

**NEUROPSYCHOLOGICAL ASSESSMENT OF
ZULU-SPEAKING SCHOOL CHILDREN:
SOME NORMATIVE DATA**

GARY VILJOEN

B.A.(Honours)(University of Natal)

**A dissertation submitted in partial fulfilment
of the requirements for the degree of
Master of Arts in Clinical Psychology**

**Department of Psychology
University of Cape Town
1992**

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to the following people:

Ann Levett, my supervisor, for her constructive advice with the collation of this study

Colin Tredoux for his patience and invaluable assistance in the statistical analysis, and for partaking in much needed leisure activities

Gerard Drenan for the valuable hours spent scoring protocols and providing cheerful support

Stuart Anderson for introducing the research area and for initial useful discussions

The headmaster, staff, and pupils of all the schools visited for being so accommodating, and allowing the present study to be conducted in their schools

The Department of Education and Training for allowing this study to be conducted.

This dissertation is the result of my own work except, where otherwise indicated.

ABSTRACT

This study investigates the Bender Gestalt and the Draw-a-Bicycle tests in South Africa by evaluating the applicability of the existing norms (Koppitz, 1975; Kolb & Whishaw, 1985) and by exploring the variables age, sex, educational level, and geographical area as they affect test performances in Zulu-speaking school children. Zulu-speaking subjects (N = 882), ranging in age from 6 - 18 years, completed both tests. Tests were administered in group form and independently scored. The study sample's performances on both tests were compared to the respective foreign norms and some significant discrepancies were apparent in each case, thereby confirming the hypothesis that these foreign norms are inappropriate for use with Zulu-speaking children. Initial normative data are presented for both tests for Zulu-speaking children. Results indicated that age accounted for most of the variance in both tests. For the Bender Gestalt test, sex and geographical area were also significantly correlated to performance but their respective effect sizes were small enough to warrant excluding them as significant variables from the normative data. For the Draw-A-Bicycle test, sex was highly correlated to performance, and was therefore included in the normative data, with geographical area's small effect size warranting exclusion. Performances per age group, on both tests were found to correlate highly, suggesting inter-test reliability. The implications of these findings for further clinical use of the Bender Gestalt and the Draw-a-Bicycle tests, are considered. Limitations of the present study, in conjunction with suggestions for further research, are discussed.

TABLE OF CONTENTS

Introduction	1
1. Review of the Literature.....	2
1.1 Bender Gestalt.....	2
1.2 Draw-a-Bicycle.....	4
2. Aims and Statement of the Problem.....	7
Methodology	10
1. Subjects.....	10
2. Procedure.....	11
Results	12
1. Descriptive and Comparative Statistics.....	12
1.1 Bender Gestalt.....	12
1.1(i). Normative Evaluation.....	13
1.1(ii). Internal Analysis.....	18
1.2 Draw-a-Bicycle.....	22
1.2(i). Normative Evaluation.....	23
1.2(ii). Internal Analysis.....	27
2. Discussion.....	31
2.1. Bender Gestalt.....	31
2.2. Draw-A-Bicycle.....	34
2.3. Comparison of BG and DAB.....	36
Conclusion	37
References	x
Appendix I	i
Appendix II	ii
Appendix III	iii
Appendix IV	iv
Appendix V	v
Appendix VI	vi

INTRODUCTION

In general, a lack of normative data for South African population groups in the area of psychological testing exists. This problem is highly evident in neuropsychological assessment where in private, clinical and hospital settings, the clinician finds him\herself making inferences using normative data based on foreign population groups, or using data from adults for children, or vice versa. In addition to these problems, there exists a distinct lack of standardisation for administration, scoring, and interpretation of certain tests used in the neuropsychological setting, for example, the Draw-a-Bicycle test.

Both, the Bender Gestalt test (hereafter referred to as the BG) and the Draw-a-Bicycle test (hereafter referred to as the DAB), are considered to be simple and economical neuropsychological screening devices in the clinical setting and are regarded by some as sensitive detectors of organicity (Koppitz, 1960 ; Kolb and Wishaw, 1985; Lezak, 1983). The strength of this statement, however, rests on a single proviso - namely, *that these tests only be administered to those for whom they have been standardised*. Research has suggested that variables such as Age, Intelligence and Educational Level are critical in determining test performance, and that existing normative data, therefore, may not necessarily cross-validate to other settings. Thus, there appears to be a strong need to establish appropriate norms for each setting in which these tests are to be used. For these tests to have assured validity and to be used effectively in South Africa they need to be standardised on the specific population in which they are to be used.

The BG and the DAB were specifically chosen as the focus "tools" for this research for various reasons: Both tests are reported to be sensitive instruments in differentiating between brain-damaged and non brain-damaged patients (Kolb and Wishaw, 1985; Lezak, 1983), and therefore have positive diagnostic value in the neuropsychological setting; both tests are cost effective in that they are easily administered and scored for quantitative analysis and in that they require inexpensive test materials, an important consideration in the South African context; both tests have scoring systems which are appropriate for evaluating performance of children and adolescents; and finally, both are measures of visual-motor perceptual ability, which means that inter-test reliability can be assessed.

Zulu-speaking children were chosen for this study because the general paucity of research investigating the use of neuropsychological tests with black children highlighted the need for investigation in this area, and Zulu-speaking children represented an opportunity sample in Natal which was potentially representative of all black children in South Africa.

1. REVIEW OF THE LITERATURE

1.1 Bender Gestalt:

The BG is a test of visual-motor perceptual ability and consists of nine figures (see Appendix i) which are presented one at a time and which the subject is asked to copy on a blank piece of paper. The original designs were used by Wertheimer as far back as 1923 (cited in Koppitz, 1975) to demonstrate the principles of Gestalt Psychology as related to perception. Bender (1938; cited in Koppitz, 1975), adapted these figures and used them as a visual-motor test, pointing out that the perception and reproduction of the figures are determined by biological principles of sensory motor action and vary depending on the growth pattern and maturation level of an individual and his/her pathological state.

One of the advantages of the BG Test lies in the fact that it can be interpreted in several different ways. Apart from the developmental and clinical approaches suggested by Bender, Hutt (1950, 1960; cited in Koppitz, 1975), used the BG as a projective test, interpreting the results in accordance with psychoanalytic theory.

As the BG became more widely used, many psychologists felt a great need for an objective scoring system that was both reliable and valid. As a result, several scoring systems were developed (e.g. Gobetz, 1953; Keller, 1955; Kitay, 1950; Peek and Quast, 1951; Stewart and Cunningham, 1958: cited in Koppitz, 1975). What is apparent, however, is that all these systems were designed for use with adult psychiatric patients or retarded institutionalised children and not for children of "normal" intelligence. Later (post-1960), a growing awareness of the value of the BG for children yielded research and adaption in which this test saw wider application, such as screening for school readiness, diagnosing reading and learning problems, predicting school achievement, determining the need for psychotherapy, and studying retardation

(Koppitz, 1975). Reliability and validity measures of the BG in these varied applications do not appear to be well researched and clearly need to be viewed with caution.

The use of this test as a "tool" for screening for brain damage in adults and in children however was also discovered and research in this area and use of the test in this capacity has increased dramatically (Koppitz, 1960; Lezak, 1983; and Kolb and Whishaw, 1985). According to Lubin (1971: cited in Lezak, 1983) the BG is the most frequently used psychological test in the United States.

The literature reveals that little effort has been devoted to the systematic use of the BG with subjects between the ages of 11 - 17 years. Some of the scoring systems for adults include subjects from 15 years old (Pascal & Suttel, 1951; Hain, 1964: cited in Edwards, 1988), and others for children, such as the Koppitz Developmental Scoring System (Koppitz, 1975), cater for children up to 11 years. There appears to be a lack of normative data for adolescents, and the clinical utility of both the "adult" and "child" scoring systems is deemed limited and unreliable for this age range (Pascal & Suttel, 1951; Koppitz, 1975). Edwards (1988) investigated the relative efficacy of two different scoring systems for use in evaluating BG performance by adolescents. His study showed a high degree of consistency between the Koppitz Developmental Scoring System (Koppitz, 1975) and the Lacks Scoring System for Adults (1984; cited in: Edwards, 1988), which suggests that either scoring system is suitable for evaluating BG performance of adolescents. According to Edwards (ibid), either scoring system can be used reliably with adolescents *as long as normative data are available for the sample being evaluated.*

According to the Koppitz (1975) and Lezak (1983) standardization procedure, the test is administered to the subject on an individual basis. Each card is independently presented and the subject is required to reproduce the figure on a blank A4 sheet of paper. This task is given no specific time limitation although excessive deviations are noted for qualitative analysis. The instructions read as follows:

"I have nine cards here with designs on them for you to copy.
Here is the first one. Now go ahead and make one just like it."

When the subject has finished drawing the figure, the card with the stimulus design is removed and the next card is put in front of him\her and so on. All nine cards are presented in this fashion in orderly sequence. The answers to any questions by the subject should remain nonprescriptive - e.g. "Make it look as much like the picture on the card as you can." The method of scoring employed in this study is in accordance with Koppitz's Developmental Bender Scoring System (1975). All items are scored as either one or zero (i.e. as either present or absent). Since the system is designed for young children with as yet immature fine motor control, minor deviations are ignored, and only clearcut deviations are scored. In case of doubt an item is not scored. All scoring points are added into a composite score upon which the normative data are based (ibid).

1.2 Draw-A-Bicycle:

The DAB is a very economical test of visual memory and visual-motor perception ability, requiring minimal test material and administration time.

The origins of the DAB test are unclear but it appears to have some of its roots in the field of clinical neurology. Paterson and Zangwill did a study in 1944 in which they employed the "free-hand drawing of a bicycle" to assess "construction and visual space perception" in two patients with right hemisphere lesions (Paterson & Zangwill, 1944). Taylor (1959), on the other hand, attributes the concept of using free-hand drawings for assessment of visuographic functioning to Piaget.

Although gaining prominence in the field of neuropsychological testing, the DAB is essentially a new recruit to this field, evidenced by the lack of reported research. An exhaustive literature search has yielded exceptionally little information on this test in general and it appears that no manual for administration, scoring, and interpretation exists. Bearing in mind the positive practical efficacy which this test is afforded by clinicians, it is clear that such data is in demand; especially context-specific information in South Africa.

An unpublished study by Nichols (1980) examined the reliability and validity of the DAB. 141 brain-injured subjects were selected to complete the DAB, the Wechsler Adult Intelligence Scale, and the General Aptitude Test Battery.

Significant correlations between DAB performance and certain subtests of the other tests served as evidence for the validity of the DAB. Internal consistency reliability was also revealed.

According to Kolb and Wishaw (1985), the DAB test has distinct advantages because it allows the subject to relax with a more familiar task and because it is sensitive to many kinds of disorders. The draw-a-bicycle task is used since bicycles are well known to almost everyone and because they are very complex. One difficulty with using drawing tests is that there are seldom any objective criteria for their scoring.

Kolb and Wishaw (1985) devised a scoring protocol for bicycles (see Appendixii) and developed norms for the DAB in a study in 1983 (Wishaw, 1992). 2665 Subject (ages 2 - 50 years) were selected for the study. The majority (2276 subjects) were school children, randomly selected from the Catholic school system in Lethbridge, Canada. Subjects were predominantly white, middle-class with no specific control for brain damage or psychiatric illness. Results of the study are published as a normative table (Kolb & Wishaw, 1985).

According to Lezak (1983), bicycle drawings can serve as a test of mechanical reasoning as well as visuographic functioning. The draw-a-bicycle test tends to bring out the drawing distortion characteristic of lateral damage. Right hemisphere patients tend to reproduce many of the component parts of the machine, sometimes with much elaboration and care, but misplace them in relation to one another (Paterson & Zangwill, 1944; McFie, Piercy & Zangwill, 1950). Left hemisphere patients are more likely to preserve the overall proportions but simplify the drawing (McFie & Zangwill, 1960). Severely impaired patients, regardless of the site of the lesion, perform this task with great difficulty, producing incomplete and simplistic drawings (Lezak, 1983; Kolb & Wishaw, 1985).

The DAB is an economical test to use: It requires a pencil and a blank A4 sheet of paper for the subject, and the instruction "draw a bicycle". The subject is allowed time to finish the drawing. The method of scoring used in this study is in accordance with the 20-point scoring system proposed by Kolb and Wishaw (1985) (See Appendixii).

Interpreting performance on visual-motor perception tests such as the BG and DAB include the following considerations:

The reproduction of designs on visual-motor tests requires appropriate fine motor development, perceptual discrimination ability, plus the ability to integrate perceptual and motor processes. Thus, inadequate performance may be the result of misperception (faulty interpretation of input information), execution difficulties (faulty fine motor response output), or integrative or central processing difficulties (faulty memory storage or retrieval systems) (Lezak, 1983; Whishaw, 1992).

Variables that may lead to poor perceptual-motor functioning and thus affect drawings made on visual-motor tests include (a) physiological limitations, sensory deprivation, muscular weakness, and other physical handicapping conditions; (b) physiological disruptions, such as illness, injury, or fatigue; (c) immediate or temporary environmental stress; (d) poor motivation; (e) affective disruptions; (f) intellectual retardation; (g) social or cultural deprivations; and (h) limited experiences (Koppitz, 1975; Lezak, 1983).

It may be possible to discern whether the difficulty lies in the output (motor or expressive functions) or the input (perceptual or receptive functions) process. Generally, if children struggle to reproduce the designs, the difficulty is likely to be motor or expressive, whereas if they draw them quickly and easily but with errors that are not recognized, the difficulty may be perceptual or receptive and not entirely motor. When children cannot see their errors, the trouble may lie in the input mechanisms. When, on the other hand, they can acknowledge their errors but cannot correct them, the difficulty may be due to faulty output mechanisms (Lezak, 1983; Whishaw, 1992).

Mildly confused, distractable, and structure-dependant patients who consider their drawings complete when it still lacks a necessary element such as pedals, drive chain or a seat, will usually note and repair the omission on questioning. Patients with problems of visual neglect, visual scanning, or more than mild confusion may refer to the missing component but remain satisfied with the incomplete drawing, or may overlook the missing part but add an inconsequential detail or superficial embellishment (Lezak, 1983).

Careful observation is needed to distinguish between perceptual failures, apraxias (disturbances of purposeful, expressive functions), spatial confusion, or attentional or motivational problems (ibid).

2. AIMS AND STATEMENT OF THE PROBLEM

This study was motivated by informal discussion with clinical psychologists and neuropsychologists in Pietermaritzburg and Cape Town who expressed their concern about the lack of local normative data available in the field of psychological testing. This concern was substantiated for the clinician by the paucity of research found in the literature in the area of test development and standardisation in South Africa. Specifically, there appeared to be a scarcity of local norms in two areas: (i) for black South Africans and (ii) for tests of organicity.

Variables such as Age, Educational Level, Socio-economic Background, and Intelligence, are considered to play an important role in test performance (Cole and Means, 1981). When it is considered that educational, social and economic factors are thorny and controversial issues in South Africa, with the educational level and the socio-economic level reached by the majority of individuals not necessarily being determined by scholastic ability or intelligence, it appears that existing test norms cannot automatically be used in the South African setting.

Concerning the question of culture bias, there appears to be an assumption based in many studies that non-verbal stimuli are "culture free" (ibid). Thus, many tests such as the BG and the DAB are used with equal aplomb for subjects of all cultures. Other researchers (Harrower: cited in Peixotto, 1954; Taylor and Thweatt, 1972; Snyder, Holowenzak & Hoffman, 1971; Marmorale and Brown, 1977; Taylor and Partenio, 1984), however, have found the opposite. In Peixotto's (1954) study, in which seven different ethnic groups were administered the BG, the results indicate significant variances at the 5% level of confidence, suggesting the probability that various ethnic groups will produce different protocols.

With respect to the sex-related variable, studies have indicated that no significant sex differences exist when the mean raw scores achieved on the BG were compared. Koppitz (1960), noted that girls appeared to mature somewhat

earlier than boys in visual-motor perception but found no significant differences between the sexes at any age level. Studies by Keogh and Smith (1963: cited in Fiedler & Schmidt, 1969), and Fiedler and Schmidt (1969), have yielded similar results. In a study on the BG, Dierks and Cushna (1969: cited in Fiedler & Schmidt, 1969) however, described significant differences between sexes in a clinical population, with boys being superior to girls in performance, and recommended that similar studies be carried out with non-clinic populations.

The normative study on the DAB by Kolb and Whishaw (Whishaw, 1990) produced results which clearly indicated sex differences, with males generally performing better than females. Their final normative data therefore included a breakdown by age and sex.

Based upon the few studies reported (Carlson, 1966; Mosely, 1969; cited in Willis and Pishkin, 1974; Zuelzer and Stedman, 1976), it appears that the BG performance may be influenced by socio-economic factors. Comparative studies such as the one by Willis and Pishkin (1974), have claimed to confirm that perceptual and linguistic differences do exist between different socio-economic classes. Their explanations offered for the disadvantaged child's lack of normal development of cognitive and/or perceptual processes center around environmental deficiencies, and a general lack of common experiences with middle-class children. Although no pertinent studies exist for the DAB, one can tentatively assume that these findings are equally valid for the DAB.

Concerning Intelligence and Educational Levels, research findings have suggested that significant correlations exist between these variables and performance on the BG. In her normative study, Koppitz (1960), found strong correlations between educational level and performance and proceeded to establish norms for this variable, which have been included in her Scoring Manual for the Developmental Bender Scoring System (Koppitz, 1975). Although not as conclusive, Johnson and Lottman (1971), and Koppitz (1960), have both indicated that intelligence is a correlative variable. They propose that the BG test may also prove useful as a non-verbal short intelligence test for young children, considering their findings which indicate a close relationship between the BG and both the WISC Performance Scale IQ and the Arithmetic Subtest Score. No studies were found investigating the relationship between intelligence and DAB test performance, but Kolb's and

Whishaw's study (Whishaw, 1992) showed a very high correlation between age and educational level.

Finally, it is apparent that the age variable has received the most attention in research for, both, the BG and the DAB. It appears to have the clearest and most significant correlation to test performance on both tests. In the majority of normative studies done, the researchers have used age as the normative variable (Pascal and Suttel, 1951; Koppitz, 1960; Hutt and Briskin, 1960; Lezak, 1983; and Kolb and Whishaw, 1985).

The aim of this study, then, is twofold: Firstly, it sets out to provide comparative normative data on the BG and the DAB, by establishing local norms, which can then be compared to existing foreign norms. It is hoped that this will facilitate the identification of any discrepancies, and, thus, the evaluation of applicability and usefulness of the local norms.

Secondly, it sets out to explore some of the variables which research has identified as playing a crucial role in test performance - for example, age, sex, educational level and geographical area. Here, a comprehensive analysis is not attempted, but simply the highlighting of some areas requiring more detailed and ambitious research.

In sum, it is hoped that this will serve to provide initial normative data (for the BG and DAB) for clinical use in the South African context.

The independent variables explored in this study, then, are Age (6 -18 yrs), sex, Educational Level and Geographical Area (of the school), with the dependent variable being Test Performance.

Although socio-economic class, as a variable, was not controlled for, certain possible assumptions can be made concerning the sample group (dealt with in the discussion). Further, it should be noted that despite the obvious importance of intelligence as an independent variable (borne out by research), it was beyond the scope of this study to investigate the effects of intelligence, as measured by intelligence tests, on test performance. Clearly, this remains an important area for future research.

The researcher hypothesised that:

- (a) there would be a significant discrepancy between the performance, per age, of the study sample's subjects and the performance of the respective normative samples' subjects for the BG and the DAB tests;
- (b) Age would have a statistically significant effect on performance for each test;
- (c) Educational Level would have a statistically significant effect on performance for each test;
- (d) Sex and Geographical Area would not have any significant effects on performance for either test; and
- (e) there would be a correlation between the means of the age groups for BG performance and DAB performance because both are measures of visual-motor perceptual ability.

METHODOLOGY

1. SUBJECTS

The participants in the study were 905 Zulu-speaking school children between the ages of 6 and 18 years. As a result of inadequate demographic information, 28 subjects' results were rejected and thus 882 subjects, 434 males and 448 females, were selected for further investigation. The subjects were drawn from two geographical areas in Natal - namely; the "rural" geographical area of Lidgeton West (one pre-primary, one primary, and one middle school), which comprised 438 pupils; and the "urban" geographical area of Edendale (one pre-primary, one preparatory, and one high school), which comprised 444 pupils. It is noted that this was an opportunity sample, with no control for variables such as learning disorders, school failure, head injury, epilepsy, or severe emotional or behavioural problems.

2. PROCEDURE

The Department of Education and Training (DET) was contacted for permission to do research in schools which fall under this department's control. This was granted and contact was made with the Regional Branch of the DET (Pietermaritzburg) to finalise the selection of schools for research and to obtain the official referrals. Finally the various principals were approached for permission to collect data in their schools and to arrange dates and times to meet the relevant staff members and to execute the research. The principals, teachers, and (indirectly) the subjects, were asked if they would participate in a project directed at gaining normative data and assessing the applicability of two psychological tests used by South African clinicians.

Once all parties had agreed to participate, administration preparations began. The problem of procuring translators for the administration was overcome by acquiring the help of the available class teachers, who were given a standard briefing prior to each administration session. This proved invaluable in that a comfortable rapport existed between teachers and pupils, which seemed to minimise test anxiety among subjects. The single most important shortcoming in the study revolves around the question of age. Enquiry after exact dates of birth proved problematic, as school records were either incomplete or non-existent, and as subjects' knowledge of own dates of birth were unreliable. Thus, ages are recorded simply in years. It is hoped that this shortcoming has been accommodated, to an extent, by the large sample size.

It was decided that a group administration procedure, in accordance with a study by Keogh and Smith (1971) would be adopted for both the BG and the DAB (see Appendix iii for summary). This meant that the BG and the DAB were administered to a class at a time. Subjects were required to complete a basic printed demographics questionnaire to establish age, grade and sex (teachers assisted with those who had difficulties).

For the BG, each subject was then handed a blank A4 sheet of paper and a pencil with an attached eraser. Instructions were read in English and in Zulu. The Bender designs, reproduced to scale on A4 cards, were then presented to the class one at a time in sequence (clear visibility of the designs for each subject was ensured before commencing administration with

every class). Each design was presented until all subjects had completed their drawings. All administration variables and queries by subjects were dealt with according to the recommended Developmental Bender Scoring System (Koppitz, 1975).

For the DAB, each subject was handed a blank A4 sheet of paper and a pencil with eraser and instructions were read in English and in Zulu. All subjects were allowed time to finish their drawings. Administration was in accordance with Kolb's and Whishaw's administration procedure during their normative study (Whishaw, 1992) and in accordance with Lezak's (1983) recommended method.

Data were collected over two separate periods of seven days each for the rural area and urban area respectively. The method of scoring employed was in accordance with Koppitz's Developmental Bender Scoring System for the BG (Koppitz, 1975), and in accordance with Kolb's and Whishaw's scoring method for the DAB (Kolb and Whishaw, 1985) (see Appendix ii).

RESULTS

1. DESCRIPTIVE AND COMPARATIVE STATISTICS

1.1 Bender Gestalt:

Descriptive and comparative statistics for the BG is divided into two sections - namely, normative evaluation and internal analysis of the results. In **normative evaluation** the central question is:

- (1) "Are Koppitz's norms appropriate for use with Zulu-speaking children?"

This question is answered by means of a process of statistical comparisons between the results of Koppitz's original normative sample and the current study sample.

Internal analysis deals with the analysis of subject attributes and the relationship between these and the BG. Two questions are asked:

(1) "Can these norms be reported without respect to sex or geographical area?"; and

(2) "How reliable are the study sample's norms?".

These questions are answered by performing a 3-way ANOVA and assessing inter-rater reliability. The independent variables are **Age, Sex, Geographical Area,** and **Educational Level** and the dependent variable is **Test Performance** on the BG.

1.1(i) Normative Evaluation:

Performances by the Koppitz's normative sample were compared to the performances of (a) the "rural geographical area" sample, (b) the "urban geographical area" sample and (c) the "total study" sample. Discrepancies between the means of each age group (between Koppitz's sample and the respective study group) were analysed using the Student's t-Test to determine whether these differences were statistically significant.

For each, a test for the equality of the slope coefficients of the regression equations relating age to performance, between the Koppitz's sample group and each of the relative study sample groups, was performed in order to determine if there were any differences in rate of improvement across the age range - ie, did the Zulu-speaking subjects improve their performance with age at a faster or slower rate than Koppitz's subjects, and was this difference statistically significant?

Koppitz's normative data for the Developmental Bender Scoring System for children (Koppitz, 1975) only caters for ages 5 - 10 years. According to Koppitz:

"(Results) show a marked drop in scores between the age of 5 and the age of 7. Thereafter the decrease is more gradual and seems to level off at age 9 when sufficient maturity in visual motor perception has been acquired by most Ss to draw the Bender correctly without any distortions." (Koppitz, 1960).

Based on this explanation of the performance trend, all ages beyond 10 years in the Koppitz's sample have been assigned equivalent means, standard deviations and sample sizes in the present evaluation, in order to facilitate comparisons with the study sample's age groups.

(a) Koppitz's normative sample compared to the "rural geographical area" sample, per age group:

The comparison between the means of these two samples can be seen in graphic form in Figure 1. Results of the Student's t-Tests per age group can be seen in Table 1 (Appendix iv). These results indicate that, on average, Zulu-speaking children at the 8, 9, 10, 11 and 12 year old level perform significantly more poorly than their American counterparts.

Linear Regression Equations by the Ordinary Least Squares method was determined for (1) the Koppitz's normative sample and (2) the "rural geographical area" sample for ages 6 - 10 years. Equations were constructed by regressing age levels against mean BG scores¹. The slope coefficient of the equations represents the rate of change in BG score as a function of age - i.e. the steeper the slope, the faster the improvement of performance with age (for the BG the slope has a negative gradient because the lower the score the better the performance).

A test for equality of the slope coefficients was then performed and a statistically significant difference was found (Table 2, Appendix iv). It was found that Koppitz's sample's slope was steeper than the "rural geographical area" sample's slope, suggesting that the rate of improvement of performance with age was quicker for the American sample (average, 1.5 units per year) than for the Zulu-speaking, rural sample (average, 1 unit per year). The slope is negative because the BG is scored for errors, which means that the lower the score the better the performance (Figure 2).

¹Important to note is that the mean scores and not the raw scores were correlated and regressed against age levels throughout the calculations in this study. The result of this is that both regressions and correlations tend to be more significant than would be the case using raw scores.

Figure 1

BENDER GESTALT SCORES BY AGE AND G.AREA

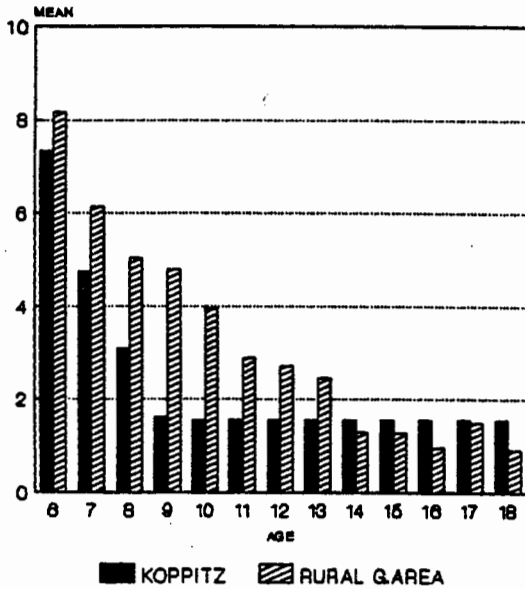
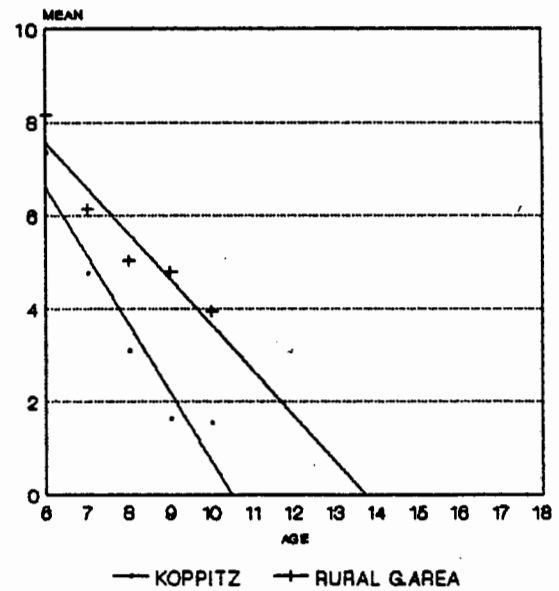


Figure 2

BENDER GESTALT SCORES BY AGE AND G.AREA



(b) Koppitz's normative sample compared to the "Urban Geographical Area" sample, per age group:

The comparison between the means of these two samples can be seen in graphic form in Figure 3. Results of the Student's t-Tests per age group can be seen in Table 3 (Appendix iv). These results indicate that, on average, urban, Zulu-speaking children outperform their American counterparts at age 6, but perform equivalently at all other ages.

Linear Regression Equations by the Ordinary Least Squares method was determined for the Koppitz's normative sample and the "urban geographical area" sample for ages 6 - 10 years. A test for equality of the slope coefficients was then performed and a statistically significant difference was found (see Table 4, Appendix iv). It was found that Koppitz's sample's slope was steeper than the "urban geographical area" sample's slope, suggesting that the rate of improvement of performance with age was quicker for the American sample (average, 1.5 units per year) than for the Zulu-speaking sample (average, 0.6 units per year) (see Figure 4).

Figure 3

BENDER GESTALT SCORES BY AGE AND G.AREA

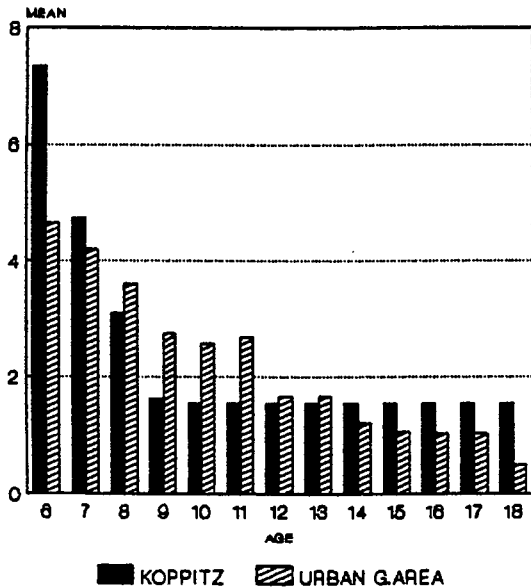
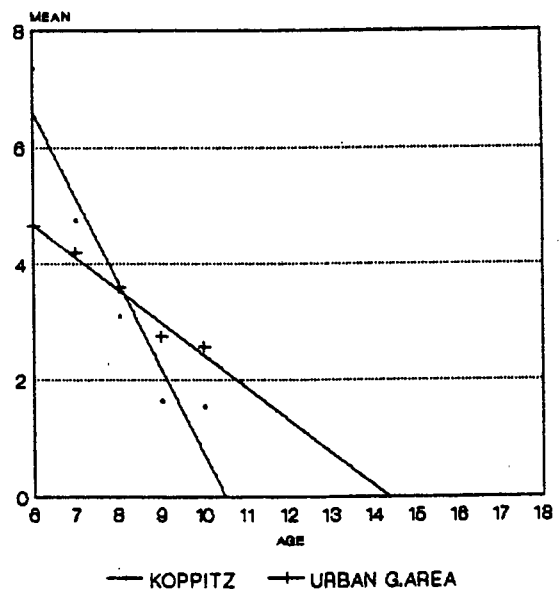


Figure 4

BENDER GESTALT SCORES BY AGE AND G.AREA



(c) Koppitz's normative sample compared to the Total study sample, per age group:

The comparison between the means of these two samples can be seen in graphic form in Figure 5. Results of the Student's t-Tests per age group can be seen in Table 5 (Appendix iv). These results indicate that, on average, a combination of rural and urban, Zulu-speaking children perform significantly more poorly at ages 8, 9, 10 and 11 than their American counterparts. It appears that this overall result is strongly influenced by the rural sample's performance, as opposed to the urban sample's performance which is similar to that of the Koppitz's normative sample. The significant discrepancy at age 18 needs to be viewed with caution because of the small cell size ($n=16$) in the study sample and the fact that these results have been compared to an "average" mean for the Koppitz's age group.

Linear Regression Equations by the Ordinary Least Squares method was determined for the Koppitz's normative sample and the "total study" sample for ages 6 - 10 years. A test for equality of the slope coefficients was then

performed and a statistically significant difference was found (see Table 6, Appendix iv). It was found that Koppitz's sample's slope was steeper than the "total study" sample's slope, suggesting that the rate of improvement of performance with age was quicker for the American sample (on average, 1.5 units per year) than for the Zulu-speaking sample (on average, 0.5 units per year) (see Figure 6).

Figure 5

BENDER GESTALT MEAN SCORES BY AGE

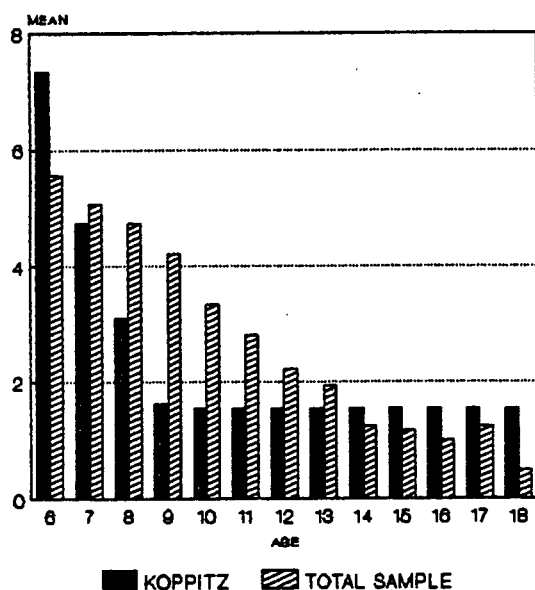
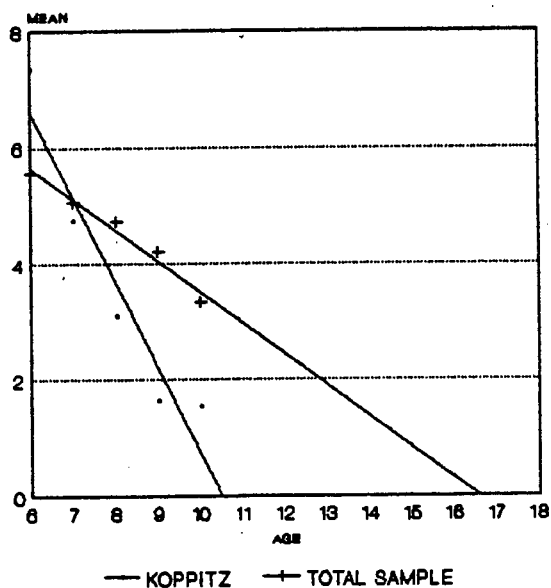


Figure 6

BENDER GESTALT MEAN SCORES BY AGE



A comparison of the mean BG scores per age group for the Koppitz's sample and the Total Study sample, using the Student's t-Test, clearly indicated that performances are not equivalent. In addition, testing the equality of the slope coefficients which represent the rate of change in BG score as a function of age, clearly showed that these two samples do not improve performance with age at an equivalent rate. On average, the American children improve their performance by 1.5 units per year between the ages of 6 - 10 years, whereas their Zulu-speaking counterparts do so at 0.5 units per year. Also, the American children appear to reach "maturity" (the point at which mean scores for the age groups level out) for this task at approximately 9 years of age whereas the Zulu-speaking children do so at approximately 13 years of age.

1.1(ii) Internal Analysis:

A 3-way Analysis of Variance was performed in order to elicit effects of the variables Age, Sex and Geographical Area, to determine the magnitude of each, and to elicit the complexities of the interactions (Table 7). Because of the high correlation between Age and Educational Level ($r = 0.873$; $p < 0.001$) further analysed of Educational Level was deemed superfluous.

Table 7
BG: 3-way ANOVA Summary Table for Independent Variables
Age, Sex, Geographical Area, on Dependent Variable BG Score

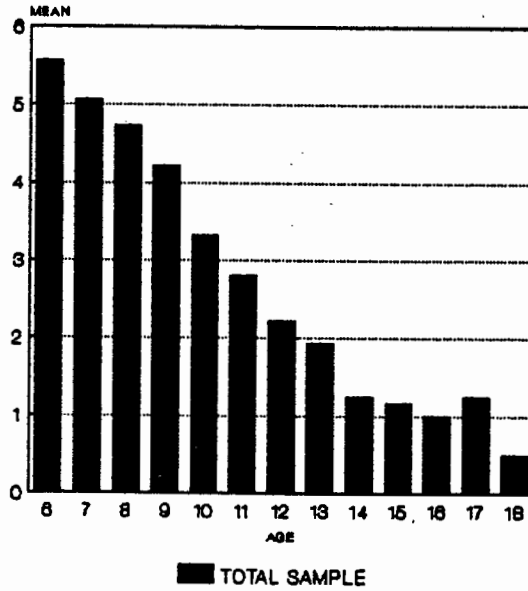
Source	df	Seq SS	Adj SS	Adj MS	F	P	eta ²
Age	11	1543.82	1241.42	112.86	29.02	0.0004	0.306
Sex	1	7.16	21.68	21.68	5.57	0.018	0.001
G. Area	1	107.43	169.88	169.88	43.68	0.0004	0.021
Age*Sex	11	51.66	50.85	4.62	1.19	0.291	
Age*G. A	11	119.88	114.93	10.45	2.69	0.002	0.024
Sex*G. A	1	2.48	6.00	6.00	1.54	0.215	
Age*Sex*G. A	11	32.66	32.66	2.97	0.76	0.677	
Error	818	3181.43	3181.43	3.89			
Total	865	5046.51					

It is noted that Age, Sex and Geographical Area are significant main effects. In order to determine the amount of variance due to the respective effects, eta² was calculated (eta² = the sequential sum of squares of the variable divided by the total sum of squares of all variables). Each main effect will be discussed independently:

Age is a main effect responsible for significant variance across the Study Sample, shown by the large eta² value. An eta² value of 0.306 means that the variable age explains approximately 31% of the variation in the performance within the sample. The regression equation (BG mean scores regressed against age groups) had a slope coefficient of -0.176 , which means that the BG score for Zulu-speaking children between the ages of 6 - 18 years decreases (equals improved performance) at a relatively constant rate of 0.176 units per year. This relationship is graphically represented in Figure 7.

Figure 7

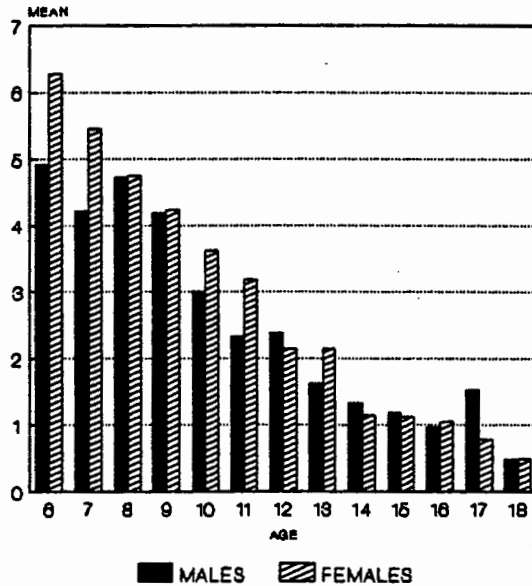
BENDER GESTALT MEAN SCORES BY AGE



Although Sex is reflected as a main effect by the ANOVA, its relative effect size is small, as shown by the η^2 value of 0.001. This means that the variable sex explains approximately 0.1% of the performance variation within the sample. Comparison of the means of males and females, using a Student's t-Test, delivered a $t(\text{obs}) = -2.222$ with $df = 893$. This was not significant at $p < 0.01$. This comparison is graphically represented in Figure 8. A test for equality of the slope coefficients (males against females) was performed and no statistically significant difference was found (see Table 8, Appendix iv).

Figure 8

BENDER GESTALT SCORES BY AGE AND SEX



TOTAL STUDY SAMPLE

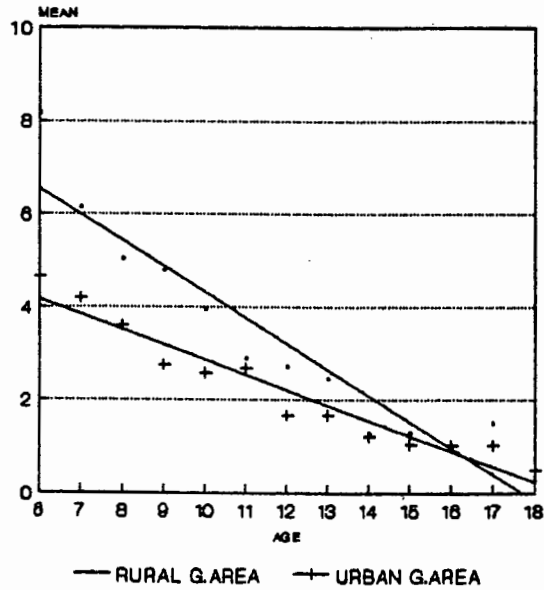
Geographical Area shows up in the ANOVA as a significant main effect. An η^2 value of 0.021, however, indicates that geographical area only accounts for approximately 2% of the overall variance.

Once again a linear regression was performed on the "Rural Geographical Area" sample and the "Urban Geographical Area" sample for ages 6 - 18 years. A statistically significant difference was found when the equality of the slope coefficients were tested (see Table 9, Appendix iv).

The ANOVA shows that Age and Geographical Area interact and account for approximately 2% of the variance. This, plus the graphic depiction of the interaction, provides an interpretation of the nature of the relationship: At the early ages the rural children are quite clearly weaker than the urban children at the BG task. The difference however decreases systematically with age until it disappears at around 15 years, indicating that visual-motor skills are equally developed for rural and urban children by the time they reach 15 - 16 years of age (Figure 9).

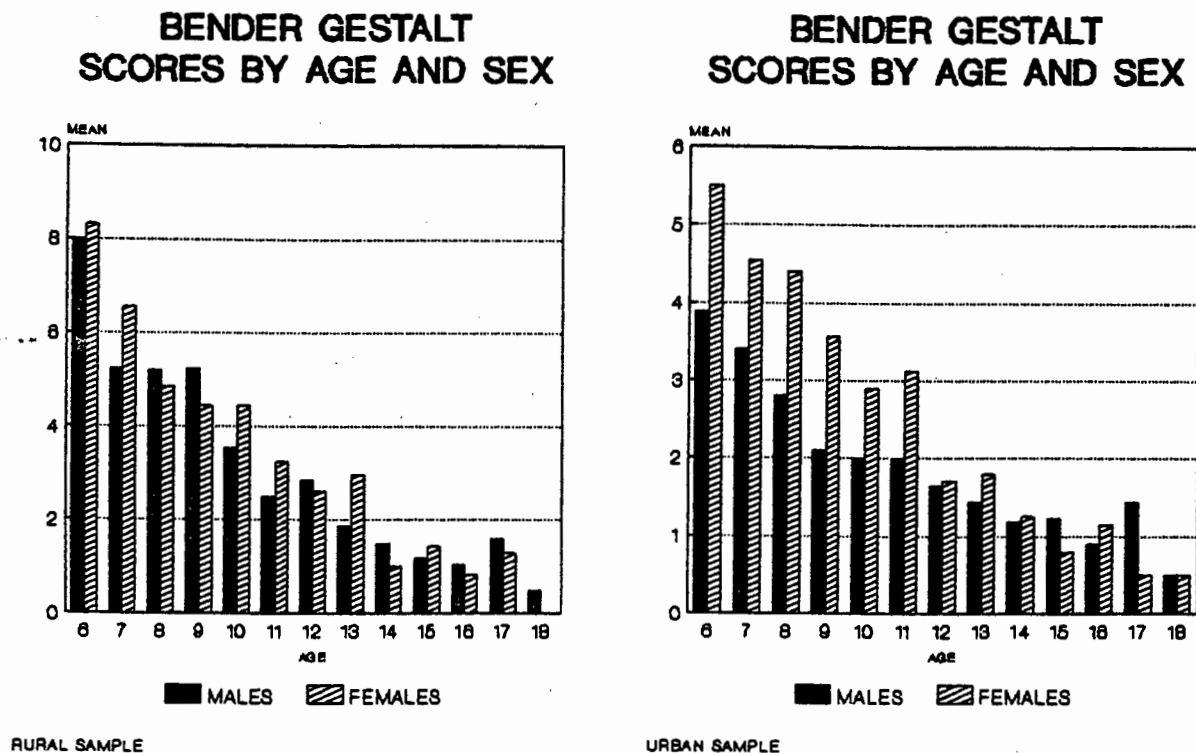
Figure 9

BENDER GESTALT SCORES BY AGE AND G.AREA



The Age x Sex x Geographical Area interaction did not show up as significant on the ANOVA (Table 7). Noteworthy, however, is that the male-female performance discrepancy appeared (graphically) more significant in the "Urban Geographical Area" sample than the "Rural Geographical Area" sample (see Figure 10). It is however clear that the current cell sizes are too small to allow for reliable analysis of this phenomenon and so this was not pursued - for example, in the rural area there were 4 males and 0 females in the 18 year old age group, and 3 males and 3 females in the 6 year old age group).

Figure 10



Based on the above analysis it is clear that age is the predominant variable in terms of its magnitude of effect. Both sex and geographical area are small enough in terms of their effect size to warrant collapsing the norms across them and reporting age-related norms only. This is in accordance with Koppitz's findings and manner of reporting her norms.

Thirty-two randomly selected BG protocols (minimum of 2 per age group, per geographical area) were independently scored by a clinician using the Koppitz's standardised scoring method. A correlation was performed to establish inter-rater reliability and yielded $r = 0.974$ ($p < 0.001$).

1.2 Draw-A-Bicycle:

Descriptive and comparative statistics for the DAB is divided into two sections - namely, normative evaluation and internal analysis of the results. In **normative evaluation** the central question is:

"Are Kolb and Whishaw's norms appropriate for use with Zulu-speaking children?".

This question is answered by means of a process of statistical comparisons between the results of Kolb and Whishaw's original normative sample and the current study sample. In **internal analysis** of the results two questions are asked:

"Can these norms be reported without respect to sex or geographical area?",
and

"How reliable are the study sample's norms?".

These questions are answered by performing a 3-way ANOVA and assessing inter-rater reliability. The independent variables are **Age, Sex, Geographical Area,** and **Educational Level** and the dependent variable is **Test Performance** on the DAB.

1.2(i) Normative Evaluation:

Performances by Kolb's and Whishaw's normative sample were compared to the performances of (a) the "rural geographical area" sample, (b) the "urban geographical area" sample and (c) the "total study" sample. Discrepancies between the means of each age group of Kolb's and Whishaw's sample and the respective study sample were analysed using the Student's t-Test to determine whether these differences were statistically significant.

For each, a test for the equality of the slope coefficients of the regression equations relating age to performance, between Kolb's and Whishaw's sample group and each of the relative study sample groups, was performed in order to determine if there were any differences in rate of improvement across the age range -ie, did the Zulu-speaking subjects improve their performance with age at a faster or slower rate than Kolb's and Whishaw's subjects, and was this difference statistically significant?

(a) Kolb's and Whishaw's normative sample compared to the "Rural Geographical Area" sample, per age group:

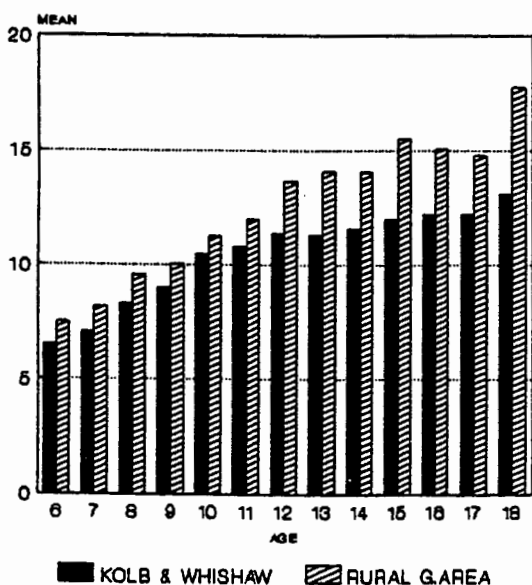
The comparison between the means per age-group of these two samples can be seen in graphic form in Figure 11. Results of the Student's t-Tests per age group can be seen in Table 10 (Appendix iv). These results indicate that, on average, Zulu-speaking children perform significantly better than their American counterparts at the 12, 13, 14, 15, 16, and 17 year old level.

Linear Regression Equations by the Ordinary Least Squares method was determined for Kolb's and Whishaw's normative sample and the "rural geographical area" sample for ages 6 - 18 years. Equations were constructed by regressing age levels against mean DAB scores. The slope coefficient of the equations represent the rate of change in DAB score as a function of age - i.e. the steeper the slope, the faster the improvement of performance with age (for the DAB the slope has a positive gradient because the higher the score the better the performance). A test for equality of the slope coefficients was then performed and a statistically significant difference was found (see Table 11, Appendix iv). It was found that the "rural geographical area" sample's slope was steeper than Kolb's and Whishaw's sample's slope, suggesting that the rate of improvement of performance with age is quicker for the rural, Zulu-speaking sample than for the Canadian sample. At age 6 years their performance is similar but after that the Zulu-speaking children exhibit an average increase of 0.77 units per year compared to the Canadian children's average of 0.51 units per year (see Figure 12).

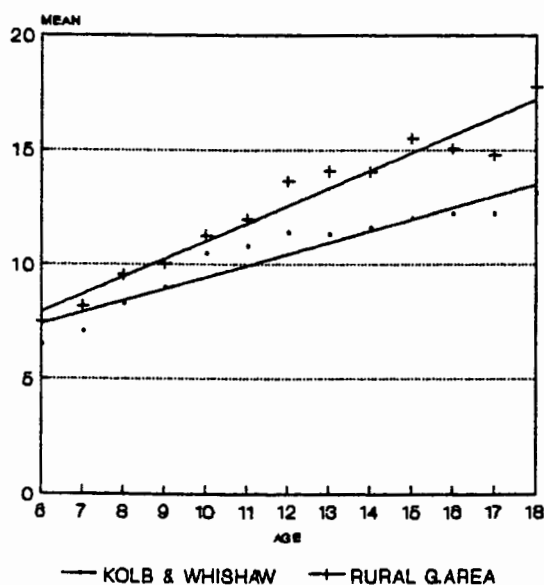
Figure 11

Figure 12

DRAW-A-BICYCLE SCORES BY AGE AND G.AREA



DRAW-A-BICYCLE SCORES BY AGE AND G.AREA



(b) Kolb's and Whishaw's normative sample compared to the "Urban Geographical Area" sample, per age group:

The comparison between the means, per age, of these two samples can be seen in graphic form in Figure 13. Results of the Student's t-Tests per age group can be seen in Table 12 (Appendix iv). These results indicate that, on average, urban, Zulu-speaking children outperform their Canadian counterparts at ages 14 and 15, but perform equivalently at all other age levels.

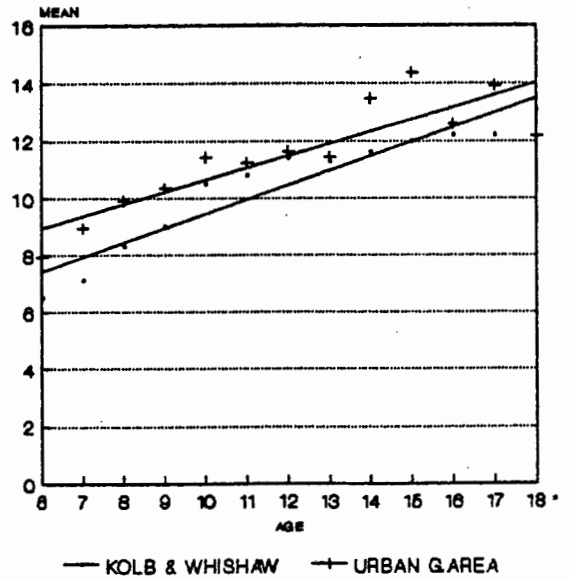
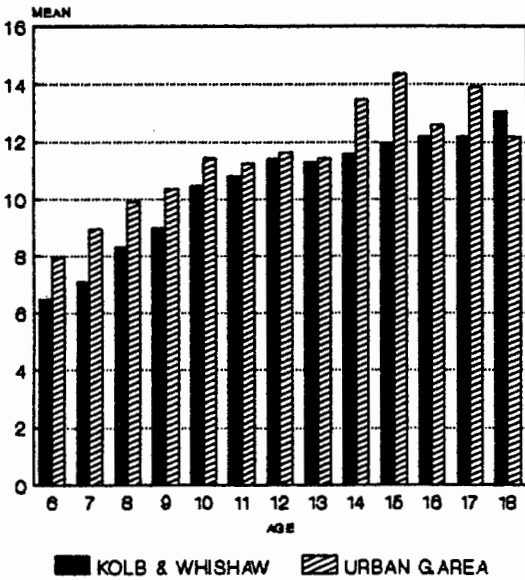
Linear regression equations were determined for Kolb's and Whishaw's normative sample and the "Urban Geographical Area" sample for ages 6 - 18 years. A test for equality of the slope coefficients was then performed and no statistically significant difference was found (see Table 13, Appendix iv). This suggests that the rate of improvement of performance with age can be regarded as equivalent for the Canadian and Zulu-speaking sample. The urban, Zulu-speaking sample exhibited an average increase of 0.43 units per year and the Canadian sample an average of 0.51 units per year (see Fig. 14).

Figure 13

Figure 14

**DRAW-A-BICYCLE
SCORES BY AGE AND G.AREA**

**DRAW-A-BICYCLE
SCORES BY AGE AND G.AREA**



(c) Kolb's and Wishaw's normative sample compared to the Total study sample, per age group:

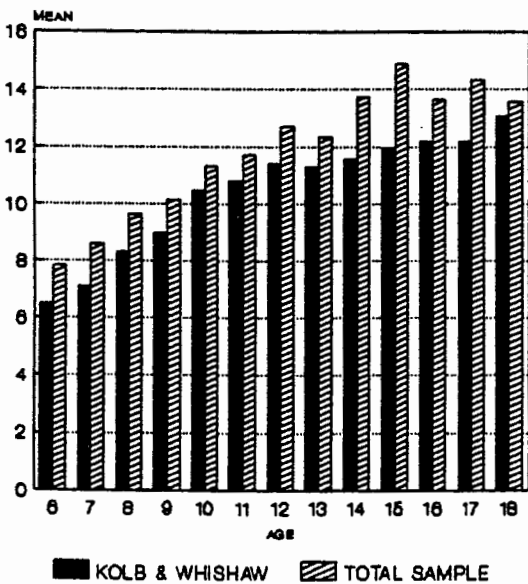
The comparison between the means per age of these two samples can be seen in graphic form in Figure 15. Results of the Student's t-Tests per age group can be seen in Table 14 (Appendix iv). These results indicate that, on average, Zulu-speaking children perform significantly better at ages 9, 14, 15, 16 and 17 than their Canadian counterparts.

Linear regression equations were determined for Kolb's and Wishaw's normative sample and the "total study" sample for ages 6 - 18 years. A test for equality of the slope coefficients was then performed and no statistically significant difference was found (see Table 15, Appendix iv). This suggests that the rate of improvement of performance with age can be regarded as equivalent for the Canadian and combined rural and urban, Zulu-speaking sample. The Zulu-speaking sample exhibited an average increase of 0.54 units per year and the Canadian sample an average of 0.51 units per year (see Figure 16).

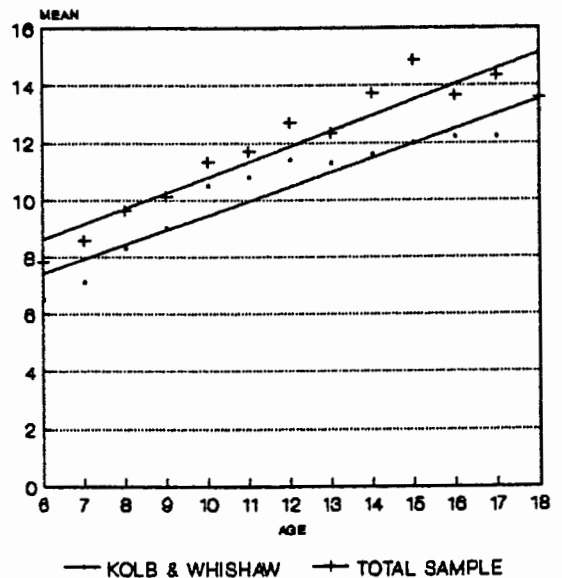
Figure 15

Figure 16

**DRAW-A-BICYCLE
MEAN SCORES BY AGE**



**DRAW-A-BICYCLE
MEAN SCORES BY AGE**



A comparison of the mean DAB scores per age group for the Kolb's and Whishaw's sample and the Total Study sample, using the Student's t-Test, clearly indicated that performances can not be regarded as equivalent. At 5 out of the 13 age levels, the Zulu-speaking children outperformed their Canadian counterparts. Testing the equality of the slope coefficients, which represent the rate of change in DAB score as a function of age, showed that these two samples improve performance, with age, at an equivalent rate. Thus, although the children in both samples develop visual-motor skills at an equivalent rate, the Zulu-speaking children appear to start at a higher level at age 6 years and appear to maintain this advantage throughout their schooling. For this reason the normative data cannot be regarded as equivalent. (Kolb's and Whishaw's normative data is therefore evidently inappropriate for use on Zulu-speaking subject).

1.2(ii) Internal Analysis:

A 3-way Analysis of Variance was performed in order to elicit the effects of the variables Age, Sex and Geographical Area, to determine the magnitude of each, and to elicit the complexities of the interactions (Table 16). Educational Level, as a variable was not included because of its high correlation to Age ($r=0.873$; $p<0.001$).

Table 16
DAB: 3-way ANOVA Summary Table for Independent Variables
Age, Sex, Geographical Area, on Dependent Variable DAB Score

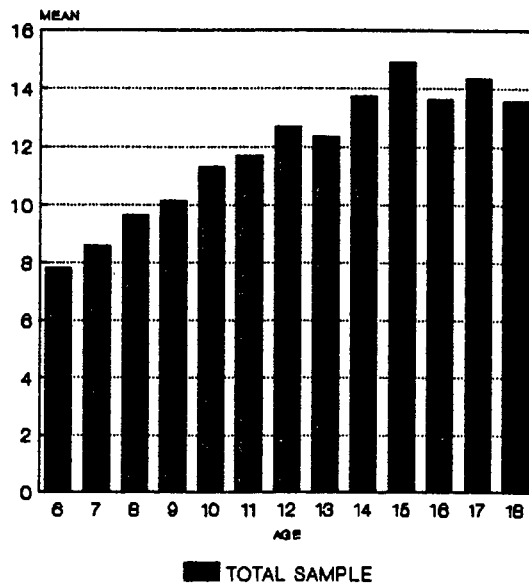
Source	df	Seq SS	Adj SS	Adj MS	F	P	eta ²
Age	11	2851.87	2106.05	191.46	28.28	0.0004	0.261
Sex	1	1755.29	824.36	824.36	121.76	0.0004	0.161
G. Area	1	139.14	33.79	33.79	4.99	0.026	0.013
Age*Sex	11	326.91	248.57	22.60	3.34	0.0004	0.030
Age*G. A	11	176.02	156.38	14.22	2.10	0.018	0.016
Sex*G. A	1	19.72	0.86	0.86	0.13	0.722	
Age*Sex*G. A	11	109.87	109.87	9.99	1.48	0.135	
Error	818	5538.30	5538.30	6.77			
Total	865	10917.13					

It is noted that Age, Sex and Geographical Area are significant main effects. In order to determine the amount of variance due to each, η^2 was calculated. Each significant effect will be independently discussed:

Age is a main effect responsible for significant variance across the Study Sample, shown by the large η^2 value. An η^2 value of 0.261 means that the variable age explains approximately 26% of the variation in the performance within the sample. The regression equation of DAB score against age had a slope coefficient of 0.544, which means that the DAB score for Zulu-speaking children between the ages of 6 and 18 years increases at a relatively constant rate of 0.544 units per year. This relationship is graphically represented in Figure 17.

Figure 17

DRAW-A-BICYCLE MEAN SCORES BY AGE



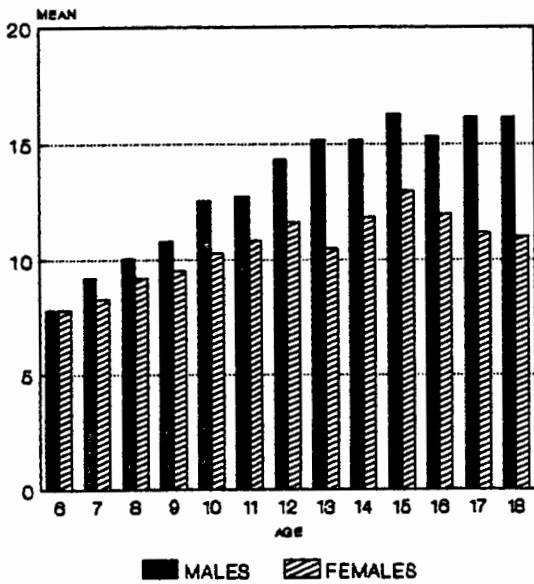
Sex is a main effect responsible for significant variance across the Study Sample, shown by the large η^2 value. An η^2 value of 0.161 means that the variable sex explains approximately 16% of the variation in the performance within the sample.

The comparison between the means, per age, of males and females can be seen in graphic form in Figure 18. Results of the Student's t-Tests per age group can be seen in Table 17 (Appendix iv). These results indicate that, on average, Zulu-speaking boys outperform their female counterparts at ages 9, 10, 11, 12, 13, 14, 15, 16, 17, and 18 years.

A Linear regression was performed for males and females for ages 6 - 18 years. A test for equality of the slope coefficients was then performed and a statistically significant difference was found (see Table 18, Appendix iv). These results indicate that males develop their bicycle-drawing skills more rapidly between the ages of 6 - 18 years than do females. On average, males improve their performance by 0.71 units per year as opposed to 0.32 units per year for females. This is graphically represented in Figure 19.

Figure 18

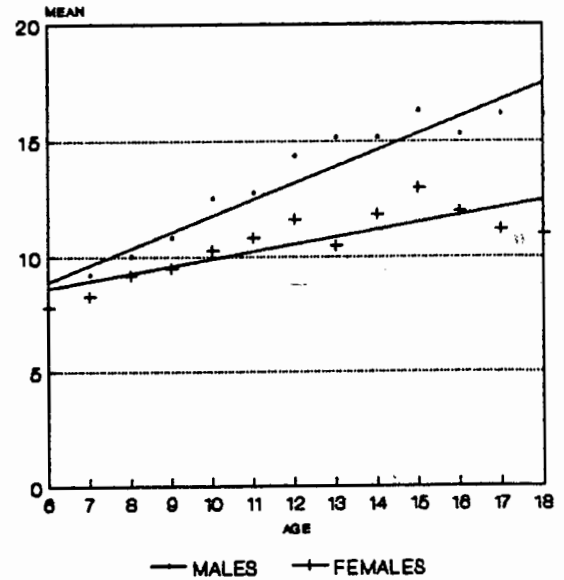
**DRAW-A-BICYCLE
SCORES BY AGE AND SEX**



TOTAL SAMPLE

Figure 19

**DRAW-A-BICYCLE
SCORES BY AGE AND SEX**



TOTAL SAMPLE

Geographical Area is reflected as a main effect by the ANOVA. Its relative effect size is small but significant, as shown by the eta² value of 0.013. This means that the variable Geographical Area explains approximately 1.3% of the performance variation within the sample. The comparison between the means, per age, of "Rural Geographical Area" sample and the "Urban

Geographical Area" sample can be seen in graphic form in Figure 20. Results of the Student's t-Tests per age group can be seen in Table 19 (Appendix iv). These results indicate that, on average, children from the rural areas perform better than their urban counterparts at the 12, 13, 16, and 18 year old age levels.

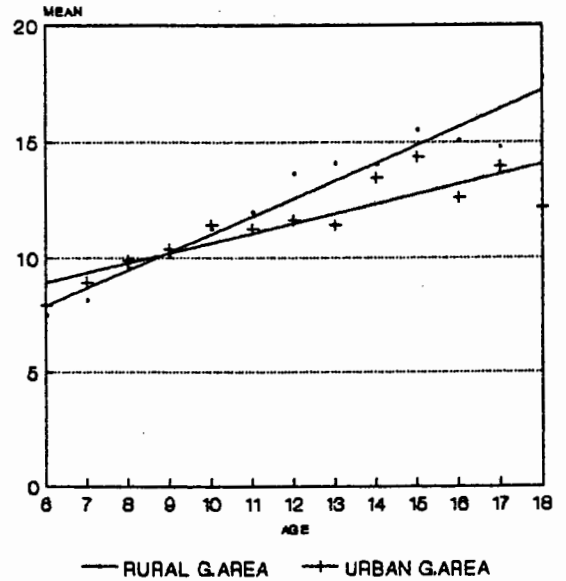
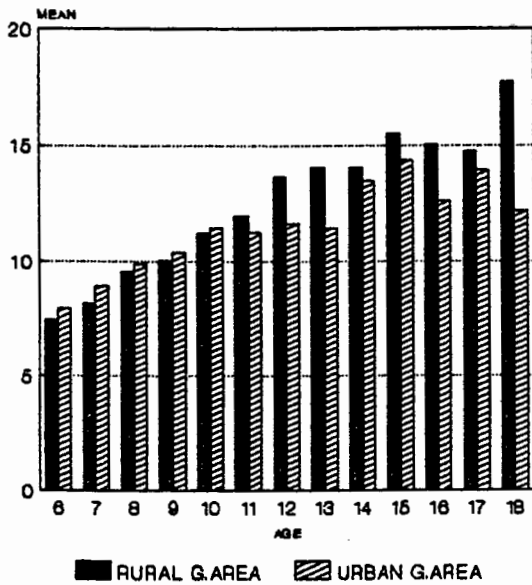
A Linear regression was performed on the "Rural Geographical Area" sample and the "Urban Geographical Area" sample for ages 6 - 18 years. A test for equality of the slope coefficients was then performed and a statistically significant difference was found. On average, children from rural areas exhibit an increase of 0.77 units per year as opposed to 0.43 units per year for urban children (see Table 20, Appendix iv). The relationship of the slopes is graphically represented in Figure 21.

Figure 20

Figure 21

**DRAW-A-BICYCLE
SCORES BY AGE AND G.AREA**

**DRAW-A-BICYCLE
SCORES BY AGE AND G.AREA**



Based on the above analysis it is clear that age is the predominant variable in terms of its magnitude of effect. Sex and geographical area are significant main effects and interact significantly with age. The relatively small effect of Geographical Area would appear to warrant collapsing the norms across this variable and warrant reporting norms per age group and

norms for each sex, per age group, only (see Appendix v). This is in accordance with Kolb's and Whishaw's findings and manner of reporting their norms (Kolb & Whishaw, 1985).

Thirty-two randomly selected DAB protocols (minimum of 2 per age group, per geographical area) were independently scored by a clinician using the Kolb and Whishaw standardised scoring method. A correlation was performed to establish inter-rater reliability and yielded $r = 0.936$ ($p < 0.001$)

Finally, it should be mentioned that a statistically significant correlation between the performances on the BG and the DAB became evident when a correlation was performed, yielding a coefficient of $r = -0.910$ ($p < 0.001$). Considering that the BG and the DAB are both measures of visual-motor ability, this result is not surprising.

2. DISCUSSION

2.1 Bender Gestalt:

A comparison of the mean BG scores, per age group, between the Koppitz's sample and the Total Study sample, using the Student's t-Test, clearly indicated that performances can not be regarded as equivalent. The results suggest that these two samples do not improve performance, with age, at an equivalent rate but that the American children improve their performance more quickly over the 6 - 10 year old period than do their Zulu-speaking counterparts. The American children appear to reach "maturity" for this task at approximately 9 years of age whereas the Zulu-speaking children do so at approximately 13 years of age. For this reason the normative data can not be regarded as equivalent and clearly Koppitz's normative data are inappropriate for use with Zulu-speaking children.

A number of factors are considered in an attempt to understand these differences:

Firstly, differences in sampling procedure may well have contributed to these discrepancies. As previously mentioned, the sample in the present study must be construed as an opportunity or convenience sample in that all pupils

present at the time of administration were unconditionally selected for participation. Thus, there was no screening in order to exclude those with a history of organic pathology or psychiatric disturbance. Additionally, although equivalent numbers were selected from rural and urban areas, it must be noted that a large proportion of the "urban sample" were in fact rurally-based children attending boarding school in an urban area. So, although they were receiving an "urban education", many had spent at least 6 years of their early development in a "rural" environment (the known effects of "early development within a rural environment" on visual-motor ability is unclear and beyond the scope of this study, but it may be speculated, for example, that limited exposure to pen and paper tasks in early life may result in poor performance on tests such as the BG later on).

It could therefore be argued that the sampling method employed in this study was not rigorous enough to reflect a true cross-section of the Zulu-speaking population. The Koppitz's normative sample appeared to be more tightly controlled and comprised subjects with a wider range of geographical backgrounds (Koppitz, 1975). Future studies in South Africa may be strengthened by controlling for geographical area and socio-economic background.

Important to note however is that the above recommendation may be more difficult to execute than expected in the current socio-politico-economic climate where black people appear to be geographically and socially more mobile than other population groups. For example, rapid urbanisation of many rurally-based communities has seen the establishment of urban "townships" and a radical change from rural to "urban" schooling for the children. Violence, in rural and urban areas, has resulted in massive upheaval for children, often altering their geographical and socio-economic status. Further, because of the very poor educational resources offered in the rural areas, many rural families choose to send their children to "urban" schools, thereby often changing their geographical and socio-economic status. Adequate controlling for these variables therefore may be extremely difficult. With this sociological phenomenon in mind, the representativeness of this study's sample could possibly be positively re-appraised - i.e. this sample of black children drawn from "rural" and "urban" schools may well be a valid representation of all black school children in South Africa.

Secondly, discrepancies may be attributed to methods of administration. It has been explained that a group-administration procedure was adopted in the present study, which is non-consistent with the existing standardised recommendations (Koppitz, 1975). Although the study by Keogh and Smith (1971) showed that no significant difference resulted between individual and group administration methods, no additional studies supporting this claim have been reported.

Also, the fact that different teachers were used as (interpreting) instructors, brings into question the standardisation of the administration procedure. It may be a recommendation for future research of this nature to control for these variables.

Thirdly, as was briefly mentioned in the introduction, the South African State education policy has been ideologically bound, which means that factors such as self-esteem, class, and socio-economic standing are automatically implicated when social sub-groups are considered. As a group, the study sample, drawn from Zulu-speaking, rural, "semi-rural" and "urban" school children, is likely to represent a predominantly low socio-economic class. If the possibility is considered that poorer cognitive and/or perceptual development is linked to lower socio-economic background (Willis and Pishkin, 1974), then it becomes apparent that the study sample may not be well correlated to the normative sample for the variable, "level of intelligence", as defined by Koppitz (Koppitz, 1975).

Lastly, the nature of the "Bantu Education System", which has operated for more than 40 years in South Africa, has resulted in an inferior system of tuition for blacks. The lack of resources available (as witnessed by the clinician) at all the schools visited, from the "smartest" urban school to the "poorest" rural school, served as evidence of this ideologically-based, inferior system. It may be speculated that educational deprivation of this nature could, amongst other things, manifest itself in many areas of cognitive functioning such as delays or deficits in visual-motor ability. A question which then needs to be asked is:

"Do these norms represent the "true" capabilities of Zulu-speaking children - ie, a combination of the current and potential capabilities - or merely the current capabilities of

the majority of Zulu-speaking children who may not have reached their potential as a result of a deprived education system?"

Addressing this question is clearly not within the scope of this study, but it does however raise some important methodological points. The need for future studies of this nature in South Africa to control for variables such as socio-economic background, level of intelligence, and system of education, is essential.

Of importance is that, from a practical point of view, it would seem that the norms developed here (Appendix v) are clinically useful in that they are a realistic reflection of the current capabilities of the majority of Zulu-speaking children between the ages of 6 - 18 years. With recent political changes and an evolving education system in South Africa a 10-year follow-up study to determine the reliability of these norms may well be indicated.

Based on the analysis of the variables, it is clear that age is the predominant variable in terms of its magnitude of effect. Both sex and geographical area are statistically significant variables according to the 3-Way ANOVA, but their relative effect size is negligible according to their η^2 values. This warrants collapsing the norms across them and thus reporting only age-related norms. Because of the high correlation between age and educational level, the reporting of grade norms is deemed superfluous.

Concerning the reliability of the scoring system employed - Koppitz's Developmental Bender Scoring System (Koppitz, 1975) - a high correlation, and thus inter-rater reliability, was found between two independent scorers of the study sample's protocols. This adds to the clinical usefulness of the norms developed here in that clinicians are assured that their scoring is reliable if they stick to the above scoring system.

2.2 Draw-A-Bicycle:

A comparison of the mean DAB scores, per age group, between Kolb's and Whishaw's sample and the Study sample, using the Student's t-Test, clearly indicated that performances can not be regarded as equivalent. The results suggest that the children in both samples develop visual-motor perceptual skills at an equivalent rate between the ages of 6 - 18 years, but that the

Zulu-speaking children perform at a higher level in the early years and appear to maintain this advantage throughout their schooling. For this reason the normative data cannot be regarded as equivalent and clearly Kolb's and Whishaw's normative data are inappropriate for use with Zulu-speaking children.

A major contributing factor to the above results would appear to be the low socio-economic status of black people in South Africa. One repercussion thereof is that the black population has historically used the bicycle as one of its primary sources of transport. The result is that the majority of black children see bicycles being used by their own families or communities on a daily basis throughout their lives, resulting in a high level of familiarity. Additionally, it is speculated that bicycles become status symbols under these circumstances and are regarded as prized possessions by many black children, so also increasing their awareness and familiarity with these objects.

A factor which may, to some extent, explain the significantly higher performance of 12 - 17 year old children in the "rural geographical area" sample compared to those in the "urban geographical area" sample is a practically related one. At the rural high school visited, a system was in operation in which four bicycles, belonging to the school, served as "pool vehicles" for the pupils (for buying food and drinks at the local trading store for teachers and pupils during lunch times). This meant that all children were constantly exposed to the bicycles, with males being exclusively afforded the opportunity to ride them (females were only occasionally allowed to ride the bicycles, according to the teachers). It is suggested, then, that the relative familiarity with the bicycles - for the males, actually riding them and for the females, constantly seeing them - is another example of the kind of exposure black children have and may go a long way in explaining the superior bicycle drawing ability of these children. This increased visual-spatial memory ability may result from what Vygotsky (cited in Thomas, 1985) termed the child's "culture" or "language environment". According to him, the history of that particular "culture" and the history of the child's experience within it, determine the child's cognitive abilities.

Based on the analysis of the variables it is clear that Age is the predominant variable in terms of its magnitude of effect. Sex and Geographical Area are significant main effects and interact significantly with age. The relatively small effect of Geographical Area however would appear to warrant collapsing the norms across this variable and warrant reporting norms per age group and norms for each sex, per age group, only. This is in accordance with Kolb's and Whishaw's findings and manner of reporting their norms. Because of the high correlation between Age and Educational Level, the reporting of grade norms is deemed superfluous.

A clear phenomenon emerging from the results is the difference in performance between sexes, where males clearly outperform females from the age of 9 years upwards. It would appear that the predominant contributing factor here is socio-cultural. The above example concerning access to the communal bicycles at school is evidence of this and goes some way in explaining why males are able to produce more competent drawings of bicycles. It is speculated that the effects of socialisation do not manifest clearly in tasks such as the DAB before the age of 9 years and thus males' and females' performances do not differ significantly until such time.

Concerning the reliability of the DAB, a high correlation was found between the results of two clinicians who independently scored thirty protocols. This indicates high inter-rater reliability. Additionally, the group administration system used in this study is in accordance with that used by Kolb and Whishaw in their original normative study (Whishaw, 1992). These factors add to the clinical usefulness of the norms developed here in that clinicians are assured that their testing is relatively reliable if they stick to the standardised administration and scoring system (Kolb and Whishaw, 1985).

2.3 Comparison of the BG and the DAB:

A statistically significant correlation between the performances on the BG and the DAB became evident when a correlation of mean scores across age groups was performed, yielding a coefficient of $r = -0.910$ ($p < 0.001$). Upon the assumption that both tests are valid tools for assessing visual-motor perceptual ability, this result can be seen as an indicator of inter-test reliability. Clinically, this means that these tests can be used

independently, with more confidence, to measure visual-motor perceptual ability. Practically, this means that they can be used in conjunction with one another to back-up neuropsychological findings.

Although it has fallen beyond the scope of this study, Intelligence, as a variable, needs serious consideration. The strong correlation between Intelligence and performance on the BG and similar tests, is borne out by previously mentioned studies (Johnson & Lottman, 1971; Koppitz, 1975). It seems, then, that there is a need for the investigation into and development of "IQ" corrected norms for, both, the BG and the DAB.

CONCLUSION

Crawford-Nutt (1977) warns against comparing a person's test score with an inappropriate norm group because of the real danger of getting unreliable or totally invalid assessment results. This would appear to be the very danger faced by most South African clinicians in their dealings with the diversity of cultures and sub-groups within South Africa. The scarcity of locally developed norms results in clinicians having to make inferences using normative data primarily based on foreign population groups.

This dissertation, then, has addressed this issue and explored the use of two visual-motor perceptual ability tests (the BG and the DAB) within the sample group of Zulu-speaking school children aged 6 - 18 years. It has investigated this population's performance on both tests and investigated the appropriateness of using the existing, foreign normative data.

The results clearly indicated the danger of using foreign norms, in that, for both the BG and the DAB, the performance of the population group investigated proved to be significantly different to the respective performances of the foreign normative samples (namely, American children for the BG, and Canadian children for the DAB). Using these foreign norms for Zulu-speaking children therefore would be unreliable and inappropriate, and would provide invalid assessment information.

This study showed that Age was the variable most highly correlated with test performance for both tests and that it accounted for most of the variance.

Educational Level was also identified as correlating highly with performance but because of its very high correlation with Age, was not dealt with or reported separately. For the BG, the variables Sex and Geographical Area were also correlated with performance, but in both cases this was small enough to warrant collapsing the norms across these variables and only reporting normative data per Age-level. For the DAB, Geographical Area was correlated with performance but was also small enough to warrant collapsing the norms across this variable. The variable Sex, however, was highly correlated with performance and so normative data was reported for Age, broken down by Sex (see Appendix v). Additionally, this study demonstrated a high correlation between performance on the BG and the DAB, per age group. It is proposed that this is an indication of the high inter-test reliability, i.e. an indication that these tests are of similar strength and efficiency in measuring visual-motor perceptual skills.

Of clinical significance is the fact that this study has resulted in useful normative data for the BG and the DAB tests for Zulu-speaking children aged 6 - 18 years. This means that clinicians will be able to employ these tools of visual-motor perceptual functioning assessment with far greater confidence than in the past.

As was mentioned earlier, these norms may need reviewing in 10 - 15 years time. It would be unwise to regard these norms as "stable" considering the progressive policy proposals for black education in South Africa. Political, economic and social changes may well have a significant impact on the development of children from backgrounds which were formerly "deprived".

These norms were developed with regard to a specific sub-population, namely, Zulu-speaking children. A question which must be asked is whether or not these norms are applicable to other sub-populations in South Africa who share a similar background - i.e. children who have been exposed to similar social, political, economic and educational conditions? This would include most African children, for example, Xhosa-speaking and Sotho-speaking children. It is hypothesised that these norms would prove to be appropriate for a far wider population group than Zulu-speaking children only - for example, all "African children in South Africa". In order to establish this, comparative research is needed. Clearly, further assessment of these norms in terms of their appropriateness for use with other sub-groups of children in South

Africa, would serve to establish their clinical value and applicability.

In the light of the above comments, it becomes apparent that studies aimed at developing local normative data, and thus increasing the usefulness of psychological tests in general in the South African context, are essential if we hope to provide an adequate service to the communities in this country in the future.

REFERENCES

x

- Cole, M. & Means, B. (1981). **Comparative studies of how people think.** Cambridge, Mass: Harvard University Press.
- Crawford-Nutt, D.H. (1977). The effect of educational level on the test scores of people in South Africa. **Psychologia Africana**, 17, 49-59.
- Edwards, G.L. (1988). The Bender Gestalt with adolescents: Comparisons of two scoring systems. **Journal of Clinical Psychology**, 44, 226-230.
- Fiedler, M. & Schmidt, E. (1969). Sex differences in Bender-Gestalt drawings of seven year old children. **Perceptual and Motor Skills**, 29, 753-754.
- Hutt, M. & Briskin, G. (1960). **The Clinical Use Of The Revised Bender Gestalt Test.** New York: Grune & Stratton, Inc.
- Johnson, J. and Lottman, T. (1971). The relationship between intelligence, brain damage, and Hutt-Briskin errors on the Bender Gestalt. **Journal of Clinical Psychology**, 21, 84-85.
- Keogh, B. and Smith, C. (1971). Group techniques and proposed scoring system for the Bender Gestalt Test with children. **Journal of Clinical Psychology**, 21, 172-175.
- Kolb, B. and Whishaw, I. (1985). **Fundamentals of Human Neuropsychology.** New York: Freeman and Company.
- Koppitz, E. (1960). The Bender Gestalt Test for children: A normative study. **Journal of Clinical Psychology**, 16, 432-435.
- Koppitz, E. (1975). **The Bender Gestalt Test for Young Children.** New York: Grune and Stratton, Inc.
- Lezak, M.D. (1983). **Neuropsychological Assessment.** New York: Oxford University Press, Inc.

- Marmorale, A. and Brown, F. (1977). Bender Gestalt performance of puertorican, white, and negro children. **Journal of Clinical Psychology**, 33, 224-227.
- McCall, R.B. (1986). **Fundamental Statistics For Behavioural Sciences**. New York: Harcourt Brace Jovanovich, Inc.
- McFie, J., Piercy, F. & Zangwill, O.L. (1950). Visual-spatial agnosia associated with lesions on the right cerebral hemisphere. **Brain: A Journal Of Neurology**, 73, 167-190.
- McFie, J. & Zangwill, O.L. (1960). Visual-constructive disabilities associated with lesions of the left cerebral hemisphere. **Brain: A Journal Of Neurology**, 83, 243-260.
- Pascal, G.R. & Suttel, B.J. (1951). **The Bender Gestalt Test: Quantification and Validity for Adults**. New York: Grune & Stratton.
- Nichols, M. (1980). **Reliability and Validity of the Bicycle Drawing Test**. Unpublished dissertation submitted for the degree of MA. Oregon Health Sciences University.
- Paterson, A. & Zangwill, O.L. (1944). Disorders of visual space perception associated with lesions of the right cerebral hemisphere. **Brain: A Journal Of Neurology**, 67, 331-358.
- Peixotto, H.E. (1954). The Bender Gestalt Visual Motor Test as a culture free test of personality. **Journal of Clinical Psychology**, 10, 369-372.
- Snyder, R., Holowenzak, S. and Hoffman, N. (1971). A cross-cultural item analysis of Bender Gestalt protocols administered to ghetto and suburban children. **Perceptual and Motor Skills**, 33, 791-796.
- Taylor, E.M. (1959). **The Appraisal Of Children With Cerebral Deficits**. Cambridge, Mass: Harvard University Press.
- Taylor, H. and Thweatt, C. (1972). Cross-cultural developmental performance of Navajo children on the Bender Gestalt Test. **Perceptual and Motor Skills**, 35, 307-309.

Taylor, R. and Partenio, I. (1984). Ethnic differences on the Bender Gestalt: Relative effects of measured intelligence. **Journal of Consulting and Clinical Psychology**, 52, 784-788.

Thomas, R.M. (1985). **Comparing Theories Of Child Development**. Belmont: Wadsworth, Inc.

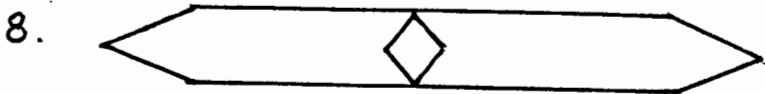
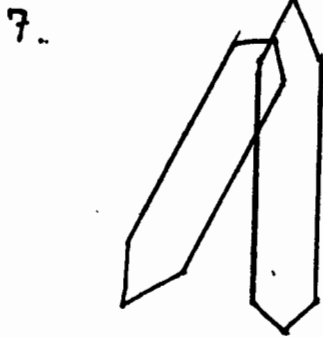
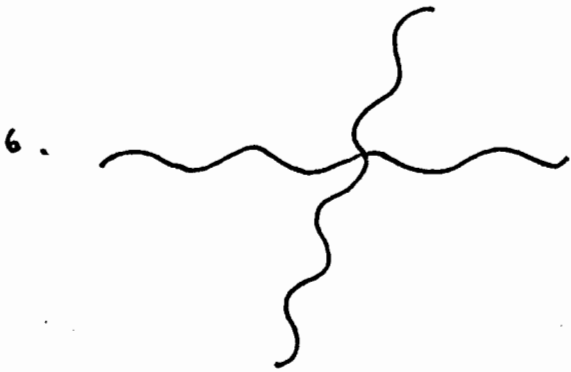
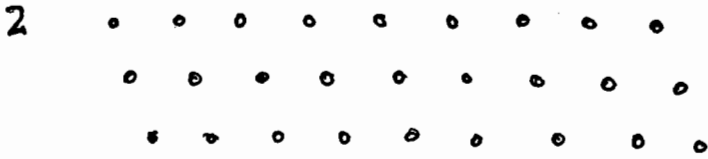
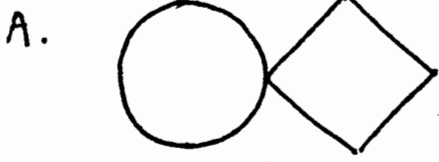
Whishaw, I. (1992). Personal communication^a.

Willis, D. and Pishkin, V. (1974). Perceptual-motor performance on the Vane and Bender Tests as related to two socio-economic classes and ages. **Perceptual and Motor Skills**, 38, 883-890. ✓

Zuelzer, M. and Stedman, J. (1976). Koppitz Bender Gestalt scores in first grade children as related to ethnocultural background, socio-economic class, and sex factors. **Journal of Consulting and Clinical Psychology**, 44, 875.

^aContact address: Department of Psychology
University of Lethbridge, Lethbridge,
Alberta, Canada, T1K 3M4

APPENDIX I



APPENDIX II

Kolb & Whishaw (1985) Scoring System for the Draw-A-Bicycle Test:

<u>Part</u>	<u>Representation</u>	<u>Score</u>
Wheels	2 Circles	2 (1 each)
	2 Tyres	1
	Spokes	2 (any spokes, 1 for each wheel)
Frame	Front Wheel	1
	Back Wheel	1
	Middle "v"	1
	Seat	1
	Handlebars	1
Drive	Rear Wheel Sprocket	1
	Front Sprocket	1
	Chain	1
	2 Pedals	1
	2 Pedal Supports	1
Function	Can Coast	1
	Can Be Driven	1
	Parts in Right Place	1
	Parts in Proportion	1
	Well Drawn	1

APPENDIX III

The purpose of the study by Keogh and Smith (1971) was to investigate the effects of modified administrative techniques on the Bender Gestalt test performance of children. The test was administered to three matched groups of children using three different methods. The first was in the standard individual way; the second was in group form, in which the subjects were presented with the individual figures, enlarged to scale, from a position at the front of the group; and the third was also in group form, in which each subject received his\her own set of figures. Results showed no significant difference in performance between the methods used. Of importance is that the latter method proved to require a far greater degree of supervision because many found it confusing.

APPENDIX IV

Table 1

Student's t-Tests¹ Per Age:
Koppitz Compared to Rural G. Area²

Age	df	t(obs)	Age	df	t(obs)
6	339	-0.522	13	96	-2.277
7	277	-1.416	14	120	0.828
8	159	-3.300**	15	100	0.728
9	152	-8.287***	16	91	1.539
10	96	-5.639***	17	80	0.120
11	106	-3.515***	18	60	1.100
12	95	-2.852**			

** = $p < 0.01$; *** = $p < 0.001$

Table 2

Linear Regression Equations and Test for Equality of
Slope Coefficient t(obs), Between Koppitz and Rural G. Area

	r (corr)	b (slope)	sd of b	t	df
Koppitz	-0.958 ³	-1.466	0.254	7.80***	1238
Rural G. A.	-0.951	-0.979	0.183		

** = $p < 0.01$; *** = $p < 0.001$

¹The large number of t-Tests performed, which increases the rate of Type-1 errors (McCall, 1986), was compensated for by selecting a "primary" significance level of $p < 0.01$ and a "secondary" significance level of $p < 0.001$.

²All standard deviations and numbers of subjects for t-Tests performed are presented in Appendix vi.

³In all tests involving slopes the correlation between the age-levels and the BG means, per age group, was calculated to establish that the relationship was linear.

Table 3

Student's t-Test Per Age:
Koppitz Compared to Urban G.Area

Age	df	t(obs)	Age	df	t(obs)
6	350	2.778**	13	132	-0.350
7	280	0.638	14	133	1.087
8	130	-0.457	15	107	1.538
9	128	-2.290	16	102	1.524
10	89	-2.523	17	84	1.283
11	84	-2.405	18	68	1.881
12	90	-0.318			

** = P<0.01: *** = p<0.001

Table 4

Linear Regression Equations and Test for Equality of
Slope Coefficient t(obs), Between Koppitz and Urban G.Area

	r (corr)	b (slope)	sd of b	t	df
Koppitz	-0.958	-1.466	0.254	7.78***	1192
Urban G.A.	-0.986	-0.558	0.055		

** = p<0.01; *** = p<0.001

Table 5

Student's t-Tests Per Age:
Koppitz Compared to Total Study Sample

Age	df	t(obs)	Age	df	t(obs)
6	356	2.107	13	172	-1.222
7	293	-0.458	14	197	1.112
8	169	-3.058**	15	151	1.423
9	168	-7.331***	16	137	1.906
10	129	-4.878***	17	108	0.966
11	134	-3.577***	18	72	2.169**
12	129	-1.940			

** = p<0.01; *** = p<0.001

Table 6

Linear Regression Equations and Test for Equality of Slope Coefficient t(obs), Between Koppitz and Total Sample

	r (corr)	b (slope)	sd of b	t	df
Koppitz	-0.958	-1.466	0.254		
Total Sample	-0.985	-0.534	0.055	10.45***	1330

** = $p < 0.01$; *** = $p < 0.001$

Table 8

Linear Regression Equations and Test for Equality of Slope Coefficient t(obs), Between Males and Females

	r (corr)	b (slope)	sd of b	t(obs)	df
Males	-0.954	-0.374	0.036		
Females	-0.978	-0.484	0.031	0.823	891

Table 9

Linear Regression Equations and Test for Equality of Slope Coefficient t(obs), Between Rural and Urban G.Area

	r (corr)	b (slope)	sd of b	t(obs)	df
Rural G.A	-0.949	-0.559	0.056		
Urban G.A.	-0.966	-0.328	0.027	14.31***	878

** = $p < 0.01$; *** = $p < 0.001$

Table 10

Student's t-Tests Per Age:
Koppitz Compared to Rural G. Area

Age	df	t(obs)	Age	df	t(obs)
6	43	-0.929	13	272	-4.594***
7	64	-1.246	14	311	-5.052***
8	75	-2.221	15	427	-6.586***
9	78	-2.114	16	422	-4.531***
10	92	-1.205	17	383	-3.344*
11	90	-1.785	18	219	-2.436
12	144	-3.385***			

** = $p < 0.01$; *** = $p < 0.001$

Table 11

Linear Regression Equations and Test for Equality of
Slope Coefficient t(obs), Between Kolb & Whishaw and Rural G. Area

	r (corr)	b (slope)	sd of b	t	df
Kolb & Whishaw	0.948	0.508	0.051	-3.532***	3099
Rural G.A.	0.973	0.774	0.055		

** = $p < 0.01$; *** = $p < 0.001$

Table 12

Student's t-Test Per Age:
Kolb & Whishaw Compared to Urban G. Area

Age	df	t(obs)	Age	df	t(obs)
6	54	-2.101	13	308	-0.279
7	67	-2.506	14	324	-4.138***
8	46	-2.001	15	434	-4.639***
9	54	-2.231	16	433	-0.714
10	85	-1.259	17	387	-2.399
11	68	-0.627	18	227	0.839
12	139	-0.339			

** = $P < 0.01$; *** = $p < 0.001$

Table 13

Linear Regression Equations and Test for Equality of Slope Coefficient $t(\text{obs})$, Between Kolb & Whishaw and Urban G. Area

	r (corr)	b (slope)	sd of b	t	df
Kolb & Whishaw	0.948	0.508	0.051	-0.95	3105
Urban G. A.	0.876	0.425	0.070		

Table 14

Student's t -Tests Per Age:
Kolb & Whishaw Compared to Total Study Sample

Age	df	$t(\text{obs})$	Age	df	$t(\text{obs})$
6	60	-2.181	13	348	-2.592
7	80	-2.473	14	388	-5.870***
8	85	-2.529	15	478	-7.492***
9	94	-2.619**	16	468	-3.369**
10	125	-1.511	17	411	-3.933***
11	118	-1.617	18	231	-0.471
12	178	-2.504			

** = $p < 0.01$; *** = $p < 0.001$

Table 15

Linear Regression Equations and Test for Equality of Slope Coefficient $t(\text{obs})$, Between Kolb & Whishaw and Total Sample

	r (corr)	b (slope)	sd of b	t	df
Kolb & Whishaw	0.948	0.508	0.051	-0.449	3543
Total Sample	0.938	0.543	0.061		

Table 17

Student's t-Test Per Age:
Males Compared to Females on the DAB

Age	df	t(obs)	Age	df	t(obs)
6	21	0.012	13	114	9.125***
7	27	0.943	14	139	6.684***
8	47	1.152	15	93	5.568***
9	54	2.621**	16	79	5.598***
10	71	3.607***	17	50	7.732***
11	76	3.295**	18	14	5.083***
12	71	3.773***			

** = P<0.01: *** = p<0.001

Table 18

Linear Regression Equations and Test for Equality of
Slope Coefficient t(obs), Between Males and Females

	r (corr)	b (slope)	sd of b	t(obs)	df
Males	0.954	0.712	0.067	4.205***	891
Females	0.824	0.317	0.066		

** = p<0.01; *** = p<0.001

Table 19

Student's t-Test Per Age:
Rural G.Area Compared to Urban G.Area on the DAB

Age	df	t(obs)	Age	df	t(obs)
6	21	-0.463	13	114	4.121***
7	27	-0.860	14	139	1.026
8	47	-0.371	15	93	1.683
9	54	-0.560	16	79	3.807***
10	71	-0.256	17	50	0.942
11	76	1.103	18	14	4.374***
12	71	2.725**			

** = P<0.01: *** = p<0.001

Table 20

Linear Regression Equations and Test for Equality of
Slope Coefficient $t(\text{obs})$, Between Rural and Urban G.Area

	r (corr)	b (slope)	sd of b	t(obs)	df
Rural G.A	0.973	0.774	0.055	3.914***	878
Urban G.A.	0.876	0.425	0.070		

** = $p < 0.01$; *** = $p < 0.001$

APPENDIX V

INITIAL NORMATIVE DATA FOR ZULU-SPEAKING CHILDREN

1. Bender Mean Scores and Standard Deviations by Age:

Age	(n)	Mean	S. D.
6	(23)	5.5652	2.9206
7	(29)	5.0690	3.0698
8	(49)	4.7347	2.5963
9	(56)	4.2143	2.7812
10	(73)	3.3288	2.2051
11	(78)	2.8077	2.1324
12	(73)	2.2329	2.0718
13	(116)	1.9397	2.0272
14	(141)	1.2482	1.6611
15	(95)	1.1684	1.3964
16	(81)	1.0123	1.4534
17	(52)	1.2500	1.2661
18	(16)	0.5000	0.7303
Total	(882)	2.2860	2.3994

2. Draw-A-Bicycle Mean Scores and Standard Deviations:

Age	Sex	(n)	Mean	S.D.
6	m	(12)	7.8	2.4
	f	(11)	7.8	1.5
	Total	(23)	7.8	2.0
7	m	(9)	9.2	3.2
	f	(20)	8.3	2.0
	Total	(29)	8.6	2.4
8	m	(25)	10.0	2.4
	f	(24)	9.2	2.7
	Total	(49)	9.6	2.5
9	m	(27)	10.8	1.7
	f	(29)	9.5	2.0
	Total	(56)	10.1	1.9
10	m	(34)	12.5	2.9
	f	(39)	10.3	2.5
	Total	(73)	11.3	2.9
11	m	(35)	12.8	2.7
	f	(43)	10.8	2.4
	Total	(78)	11.7	2.7
12	m	(29)	14.3	3.1
	f	(44)	11.6	3.0
	Total	(73)	12.7	3.3
13	m	(46)	15.2	2.9
	f	(70)	10.5	2.5
	Total	(116)	12.4	3.5
14	m	(80)	15.2	3.3
	f	(61)	11.9	2.3
	Total	(141)	13.7	3.3
15	m	(55)	16.3	2.7
	f	(40)	13.0	3.1
	Total	(95)	14.9	3.3
16	m	(41)	15.3	3.1
	f	(40)	12.0	2.2
	Total	(81)	13.7	3.1
17	m	(33)	16.2	2.6
	f	(19)	11.2	1.2
	Total	(52)	14.3	2.3
18	m	(8)	16.1	2.6
	f	(8)	11.0	1.2
	Total	(16)	13.6	3.3
ALL	m	(434)	14.1	3.6
	f	(448)	11.0	2.7
	Total	(882)	12.5	3.6

APPENDIX VI

BG MEAN SCORES, STANDARD DEVIATIONS AND CELL SIZES

1. CONTROL: garea = rural

ROWS: age COLUMNS: sex (0=m; 1=f)

	0	1	ALL
6	8.0000 3.0000 3	8.3333 1.1547 3	8.1667 2.0412 6
7	5.2500 4.1130 4	6.5556 2.5550 9	6.1538 2.9957 13
8	5.2000 2.7834 20	4.8421 2.4555 19	5.0256 2.6005 39
9	5.2222 2.9814 18	4.4545 2.6855 22	4.8000 2.8121 40
10	3.5455 2.5209 22	4.4444 1.9770 18	3.9500 2.3088 40
11	2.5000 1.9337 24	3.2308 2.1035 26	2.8800 2.0368 50
12	2.8333 1.9174 18	2.6190 2.3125 21	2.7179 2.1145 39
13	1.8947 1.3701 19	2.9524 2.3125 21	2.4500 1.9735 40
14	1.4865 1.9239 37	1.0000 1.3587 27	1.2813 1.7133 64
15	1.1739 1.4970 23	1.4286 1.4343 21	1.2955 1.4560 44
16	1.0435 1.8210 23	0.8333 0.7177 12	0.9714 1.5240 35
17	1.5882 1.1213 17	1.2857 1.3801 7	1.5000 1.1795 24
18	0.5000 0.5774 4	-- -- 0	0.5000 0.5774 4

2. CONTROL: gareea = urban

ROWS: age COLUMNS: sex (0=m; 1=f)

	0	1	ALL
6	3.8889 3.2575 9	5.5000 1.5119 8	4.6471 2.6444 17
7	3.4000 2.5100 5	4.5455 3.1421 11	4.1875 2.9262 16
8	2.8000 2.6833 5	4.4000 1.9494 5	3.6000 2.3664 10
9	2.1111 1.9003 9	3.5714 2.2991 7	2.7500 2.1448 16
10	2.0000 1.2792 12	2.9048 2.0471 21	2.5758 1.8376 33
11	2.0000 1.2649 11	3.1176 2.7587 17	2.6786 2.3262 28
12	1.6364 1.8586 11	1.6957 1.9641 23	1.6765 1.9025 34
13	1.4444 1.4763 27	1.7959 2.2637 49	1.6711 2.0159 76
14	1.1860 1.7762 43	1.2647 1.4420 34	1.2208 1.6273 77
15	1.2187 1.3850 32	0.7895 1.2727 19	1.0588 1.3478 51
16	0.8889 1.2783 18	1.1429 1.5084 28	1.0435 1.4135 46
17	1.4375 1.5042 16	0.5000 0.7977 12	1.0357 1.3189 28
18	0.5000 1.0000 4	0.5000 0.7559 8	0.5000 0.7977 12

DAB MEAN SCORES, STANDARD DEVIATIONS AND CELL SIZES

1. CONTROL: garea = rural

ROWS: age COLUMNS: sex(0=m; 1=f)

	0	1	ALL
6	8.000 1.732 3	7.000 2.646 3	7.500 2.074 6
7	9.500 4.203 4	7.556 2.007 9	8.154 2.824 13
8	9.900 2.360 20	9.211 2.974 19	9.564 2.664 39
9	11.000 1.815 18	9.273 2.097 22	10.050 2.136 40
10	11.864 2.748 22	10.500 1.948 18	11.250 2.488 40
11	13.000 3.065 24	11.000 2.498 26	11.960 2.934 50
12	14.611 3.567 18	12.810 3.572 21	13.641 3.638 39
13	15.895 2.807 19	12.476 2.960 21	14.100 3.334 40
14	14.946 3.543 37	12.815 2.512 27	14.047 3.302 64
15	16.957 2.325 23	13.905 2.256 21	15.500 2.741 44
16	16.478 2.695 23	12.333 2.605 12	15.057 3.298 35
17	16.235 2.635 17	11.286 1.254 7	14.792 3.243 24
18	17.750 2.217 4	-- -- 0	17.750 2.217 4

2. CONTROL: garea = urban

ROWS: age COLUMNS: sex (0=m; 1=f)

	0	1	ALL
6	7.778 2.635 9	8.125 0.991 8	7.941 1.983 17
7	9.000 2.739 5	8.909 1.868 11	8.937 2.081 16
8	10.600 2.702 5	9.200 0.837 5	9.900 2.025 10
9	10.444 1.509 9	10.286 1.380 7	10.375 1.408 16
10	13.750 2.768 12	10.095 2.879 21	11.424 3.317 33
11	12.273 1.902 11	10.588 2.347 17	11.250 2.303 28
12	13.909 2.300 11	10.522 1.780 23	11.618 2.511 34
13	14.667 2.948 27	9.653 1.786 49	11.434 3.300 76
14	15.349 3.169 43	11.088 1.694 34	13.468 3.366 77
15	15.813 2.879 32	11.947 3.551 19	14.373 3.638 51
16	13.778 2.861 18	11.821 2.019 28	12.587 2.544 46
17	16.063 2.744 16	11.083 1.240 12	13.929 3.333 28
18	14.500 1.915 4	11.000 1.195 8	12.167 2.209 12

BG MEAN SCORES, STANDARD DEVIATIONS AND CELL SIZES

CONTROL: total sample

ROWS: sex COLUMNS: age (0=m; 1=f)

	6	7	8	9	10	11	12	13
0	4.9167	4.2222	4.7200	4.1852	3.0000	2.3429	2.3793	1.6304
	3.5792	3.2318	2.8798	3.0260	2.2697	1.7480	1.9533	1.4354
	12	9	25	27	34	35	29	46
1	6.2727	5.4500	4.7500	4.2414	3.6154	3.1860	2.1364	2.1429
	1.9022	2.9996	2.3266	2.5864	2.1351	2.3529	2.1630	2.3237
	11	20	24	29	39	43	44	70
T	5.5652	5.0690	4.7347	4.2143	3.3288	2.8077	2.2329	1.9397
	2.9206	3.0698	2.5963	2.7812	2.2051	2.1324	2.0718	2.0272
	23	29	49	56	73	78	73	116
	14	15	16	17	18	ALL		
0	1.3250	1.2000	0.9756	1.5152	0.5000	2.1048		
	1.8403	1.4194	1.5888	1.3020	0.7559	2.3060		
	80	55	41	33	8	434		
1	1.1475	1.1250	1.0500	0.7895	0.5000	2.4605		
	1.4004	1.3810	1.3195	1.0842	0.7559	2.4759		
	61	40	40	19	8	448		
T	1.2482	1.1684	1.0123	1.2500	0.5000	2.2860		
	1.6611	1.3964	1.4534	1.2661	0.7303	2.3994		
	141	95	81	52	16	882		

DAB MEAN SCORES, STANDARD DEVIATIONS AND CELL SIZES

CONTROL: total sample

ROWS: sex COLUMNS: age (0=m; 1=f)

	6	7	8	9	10	11	12	13
0	7.833 2.368 12	9.222 3.232 9	10.040 2.389 25	10.815 1.711 27	12.529 2.863 34	12.771 2.745 35	14.345 3.120 29	15.174 2.924 46
1	7.818 1.537 11	8.300 2.003 20	9.208 1.654 24	9.517 1.975 29	10.282 2.470 39	10.837 2.419 43	11.614 2.982 44	10.500 2.541 70
T	7.826 1.969 23	8.586 2.428 29	9.633 2.531 49	10.143 1.949 56	11.329 2.873 73	11.705 2.731 78	12.699 3.303 73	12.353 3.534 116
	14	15	16	17	18	ALL		
0	15.163 3.332 80	16.291 2.699 55	15.293 3.052 41	16.152 2.647 33	16.125 2.588 8	14.105 3.615 434		
1	11.852 2.250 61	12.975 3.068 40	11.975 2.190 40	11.158 1.214 19	11.000 1.195 8	10.934 2.699 448		
T	13.730 3.338 141	14.895 3.286 95	13.654 3.127 81	14.327 3.288 52	13.563 3.286 16	12.489 3.553 882		

BG MEAN SCORES, STANDARD DEVIATIONS AND CELL SIZES

ROWS: age COLUMNS: garea (1=rural; 2=urban)

	1	2	ALL
6	8.1667 2.0412 6	4.6471 2.6444 17	5.5652 2.9206 23
7	6.1538 2.9957 13	4.1875 2.9262 16	5.0690 3.0698 29
8	5.0256 2.6005 39	3.6000 2.3664 10	4.7347 2.5963 49
9	4.8000 2.8121 40	2.7500 2.1448 16	4.2143 2.7812 56
10	3.9500 2.3088 40	2.5758 1.8376 33	3.3288 2.2051 73
11	2.8800 2.0368 50	2.6786 2.3262 28	2.8077 2.1324 78
12	2.7179 2.1145 39	1.6765 1.9025 34	2.2329 2.0718 73
13	2.4500 1.9735 40	1.6711 2.0159 76	1.9397 2.0272 116
14	1.2813 1.7133 64	1.2208 1.6273 77	1.2482 1.6611 141
15	1.2955 1.4560 44	1.0588 1.3478 51	1.1684 1.3964 95
16	0.9714 1.5240 35	1.0435 1.4135 46	1.0123 1.4534 81
17	1.5000 1.1795 24	1.0357 1.3189 28	1.2500 1.2661 52
18	0.5000 0.5774 4	0.5000 0.7977 12	0.5000 0.7303 16

DAB MEAN SCORES, STANDARD DEVIATIONS AND CELL SIZES

ROWS: age COLUMNS: garea (1=rural; 2=urban)

	1	2	ALL
6	7.500 2.074 6	7.941 1.983 17	7.826 1.969 23
7	8.154 2.824 13	8.937 2.081 16	8.586 2.428 29
8	9.564 2.664 39	9.900 2.025 10	9.633 2.531 49
9	10.050 2.136 40	10.375 1.408 16	10.143 1.949 56
10	11.250 2.488 40	11.424 3.317 33	11.329 2.873 73
11	11.960 2.934 50	11.250 2.303 28	11.705 2.731 78
12	13.641 3.638 39	11.618 2.511 34	12.699 3.303 73
13	14.100 3.334 40	11.434 3.300 76	12.353 3.534 116
14	14.047 3.302 64	13.468 3.366 77	13.730 3.338 141
15	15.500 2.741 44	14.373 3.638 51	14.895 3.286 95
16	15.057 3.298 35	12.587 2.544 46	13.654 3.127 81
17	14.792 3.243 24	13.929 3.333 28	14.327 3.288 52
18	17.750 2.217 4	12.167 2.209 12	13.563 3.286 16