INTRODUCTION

Psychological tests as a psycho-technology are urgently needed for educational, occupational, counselling, and clinical assessment in the Sudan (Khaleefa, 1987). However, there were no standardized tests for the Sudanese local environment except an attempt carried out by Scott (1950) and Badri (1966). Psychological tests are needed for the maximum utilization of human resources in the newly developing nations in Africa and elsewhere. The rapidly expanding educational facilities in these countries require testing for admission purposes as well as for individual counselling. With increasing industrialization, there is a demand for tests to help with job selection and placement of personnel, particularly in mechanical, clerical, and professional fields (Anastasi, 1961).

Badri (1979) points out that psychometry is an area in which Western psychology has offered one of its greatest contributions to science. This is particularly true of the more objective measurements like intelligence tests, personality inventories, and vocational guidance tests. But for such Western psychological tests to be of any help in Muslim countries, a good deal of adaptation and standardization must be carried out. The great differences between European super-industrialized countries and Muslim developing societies can invalidate the results of unadapted tests.

A number of investigators from various sociocultural systems have agreed upon the limitations of Western or modern psychology when applied across indigenous or traditional cultures (Berry, 1974; Azuma, 1984; Ching, 1984; Lagmay, 1984; Sinha, 1980; Hoshino and Umemoto, 1986; Moghaddam and Taylor, 1985; Moghaddam, 1993; Georgas, 1993; Enriquez, 1993). According to Ombredane (1958), the testing of Africans assessed mainly their degree of acculturation and familiarity with European ways of thinking. Africans were still governed by their custom-
ary ways; and Western tests ran the risk of being useless because the operations on which they made demands were not relevant to the system of motivation and constraints inherent in the behavioural environment of most Africans. Faverge and Falmagne (1962) made another suggestion—that the objective of educational and occupational testing should be to assess the adaptation of African students and workers to the educational or industrial environments imported from the Western world.

Apart from the Western vs. indigenous or traditional vs. modern cultural differences that can invalidate the results of the unadapted tests, there are also regional cultural variations that can affect the performance of individuals. Scott (1950) encountered some problems when administering the Egyptian 'Gabbani-Ballard Test' to Sudanese participants. This test no doubt suited Egyptian children, but when it was used in the Sudan, even with prior adaptation, two-thirds of its items proved weak or wholly undiscriminating: some because they had become too easy, and everybody got them right; some because they had become too difficult, and everybody got them wrong; and many because of misunderstandings due to associations of ideas which vary in different geographical and social environments.

Sudan is seen as an Afro-Arab Islamic country (Khalid, 1990). Egypt and Sudan both belong to the Middle Eastern culture; they have strong historical, economic, and educational relationships as well as having the same language and religion. Regardless of these cultural similarities, however, Khaleefa (1994) found that psychological tests developed in Egypt have many problems when applied to Sudanese individuals.

The north of the Sudan is considered Arab and Muslim, and the south negroid and pagan with a Christian leadership (Deng, 1973). According to Osman (1970), the great majority of the people in the north profess Islam and speak the Arabic language; many of the southern Sudanese, especially those who have been educated at missionary schools, are Christians, while the majority adhere to their tribal indigenous religions. A multiplicity of African Nilotic languages is spoken in the south. In the southern towns English and Arabic are spoken. Due to the cultural diversity between the north and the south, the aim of this work is not to study Sudan as a whole; rather, it will be restricted to the northern Sudan, particularly to the greater majority who profess Islam and speak the Arabic language.

Apart from general and regional cultural variations, there are specific cultural differences such as the extent of literacy within the one country that can influence the performance of individuals. Badri (1966) carried out an empirical study by using the Draw Man Test with a group of 1435 pre-school and school children in the Sudan. He found that urban children scored higher means than rural children. Badri explained that
one of the difficulties that handicapped the full administration of the
test in the local environment was the inability of rural children to use
pen and pencil.

Generally in Africa Western psychological tests have been found to
favour the literate people in urban and modern sectors of society, which
were created by westernization and colonization. A number of investig-
ators articulated the limitations of psychology when administered to
individuals in the traditional African culture (Ombredane, 1958; McFie,
1961; Faverge and Falmagne, 1962; Taylor, 1962; Lule, 1967; Wober,
1974). Durojaiye (1984) found that in Africa the urban candidates often
do better than rural candidates. These wide differences within the one
culture led him to suggest the need to have different cut-off points for
urban candidates. He argues that using Western tests is unsuitable for
Africans, especially those least westernized. However, their continued
usage is justified on the grounds that most African nations are develop-
ing towards Western technology and culture and it is important to
assess individuals in terms of the abilities that have foundations for this
development. Secondly, the pattern of education is essentially Western,
and it is likely that success or failure in education will be related to the
abilities that have been enhanced by Western rather than African culture.

Psychological tests are found to favour males in the indigenous socio-
cultural system. Badri (1966) found that male children showed higher
means than female children in intelligence. It seems that the males vs.
females differences related mainly to the different ways of socialization
and education provided by the society for each sex. According to
Sanderson (1975), educational statistics of government schools for boys
and girls at the close of the colonial period showed greater disparity
than usual, even for underdeveloped countries. Non-government schools
provided education for over four times as many boys as girls.

It seems that it is quite important to ask these questions: What are
the main aspects of cultural bias of the well-known Weschler Adults
Intelligence Scaled-Revised (WAIS-R) (Wechsler, 1955, 1981; Khaleefa,
1987) in the indigenous Sudanese culture? What are the differences
between literate and illiterate individuals in intelligence? What are the
differences between males and females in intelligence? What are the
levels of cultural bias in intelligence testing?

METHOD

Subjects

The test has been administered to a group of diverse volunteers, 30 and
801 as a pilot and main study, respectively. The selection of the sample
INTELLIGENCE TESTING IN THE NORTHERN SUDAN

was based on the census of 1983 in the Sudan. The following were the strict characters of sampling for the main study: geographical distribution, sex, age-group, education. However, other variables were considered such as occupational, tribal, and urban factors.

Stratified sampling was used in dividing the Sudan into six parts which were the regions of the northern Sudan. Each region was represented by a certain number or percentage which reflects the density of population in the different areas. However, the southern part of the Sudan was not included in this study, as explained above.

The number of males was 418 (52.2 per cent), and of females was 383 (47.8 per cent). The standardized sample has been classified into nine age-groups from 16 to 75 years. Sex was considered for each age-group. Because the majority of people in the Sudan are illiterate (approximately 68.7 per cent), the sample has been classified into two broad categories of education: literate and illiterate.

**Testing**

Wechsler Adult Intelligence Scaled-Revised (WAIS-R) consists of eleven sub-tests; six are verbal and five are performance. The verbal tests are: (1) Information Test, (2) Comprehension Test, (3) Arithmetic Test, (4) Similarities Test, (5) Digit-Span Test, and (6) Vocabulary Test. The performance tests are: (7) Digits Symbol Test, (8) Picture Completion Test, (9) Picture Arrangement Test, (10) Block Design Test, and (11) Object Assembly Test. The age range of the WAIS-R is from 16 years 0 months to 74 years 11 months (Wechsler, 1981).

WAIS-R was initially translated from English into Arabic. The initial translation was translated back from Arabic to English. A primary revision of the translation was checked by an expert in Arabic and psychology. Classical Arabic was used in the instructions for the examiners, while colloquial Arabic was used for respondents, because it is the spoken language and the majority of the Sudanese are illiterate and not familiar with classical Arabic. The translated test has been administered in a pilot study for these reasons: to see the suitability of the test to the local Sudanese environment, to detect culturally biased items, to check the adequacy of the Arabic language, to estimate the statistical characteristics for further study, and finally to test the time factor.

The result of the initial study showed that there were some aspects of cultural bias of the WAIS-R: these were language, content of the test, difficulty of reading and writing, and the time factor. However, further revisions were undertaken for the test items to suit the indigenous
Sudanese culture. The adapted form has been shown to a group of external judges to see the adequacy of the test items for Sudanese subjects with respect to their age group, sex, education, and geographical distribution.

The external judges who were consulted to assess and to evaluate the test items consisted of four groups: three lecturers from the University of Khartoum, three secondary school teachers, three psychology students who had previous experience in administering psychological tests, and one lecturer from the College of Fine Art and one artist from the Educational Publishing House. The four groups generally agreed about the adequacy of the new version for Sudanese subjects. They suggested some small changes to the General Comprehension, Information, and Picture Arrangement Tests. These three tests were, however, considered to be the most culturally biased tests. Finally, consensus and validity were achieved.

Procedures

Because of the large size of the main study (801 subjects), a group of thirty psychology and clinical psychology students were well trained to administer the WAIS-R. Five sessions were held in the University of Khartoum for the training procedures. The group were selected for their personal interest in participating in the study, persistence in attending the five sessions of training, and originating from different parts of the country. Further training was held for the examiners outside the class-room, in which every examiner was required to administer the test to five subjects outside the campus. Finally, the selected group showed a good level of understanding of how to administer the test, how to record the responses, and how to score.

Examiners were told to administer the instructions verbatim for each single test. All subjects were told that they were participating in a study to test the level of intelligence. Student subjects were tested in their schools during their lecture time on an individual basis. Older subjects were allowed a brief rest in the middle of the administration. Some subjects completed the test in two separate sessions. Subjects who had no birth certificate simply told the examiners their approximate age. The maximum time spent in the full administration of the test was 150 minutes. Most of the older and illiterate subjects completed the test in more than 120 minutes; however, younger subjects, particularly secondary and university students, completed the test in between 60 and 120 minutes. There was no technical difficulty in administering the scale to illiterate participants except for the Digits Symbol Test which requires writing skills.
Table (1)
Mean and SD for literate and illiterate subjects for each sub-test

<table>
<thead>
<tr>
<th>Test</th>
<th>Mean</th>
<th>SD</th>
<th>Literate</th>
<th>Illiterate</th>
<th>Literate</th>
<th>Illiterate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information</td>
<td>9.35</td>
<td>2.68</td>
<td>6.04</td>
<td>1.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension</td>
<td>9.62</td>
<td>2.48</td>
<td>7.01</td>
<td>2.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digit-Span</td>
<td>9.55</td>
<td>2.68</td>
<td>6.10</td>
<td>1.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arithmetic</td>
<td>9.82</td>
<td>2.20</td>
<td>7.39</td>
<td>1.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Similarities</td>
<td>9.65</td>
<td>2.77</td>
<td>6.32</td>
<td>1.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>9.55</td>
<td>2.67</td>
<td>6.07</td>
<td>1.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture Arrangement</td>
<td>9.66</td>
<td>2.81</td>
<td>6.84</td>
<td>1.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture Completion</td>
<td>9.46</td>
<td>2.82</td>
<td>5.98</td>
<td>2.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Block Design</td>
<td>9.16</td>
<td>2.77</td>
<td>6.01</td>
<td>1.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Object Assembly</td>
<td>9.45</td>
<td>2.59</td>
<td>7.21</td>
<td>2.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digits Symbol</td>
<td>9.05</td>
<td>2.98</td>
<td>4.87</td>
<td>1.37</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table (2)
General mean and SD for literate and illiterate according to the scaled scores for the verbal, performance, and the full scale

<table>
<thead>
<tr>
<th>Test</th>
<th>Mean</th>
<th>SD</th>
<th>Literate</th>
<th>Illiterate</th>
<th>T-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal tests</td>
<td>9.5</td>
<td>2.6</td>
<td>6.5</td>
<td>1.8</td>
<td>19.27***</td>
</tr>
<tr>
<td>Performance</td>
<td>9.4</td>
<td>2.8</td>
<td>6.2</td>
<td>1.7</td>
<td>19.35***</td>
</tr>
<tr>
<td>Full scale</td>
<td>104.3</td>
<td>23.1</td>
<td>69.8</td>
<td>12.9</td>
<td>26.27***</td>
</tr>
</tbody>
</table>

*** P < 0.001

RESULTS

The pilot study revealed that 80 per cent of Sudanese subjects failed to respond adequately to the Picture Arrangement test; 60 per cent of the subjects did not respond suitably to the Information Test; and 50 per cent failed to respond completely to the General Comprehension Test; 36 per cent were not familiar with some items of the Similarities Test; 14 per cent of the Vocabulary Test was found to be unsuitable for the subjects. The inadequate responses elicited from the Sudanese sample to these sub-tests showed clearly the cultural bias of these tests. For the test to be administered fully, it takes between 90 and 150 minutes.

Table 1 shows the means and the standard deviation (SD) of the literate and the illiterate subjects for each of the eleven sub-tests. Table 2
Table (3)
Mean and SD for females and males in each sub-test

<table>
<thead>
<tr>
<th>Test</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Females</td>
<td>Males</td>
</tr>
<tr>
<td>Information</td>
<td>7.80</td>
<td>8.94</td>
</tr>
<tr>
<td>Comprehension</td>
<td>8.39</td>
<td>9.29</td>
</tr>
<tr>
<td>Digit-Span</td>
<td>8.12</td>
<td>8.95</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>8.64</td>
<td>9.56</td>
</tr>
<tr>
<td>Similarities</td>
<td>8.49</td>
<td>8.88</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>8.08</td>
<td>8.97</td>
</tr>
<tr>
<td>Picture Arrangement</td>
<td>8.57</td>
<td>9.09</td>
</tr>
<tr>
<td>Picture Completion</td>
<td>8.13</td>
<td>8.76</td>
</tr>
<tr>
<td>Block Design</td>
<td>7.88</td>
<td>8.58</td>
</tr>
<tr>
<td>Object Assembly</td>
<td>8.60</td>
<td>8.99</td>
</tr>
<tr>
<td>Digits Symbol</td>
<td>7.57</td>
<td>8.09</td>
</tr>
</tbody>
</table>

Table (4)
General mean and SD for females and males according to the verbal, performance, and full scale

<table>
<thead>
<tr>
<th>Test</th>
<th>Mean</th>
<th>SD</th>
<th>T-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Females</td>
<td>Males</td>
<td>Females</td>
</tr>
<tr>
<td>Verbal tests</td>
<td>8.25</td>
<td>9.09</td>
<td>2.68</td>
</tr>
<tr>
<td>Performance tests</td>
<td>8.15</td>
<td>8.70</td>
<td>2.89</td>
</tr>
<tr>
<td>Full scale</td>
<td>90.3</td>
<td>98.1</td>
<td>25.3</td>
</tr>
</tbody>
</table>

*** P<0.001
** P<0.01

shows the T-value for the verbal, performance, and the full scale for both literate and illiterate subjects. The study showed that the average mean for literate subjects on the full scale was 104.3 (SD=23.1), while it was 69.8 (SD=12.9) for illiterate. The differences between literate and illiterate subjects were highly significant. For the verbal tests t = 19.27 (P<0.001); for the performance tests t = 19.35 (P<0.001); and for the full scale t = 26.27 (P<0.001).

Table 3 shows the means and SD for each sub-test according to sex. Generally, males showed higher means in all eleven sub-tests. Table 4 shows that the mean for females in the full scale was 90.3 (SD=25.3) while it was 98.1 (SD=25.39) for males. This difference was significant between males and females in the verbal tests where t = 4.33 (P<0.001),
and in performance tests where $t = 2.66$ ($P < 0.01$), as well as in the full scale where $t = 4.35$ ($P < 0.001$).

**DISCUSSION**

The first group of results of the present study showed clearly that some of the WAIS-R sub-tests were culturally biased when administered in the indigenous Sudanese culture. There are several aspects of bias. One aspect was the language. The original language of the test is English and the test needed to be administered in the indigenous Sudanese culture where Arabic is the spoken language. The test has been translated into Arabic to overcome this difficulty. However, there were some theoretical problems concerning the suitability of the translation. For example, six of the Vocabulary Tests were thought to be unsuitable for Sudanese subjects because of the problem of the equivalent meaning of the word. A concept that is non-existent in a certain language simply cannot be translated into it, a factor which acts as a safeguard against the mechanical use of a given instrument when adapting it for a different culture (Anastasi, 1961).

The second aspect of the cultural bias of the WAIS-R was the unfamiliarity of some of the test items to Sudanese subjects. In the Information Test the Sudanese subjects were not familiar with items like: What are the colours of the American flag? Who was Amelia Earhart? Who was Armstrong? Who wrote *Hamlet*? In the Comprehension Test the Sudanese subjects were not familiar with items like: Why do some people prefer to borrow money from a bank rather than from friends? In the Sudan people would rather borrow from friends than from a bank. Further unfamiliar items included: What should you do if, while in the movies, you are the first person to see smoke and fire? Most of the Sudanese from remote rural areas have no idea what a movie is. In the Similarities Test some Sudanese subjects were not familiar with items related to: Coat-Suit, which is strange dress to some Sudanese, Button-Zipper, and Work-Play. Furthermore, the Sudanese subjects were not familiar with most of the content of the Picture Arrangement Test. It is not an easy task to translate pictures into the indigenous culture of the Sudan. According to Anastasi (1961), the language of performance is the cultural perception, but its words and grammar and syntax are not even completely understood, let alone organized in rational entities. We do not know how to translate a picture into the representative language of a different culture, but we are thoroughly familiar with the techniques and requirements of translating verbal contents.
The third aspect of the bias of the WAIS-R was the time factor. Usually the full administration of the test required a time of between 60 and 90 minutes (Wechsler, 1981). However, according to the Sudanese subjects, the test needed a time ranging between 90 and 150 minutes. This factor might be due to the different concept of time in different cultures or it might relate to the level of education: illiterate subjects usually take more time to complete the test.

We assumed that intelligence tests suffer from cultural bias when used across cultures. However, cultural bias within the one cultural region has emerged. The finding reached by Scott (1950) and Khaleefa (1994) about the inadequacy of Egyptian tests when administered in the Sudanese culture reflected a form of cultural bias within the one cultural region. Furthermore, excluding the southern part of the country from participation in a national study for cultural reasons represented another kind of cultural bias. Thus, the test can only be administered to the great majority in the north of the Sudan who speak Arabic and profess Islam.

The second group of findings of the present study shows that there are significant differences between literate and illiterate individuals. Individuals who received education showed more intelligence than those who did not. There could be several possible reasons for this. One is that intelligence is strongly correlated with education. Evidence for this has been shown by Wechsler (1955). Some of the WAIS-R tests required writing skills like the Digits Symbol Test. The present study showed that the lowest mean of the eleven tests was scored by the illiterate group in the Digits Symbol Test (Table 1). However, the majority of the Sudanese are illiterate (68.7 per cent according to the census of 1973). It must not be forgotten that the test requires some facility with literacy skills. Sudanese support for this finding was provided by the study carried out by Badri (1966), who argued that one of the difficulties that handicapped the full administration of the tests in the local environment was the inability of rural (illiterate) children to use pen and pencil. Also the study confirmed the findings reached by Durojaiye (1984) that urban candidates do better than rural candidates and that the tests are unsuitable for those who are least westernized.

It seems that using modern and Western tools to assess the masses of illiterate people in the indigenous Sudanese culture can have some serious social, ethical, and clinical consequences in categorizing the illiterate subjects as borderline in their abilities because they showed an average intelligence of 69.8 according to the measure of intelligence. However, these tools can be more appropriate within the framework of literate people in urban areas who live a modern life-style. This sector of the society has been shaped by Western or modern forces. In
this case we need two different forms of analysis and understanding, one form for the literate people in urban areas and the other form for the illiterate in rural areas; in other words, one form for the indigenous or traditional sector and another form for the modern or Western sector of the society. Here I agree with Moghaddam and Taylor (1985) who have introduced the concept of 'dual perception' to argue that psychology in the Third World tends to be limited to the modern sector and divorced from the traditional sector of a developing society.

The third major group of results showed that there were significant differences between males and females in the verbal, the performance, and the full scale of the WAIS-R. These results confirmed Badri's (1966) findings that male children in primary schools scored higher means than female. This difference might be related mainly to the differences in the level of socialization and education of each group. In the Sudan males got more education than females. With more education and more social changes this difference will probably be reduced, or may disappear entirely.

To sum up, the findings of the present study showed that intelligence tests indeed suffer from cultural bias towards their milieu of origin. As such, these tests need substantial adaptation in order to enjoy adequate applicability in the indigenous Sudanese culture. The study revealed that there are three kinds of cultural bias of intelligence tests which affect the performance of individuals and groups. First, there is the difference of general or universal cultural level caused by the differences between Western and indigenous Sudanese culture. These differences have been reflected in the language and content of the test and in the time factor. Second, there is a regional cultural effect caused by the differences within the Sudanese culture and between African and Arab culture. These differences required the exclusion of the southern part of the country and showed up the inadequacy of the Egyptian tests. Third, there is the difference of specific cultural level within the one country which caused the literate-illiterate and males-females distinction. It seems that these general, regional, and specific cultural differences and similarities are similar to what Zaccaria (1967) suggested: that each person is in some ways like all other persons, like some other persons, and like no other person.

REFERENCES


Beirut: Al-Fath Publication.


— (1994). 'The Influence of some Sociocultural Factors on Creativity among Sudanese Students'. Research in progress, University of Newcastle upon Tyne, UK.


