on the composition of the research subject group. In general, correlations between the children's overall intelligence test scores and their average school performance lie between 0.40 and 0.50. Obviously, correlations between individual subtests within the intelligence test battery and school performance are lower.

One of the most important subjects within primary education is that of reading. The assessment of this subject was chosen as a criterion by which to obtain an indication of the predictive validity of ICIT and RAKIT. Correlations between the subtests of both test batteries and the criterion were all positive. The multiple correlation coefficients of the ICIT and RAKIT were high and quite similar: 0.56 and 0.48, respectively. The ICIT's somewhat higher validity coefficient might have been the consequence of the wider spread of test and criterion scores within the ICIT group.

To conclude, the Indian and the Dutch intelligence tests are strongly comparable in terms of difficulty, reliability, psychological meaning, and predictive validity.

APPENDIX
Description of ICIT and RAKIT

The RAKIT comprises 12 verbal and nonverbal subtests. Three subtests (Analogies, Idea Production, and Story Telling) were not included in the ICIT. The purpose of each test is as follows:

Closure: To recognize and name incomplete pictures (Items 1-45, 1-50).

Exclusion: To identify the one figure that does not belong in a series (Items 1-45, 1-50).

Memory Span: To remember and reproduce the serial order of pictures (Items 1-36).

Verbal Meaning: To select from four different pictures the one that corresponds to a stimulus word (Items 1-50, 1-60).

Mazes: To trace the correct route to the exit as quickly as possible in mazes of various degrees of difficulty (Items 1-14).

Analogies: To find two word pairs that display an identical relationship (a : b = c : ?).

Quantity: To perceive quantitative concepts such as length, distance, area, volume, weight, and number (Items 1-55, 1-65).

Discs: To place two layers of discs with holes over a set of matching pins (Items 1-18).

Learning Names: To learn and recall the names associated with pictures (2×12 items).

Hidden Figures: To search for the correct simple figure embedded in a larger complex pattern (Items 1-45).

Idea Production: To mention as many things as possible belonging to a special category.

Storytelling: To tell as much as possible about a picture.

NOTE: RAKIT = Revisie Amsterdamse Kinder Intelligentie Test; ICIT = Indian Child Intelligence Test.

Note

1. Both the Adoption Centre for Research and Counselling in Utrecht and the Netherlands' Intercountry Child Welfare Organization maintain intensive and fruitful contacts with Indian authorities with regard to social welfare, child protection, and in-country and intercountry adoption.

References

Anastasi, A. (1994). Aptitude testing. In *Encyclopedia of human behavior* (Vol. 1, pp. 211-221). San Diego, CA: Academic Press.

Bleichrodt, N., Drenth, P.J.D., Zaal, J. N., & Resing, W.C.M. (1984). *Revisie van de Amsterdamse Kinder Intelligentie Test* [Revised Amsterdam Child Intelligence Test]. Lisse: Swets & Zeitlinger B.V.

Bleichrodt, N., Escobar del Rey, F., Morreale de Escobar, G., Garcia, I., & Rubio, C. (1989). Iodine deficiency, implications for mental and psychomotor development in children. In G. R. DeLong (Ed.), *Iodine and the brain* (pp. 269-289). New York: Plenum.

Bleichrodt, N., Resing, W.C.M., Drenth, P.J.D., & Zaal, J. N. (1987). *Intelligentie-meting bij kinderen* [Assessment of childrens' intelligence]. Lisse: Swets & Zeitlinger B.V.

Bookstein, F. L. (1986). The elements of latent variable models: A cautionary lecture. In M. E. Lamb, A. L. Brown, & B. Rogoff (Eds.), *Advances in developmental psychology* (Vol. 4, pp. 203-230). Hillsdale, NJ: Lawrence Erlbaum.

Born, M.P., Bleichrodt, N., & Van der Flier, H. (1987). Cross-cultural comparison of sex-related differences on intelligence tests. *Journal of Cross-Cultural Psychology, 18,* 283-314.

Burt, C. (1949). The structure of the mind; a review of the results of factor analysis. *British Journal of Educational Psychology, 19*, 110-111, 176-199.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, MA: University Press.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Desai, M. (1991). (Ed.). *Research on families with problems in India: Issues and implications*. Bombay, India: Tata Institute of Social Sciences.

Drenth, P.J.D., Bleichrodt, N., Setiono, T., & Poespadibrata, S. (1975). *Test Intelligensi Anak* [Child intelligence test]. Amsterdam: Vrije Universiteit.

Drenth, P.J.D., Van der Flier, H., Muinde, N. P., Otaala, B., Omari, I. M., & Opolot, J. A. (1980). Jaribio Akili Mtoto Afrika [African Child Intelligence Test]. Amsterdam: Vrije Universiteit.

Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

Gustafsson, J.-E. (1984). A unifying for the structure of intellectual abilities. *Intelligence, 8*, 179-203.

Gustafsson, J.-E. (1989). Broad and narrow abilities in research on learning and instruction. In R. Kanfer, P. L. Ackerman, & R. Cudeck (Eds.), *Abilities, motivation, and measurement* (pp. 203-237). Hillsdale, NJ: Lawrence Erlbaum.

Hakstian, A. R., & Cattell, R. B. (1975). An examination of adolescent sex differences in some ability and personality traits. *Canadian Journal of Behavioral Science, 7*, 295-312.

Hakstian, A. R., & Whalen, T. E. (1976). A k-sample significance test for independent alpha coefficients. *Psychometrika, 41*, 219-231.

Harman, H. H. (1976). *Modern factor analysis*. Chicago: University of Chicago Press.

Hofstede, G. (1991). *Cultures and organizations. Software of the mind*. New York: McGraw-Hill.

Hoksbergen, R.A.C. (1995). *Adopting a child*. Utrecht: Adoption Centre.

Hyde, J. S. (1981). How large are cognitive gender differences? A meta-analyses using w^2 and *d. American Psychologist, 36*, 892-901.

Khire, U., Bleichrodt, N., Hoksbergen, R.A.C., & Bharat, S. (1992). *Indian Child Intelligence Test (ICIT). Technical manual*. Pune, India: Anand Agencies.

Krishnaraj, M., & Chanana, K. (1989). (Eds.). *Gender and the household domain: Social and cultural dimensions*. New Delhi: Sage Ltd.

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.

Marshalek, B., Lohman, D. F., & Snow, R. E. (1983). The complexity continuum in the radex and hierarchical models of intelligence. *Intelligence, 7*, 107-128.

McCrae, R. R., Zonderman, A. B., Costa, P. T., Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors in the revised NEO Per-

sonality Inventory: Confirmatory factor analyses versus Procrustes rotation. *Journal of Personality and Social Psychology*, *70*, 552-566.

Miller, J. G. (1997). A cultural-psychology perspective on intelligence. In R. J. Sternberg & E. Grigorenko (Eds.), *Intelligence, heredity and environment* (pp. 269-302). New York: Cambridge University Press.

Resing, W.C.M., Bleichrodt, N., & Drenth, P.J.D. (1986). Het gebruik van de RAKIT bij allochtoon etnische groepen [Using the RAKIT for ethnic minorities]. *Nederlands Tijdschrift voor de Psychologie*, *41*, 179-188.

Sherman, J. A. (1978). *Sex-related cognitive differences*. Springfield, IL: Charles C Thomas.

Sijtsma, K., & Resing, W.C.M. (1991). Scalability of an intelligence test for different ethnic groups. In N. Bleichrodt & P.J.D. Drenth (Eds.), *Contemporary issues in cross-cultural psychology* (pp. 361-373). Lisse: Swets & Zeitlinger.

Spearman, C. (1927). *The abilities of man*. New York: Macmillan.

Sternberg, R. J. (1982). Reasoning, problem solving and intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 225-307). New York: Cambridge University Press.

Sternberg, R. J. (1988). *Advances in the psychology of human intelligence* (Vol. 4). Hillsdale, NJ: Lawrence Erlbaum.

Ten Berge, J.M.F. (1977). *Optimizing factorial invariance*. Groningen: VRB Drukkerijen.

Thurstone, L. L. (1938). Primary mental abilities. *Psychometric Monographs*, *1*.

Thurstone, L. L. (1947). *Multiple factor analysis: A development and expansion of* The Vectors of Mind. Chicago: University of Chicago Press.

Verhulst, F. C., Althaus, M., Versluis-den Bieman, H.J.M. (1992). Damaging backgrounds: Later adjustment of international adoptees. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*, 518-524.

Vernon, P. E. (1960). *The structure of human abilities* (Rev. ed.). New York: Methuen.

Nico Bleichrodt is professor of work and personnel psychology at the Free University, Amsterdam. His main interests include cross-cultural research, the development and implementation of tests in developing countries, the problems of ethnic minorities, and the consequences of micronutrient deficiencies for mental and motor development.

René A. C. Hoksbergen studied social psychology and social pedagogics at the University of Amsterdam. His Ph.D. thesis was titled, Profile of the Evening Secondary School Student *(1972). At present, he is professor in the field of the adoption of Dutch and foreign foster children at Utrecht University and managing director of the Adoption Centre of Utrecht University's social faculty. He appears in print a great deal, largely in connection with problems in the field of adoption. He gives lectures at home and abroad and advises adoptive parents who are experiencing family problems.*

Usha Khire is a professor at and head of the Institute of Psychology in Pune, India. Her research focuses on leadership development, reading improvement, creativity training, problem solving, decision making, and Guilford's Structure of Intellect Model. She is also secretary of Jnana Prabodhini, Pune, a nonpolitical organization aimed at the all-around development of the physical, mental, intellectual, and spiritual qualities of students in general and the intellectually gifted in particular.