DOCUMENT RESUME

ED 230 620 TM 830 465

AUTHOR Pfeiffer, Steven, I.; Naglieri, Jack A.
TITLE The Effects of a Group Supportive Testing Procedure and Acculturation Level on Achievement Test Scores of Culturally Diverse Students. Final Report, Revision.
INSTITUTION Northern Arizona Univ., Flagstaff.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Mar 83
GRANT NIE-G-81-0079
NOTE 19p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Acculturation; *Achievement Tests; *Cultural Differences; *Group Testing; Intermediate Grades; Mexican Americans; *Minority Group Children; Profiles; Reading Tests; Scores; Standardized Tests; *Test Bias; White Students
IDENTIFIERS *Arizona (Phoenix)

ABSTRACT

A group supportive testing procedure and acculturation level are investigated as to their influence upon achievement test scores of 150 culturally diverse inner city school students in grades four through six. The group supportive testing procedure was specifically designed as a facilitation strategy that attempted to mirror the more typical interaction style of minority children. The multidimensional measure of acculturation sought to elucidate whether students with particular acculturation profiles would present a particular test-taking performance. The experimental facilitative procedure employed did not improve either the Anglo or Mexican American children's test performance. The authors state (1) that minority group children should be administered group achievement tests under standard administration procedures, and (2) a student's sociocultural status, at least socioeconomic status and family size, appears to be a potentially relevant psychological construct that warrants further attention. (PN)

FINAL REPORT

THE EFFECTS OF A GROUP SUPPORTIVE TESTING PROCEDURE AND ACCULTURATION LEVEL ON ACHIEVEMENT TEST SCORES OF CULTURALLY DIVERSE STUDENTS

Grant No. N.I.E.- G-81-0079

Investigators:

Dr. Steven I. Pfeiffer, Senior Investigator
Dr. Jack A. Naglieri, Co-Investigator

Submission Date:

March, 1983

FINAL REPORT

ABSTRACT OF RESEARCH PROJECT

Introduction

Policy planners, educators, and test developers are all faced with increasing reports of the misuse of standardized tests. Mercer (1971) first pointed out that a disproportionately higher number of Black and Mexican American students were being labeled as mentally retarded, based solely on low intelligence test scores that have been shown to be biased against minority children. It is logical to assume that a similar phenomenon may be occurring among minority children on achievement tests (i.e., standardized achievement tests may not be accurately measuring culturally diverse students' full knowledge of academic material, due to biasing influences).

Klineberg (1935) pioneered exploring extra-test factors that affect attainment level, and found that socioeconomic status (SES), language, amount of schooling, and motivation all influenced Black children's test scores. Cole and Bruner (1971) carefully reviewed a tremendously large number of studies on minority group test performance, and concluded that the evidence does not support a deficit model. Rather, they hypothesized that the tests themselves, and the testing procedures employed, may not be permitting elicitation of minority children's full potential. Bernal (1977) stated that there is "something about the test(s) or testing situation which affects how minority subjects perform," and suggested that language proficiency, SES, testwiseness, motivation, and degree of acculturation all impact upon the testee.

Pfeiffer (1982) most recently reaffirmed that all higher psychological processes are shaped by our culturally organized experiences. By implication, we can postulate that achievement tests are inevitable cultural devices, influenced not only by their psychometric qualities, but also by the individual's particular attitudes, values, perceptions, motivation level, and distinctive sociocultural-behavioral patterns.

A number of writers have speculated that minority group children are forced to perform on standardized tests in a manner inconsistent with their sociocultural experience (e.g., Bernal, 1977; Cole & Bruner, 1971; Gay & Abrahams, 1973). Gay and Abrahams (1973) contend that "it is not the questions per se which cause Black children difficulty; it is the testing environment in general and especially the techniques that are used to ask the questions." These

writers hypothesize that the "Anglo-oriented" testing situation does not match the typical interactional style of minority cultural groups, where informality and cooperation dominate. Ramirez, Castanedo and Herold (1974) add empirical support for distinct interactional styles among culturally diverse children, showing that Mexican American families are characterized by child rearing practices which foster a field dependent socialization style.

Recently, researchers have begun to explore differences not only between a particular ethnic group and the Anglo core (or host) culture, but also to look at differences within the ethnic groups. The belief that ethnic groups are heterogeneous, and that labeling a person as simply Anglo, Black, Hispanic, or Native American oversimplifies the rich and varied sociocultural fabric existing among all individuals, led Pfeiffer and Naglieri (1980) to develop a multidimensional measure of acculturation. Their instrument (PNA) plots the individual on a number of dimensions which include language preference and usage, cultural identification and preference, ethnicity, interethnic distance and interaction, SES, family structure and size, and urban assimilation. As Berry (1980), Olmedo (1979), Olmedo and Padilla (1978), Padilla (1980), and Pfeiffer and Naglieri (1980) all point out, acculturative linguistic, psychological, and sociocultural factors must all be considered in understanding how the person fits both within his own and the host culture. And this understanding of each individual's acculturation level should help explain how culture differentially influences scholastic and test-taking performance.

Purpose of the Study

The purpose of this study was twofold: to investigate whether a group supportive testing procedure and acculturation level will influence the achievement test scores of culturally diverse students. The group supportive testing procedure was specifically designed as a facilitation strategy that attempted to mirror the more typical interaction style of minority children. And the multidimensional measure of acculturation sought to elucidate whether students with particular acculturation profiles would present a particular test-taking performance.

If a specific ethnic group, or a group of individuals within any one ethnic group which presents a particular acculturation profile, score significantly higher with the group supportive testing procedure, then one can conclude that the standard testing procedure does not permit an equitable demonstration of each student's maximal knowledge.

RESEARCH SITE

This research project was conducted during April, 1982. Wilson School District and Murphy School District, both inner-city school systems in the Phoenix area, were selected as the two sites of the research. Approximately 150 students, grades fourth through sixth, participated in the project.

TEST RESULTS

As discussed earlier, this project explored two main research questions: 1) whether a group supportive testing procedure would facilitate the achievement test performance of culturally diverse elementary grade students, and 2) if acculturation status plays a significant role in a student's achievement test score.

The numbers of children of each racial group and grade are presented in Table 1. Examination of these data reveals that there were similar numbers of children of each racial group in each grade. A nonsignificant ($p > 0.10$) chi square of 7.82 (4 df) was obtained indicating that the three grades could be combined for subsequent data analyses. Since only five black students participated in the study, these children were excluded from subsequent analyses.

A description of the Mexican American and Anglo groups is presented in Table 2. Examination of this table reveals that the Anglo and Mexican American samples were quite similar with respect to age. There was only a 1.9 month difference between the means, the SDs are very similar, as were the ranges. The ratio of males to females was about 40/60 in the Anglo and 50/50 in the Mexican American samples. The greatest difference between the samples appears to be in the schools attended. Nearly 50% of the Mexican American sample came from school #4 while none of the Anglo sample attended this school. However, due to the similarity of schools within the district, this disparity is viewed as having little impact.

Table 3 presents reading vocabulary scores for the two groups under the standard and experimental conditions. Examination of these scores reveals that the Mexican American students scored somewhat lower than the Anglo students in the standard reading vocabulary test condition, while they outperformed the Anglos in the experimental condition--however, both group mean scores were lower under the group facilitative experimental condition. Repeated measures ANOVA results yielded a significnt F ratio for group ($F=11.75$, $p < 0.001$) and group by condition interaction ($F=18.41$, $p < 0.001$). Analyses of simple main effects were used to compare the four mean scores. These analyses uncovered only two comparisons that were significant: a) the Mexican American group

mean was significantly higher than the Anglo students' mean score during the experimental condition ($F=78.91$, $p<0.001$) and b) the Anglo students' mean score was significantly higher under the standard than experimental administration procedure ($F=32.70$), $p<0.001$). Hence, these analyses indicate that the Anglo students' mean performance decreased significantly from the standard to experimental administrations, while the Mexican American childrens' mean decline was slight and nonsignificant. Additionally, the Anglo group mean was significantly lower than the Mexican American group mean under the facilitative condition.

The data presented in Table 3 yield the following conclusions: 1) neither group benefitted from the group facilitative (experimental) administration; 2) the Mexican American children performed much better than the Anglo children in the experimental condition, but about equal in the Standard condition; 3) both the Anglo and Mexican American samples evidenced substantially greater variability in their scores under the experimental procedure (note SDs two to three times larger in the experimental as opposed to standard conditions).

The increased variability in the distribution of reading vocabulary standard and experimental standard scores for the Anglo and Mexican American groups is graphically illustrated in Figures 1 and 2, respectively. Examination of Figure 1 reveals that the distribution of reading vocabulary scores under the standard condition was rather normal in shape (skewness= 0.24; kurtosis= -0.084) while the distribution of reading vocabulary scores under the experimental condition was not normal (skewness= -0.43; kurtosis= -1.6). As is evident by examination of Figure 1, the distribution of scores in the experimental condition is rather bi-modal. Moreover, while a considerable difference between the Anglo group reading vocabulary experimental and standard conditions means exists (461, 335), there is a less dramatic difference between the medians (467, 423), suggesting that the difference was greatly influenced by some very low scores.

Figure 2 graphically represents the distribution of reading vocabulary scores under both conditions for the Mexican American group. Similar to the Anglo group's distribution, the scores under the standard condition were rather evently distributed (skewness= 0.35; kurtosis= -0.42) while under the experimental condition the scores were more bi-modal (skewness= -2.21; kurtosis= 5.55). Additionally, differences between mean and median scores for this sample reveal a similar pattern seen in the Anglo sample. The Mexican American group means differed by about 15 points but the medians did not--they were identical (452). Once again, this illustrates that the interpretation of the difference in mean scores is somewhat misleading due to the influence of a few very low scores.

Examination of the percentage of children who scored extremely low under the experimental (group-facilitative) condition revealed that the Anglo and Mexican American samples differed in this respect. That is, 7% of the Mexican American sample as opposed to 36% of the Anglo sample earned scores less than 300 in the experimental condition. It is this shifting of scores which is responsible for the differences in means versus medians, the large standard deviation in the Mexican American, and the very large standard deviation for the Anglo group's reading vocabulary experimental conditions.

Examination of the correlational relationship between the reading vocabulary standard and experimental conditions with the California total reading test standard scores was examined for the group as a whole (n=138) to further understand what the reading vocabulary tests in the standard and experimental conditions were reflecting. The reading vocabulary (standard) correlated 0.8575 ($p<0.001$) with the total reading score while the experimental reading vocabulary scores correlated 0.163 (p 0.05) with total reading. (Reading vocabulary standard and experimental conditions correlated 0.018, $p<0.10$). The significance of the difference between the reading vocabulary experimental/total reading and reading vocabulary standard/total reading coefficients was tested using a t-test for the difference between correlated correlations. This test yielded a t valve of 11.69 ($p<0.001$) which indicates that the difference reflected a real disparity in correlations. Hence, the reading vocabulary experimental condition was influenced by some variable other than overall reading proficiency in contrast to the standard reading vocabulary condition, which correlated as expected with total reading. Similarly, the reading vocabulary (standard) correlated significantly with the total grammar standard score (0.511, $p<0.0001$) while the reading vocabulary experimental test condition correlated 0.083 ($p>0.10$) with grammar total. The differnce between these correlation coefficients was also significant ($t=4.13$, $p<0.001$). These findings suggest that the reading vocabulary test administered under the standard conditions was essentially functioning as intended, to measure English vocabulary proficiency. In contrast, the reading vocabulary test administered under the experimental condition did not correlate nearly as well with the reading total or grammar total as such a test would be expected to, and is likely reflecting some other variable other than vocabulary skill.

Correlations among the System of Multicultural Pluralistic Assessment (SOMPA) acculturation scales and the Reading Vocabulary Scores earned under Standard and Experimental Scores earned under Standard and Experimental conditions are presented in Table 4. Examination of these coefficients reveals that the SOMPA sociocultural subtests designed

to assess acculturation status did not correlate significantly with the reading vocabulary scores under the standard condition and only two of the four variables (socioeconomic status and family size) correlated significantly under the experimental condition. the only significant correlations were reading vocabulary experimental with family size and socioeconomic status. Since family size and socioeconomic status correlated 0.38 ($p<0.0001$), these two subtest scores (variables) should probably be viewed as reflecting a strong relationship between the reading vocabulary experimental scores with general family socioeconomic conditions. This suggests that performance on the Reading Vocabulary experimental task was significantly related to a child's level of family social and economic status. Additionally, the lack of significant correlations between the SOMPA subtests and reading vocabulary scores obtained under the Standard condition suggests that a child's socioeconomic level was differentially related to achievement test scores.

The fact that the group facilitative (experimental) administration procedure did not enhance the achievement test scores of the Mexican American children can be explained in a number of possible ways. One, the group facilitative procedure may not have accurately or realistically mirrored the more typical interaction style of culturally different children. Two, the experimental condition may not have been a powerful enough "facilitative intervention strategy" to create the testing ambience that Bernal (1983) suggests leads to "improved performance of minority students in (with) an appropriate facilitative intervention virtually eliminating initial racial and ethnic differences" (p. 9-10).

A third possible explanation for the ineffectiveness of the experimental condition is that a single "score-enhancing technique" may not be significant enough alone to produce higher average achievement test scores. Jensen's (1980) exhaustive review of the research literature on external sources of test bias (i.e., test sophistication, motivation, attitude, personality, expectations, dialect, race and sex of the examiner, individual-vs.-group administration, and times vs. untimed) concluded that "no variables in the test situation, but extraneous to the tests, have been identified that contribute significantly to the observed average test score differences between social classes and racial groups" (p 618). It may be that future research will have to look at test bias in a more complex, interactive manner; with researchers modifying conjointly a number of hypothesized facilitative factors such as small group administration, rapport building between examiner and children, practice, feedback to children regarding the accuracy of their responses, and a more relaxed interactional testing procedure. It may take three or four of these factors in concert to produce significant improvement in achievement test scores of minority group children.

The present research offers two clear implications for educational practice. First, minority group children should be administered group achievement tests under standard administration procedures. The experimental facilitative procedure employed in the present study did not improve either the Anglo or Mexican American children's test performance, and any modification in standard administration should be conducted in a highly tentative, cautious, and judicious manner.

Second, a student's sociocultural status, at least socioeconomic status and family size, appears to be a potentially relevant psychological construct that warrants more attention by both researchers and practitioners. Future refinements in the measurement of sociocultural status would appear to help educators better understand a minority group student's attitude, motivation, expectancy, and level of anxiety when approaching standardized achievement (and aptitude) tests.

REFERENCES

Bernal Jr., E.M. Assessment procedures for Chicano children: The sad state of the art. International Journal of Chicano Studies Research, 1977, 8, 69-81.

Bernal Jr., E.M. Bias in mental testing: Evidence for an alternative to the heredity-environment controversy. Chapter submitted for publication, 1983.

Berry, J.W. Acculturation as varieties of adaptation. In Padilla, A.(Ed.) Acculturation: Theory, models and some new findings. Boulder, Colorado: Westview Press, 1980.

California Achievement Tests. Monterey, California: McGraw Hill, 1978.

Cole, M., & Bruner, J.S. Cultural differences and inferences about psychological processes. American Psychologist, 1971, 26, 867-876.

Gay, Geneva, & Abrahams, R.D. Does the pot melt, boil, or brew? Black children and white assessment procedures. Journal of School Psychology, 1973, 11, 330-340.

Jensen, A.R. Bias in Mental Testing. New York: Free Press, 1980.

Klineberg, O. Race differences. New York: Harper, 1935.

Mercer, J.R. Sociocultural factors in labeling mental retardates. Peabody Journal of Education, 1971, 48, 188-203.

Olmedo, E.L. Acculturation: A psychometric perspective. American Psychologist, 1979, 34, 1061-1070.

Olmedo, E.L., & Padilla, A.M. Empirical and construct validation of a measure of acculturation for Mexican Americans. Journal of Social Psychology, 1978, 105, 179-181.

Padilla, A.M. (Ed.) Acculturation: Theory, models and some new findings. Boulder, Colorado: Westview Press, 1980.

Pfeiffer, S.I. A cross-cultural note on the superiority of team decision-making. Exceptional Children, 1982, 49 (1), 68-69.

Pfeiffer, S.I., & Naglieri, J. The development of a measure of acculturation. Unpublished manuscript, Northern Arizona University, 1980.

Ramirez, M., Castanedo, A., & Herold, P.L. The relationship of acculturation to cognitive style among Mexican Americans. Journal of Cross-Cultural Psychology, 1974, 5, 424-433.

TABLE 1

Number of Children of Each Racial Group and Grade (N=138)

Race	Grade			Row Total
	4	5	6	
White	N=13	N=19	N=7	N=39
Hispanic	N=41	N=24	N=29	N=94
Black	N= 3	N= 1	N= 1	N=5
Column total	N=57	N=44	N=37	Total N=138

TABLE 2

Descriptive Statistics of Anglo and Mexican American Samples

Variable		Anglo (N=39)	Mexican American (N=94)
Age in months	$\bar{x}$	132.8	134.7
	SD	11.7	12.1
	Range	112-161	112-172
Sex	Males	38.5%	47.9%
	Females	61.5%	52.1%
School attended	#1	76.9%	30.9%
	#2	17.9%	14.9%
	#3	5.1%	16.0%
	#4	0.0%	47.9%

TABLE 3

Means and SDs for Anglo and Mexican American Groups

Groups	Reading Vocabulary Conditions	
	Standard	Experimental
Mexican American (N=94)		
Mean	451.8	436.4
SD	49.5	100.6
Anglo (N=39		
Mean	461.5	335.4
SD	48.3	179.5

TABLE 4

Correlations Among the SOMPA

and Reading Vocabulary Scores

SOMPA Subtests	Reading Vocabulary	
	Standard Condition	Experimental Condition
Family size	0.01	0.22*
Family structure	0.10	-0.08
Socioeconomic status	0.02	0.23*
Urbal acculturation	0.02	0.15

*Correlations are significant at an experiment-wise error rate of 0.05 $\left(\frac{0.05}{8} = 0.006\right.$ level of probability need for significance of the Pearson product-moment correlation$\left.\right)$.

Figure 1

Distribution of Anglo Reading Vocabulary Standard and Experimental Conditions (N=39)

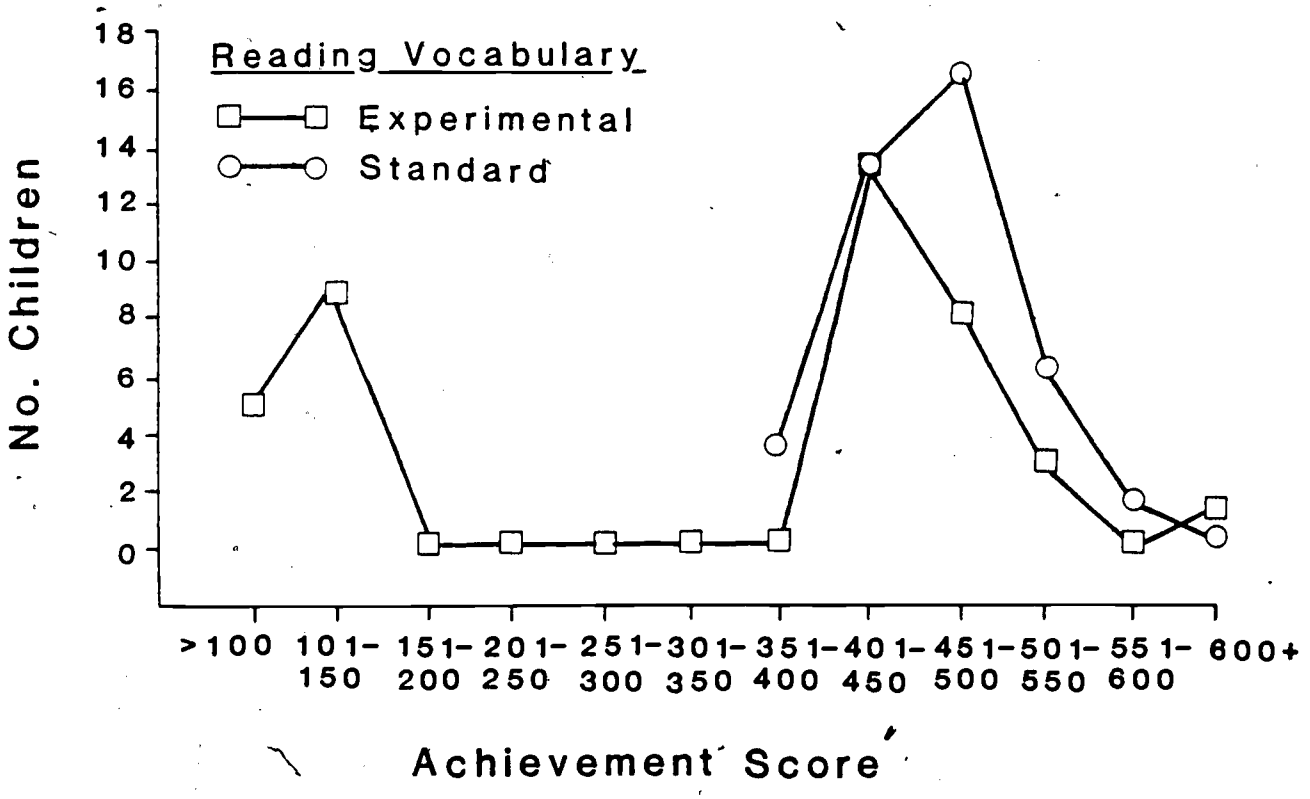

Figure 2

Distribution of Mexican American Reading Vocabulary

Standard and Experimental Conditions (N=94)

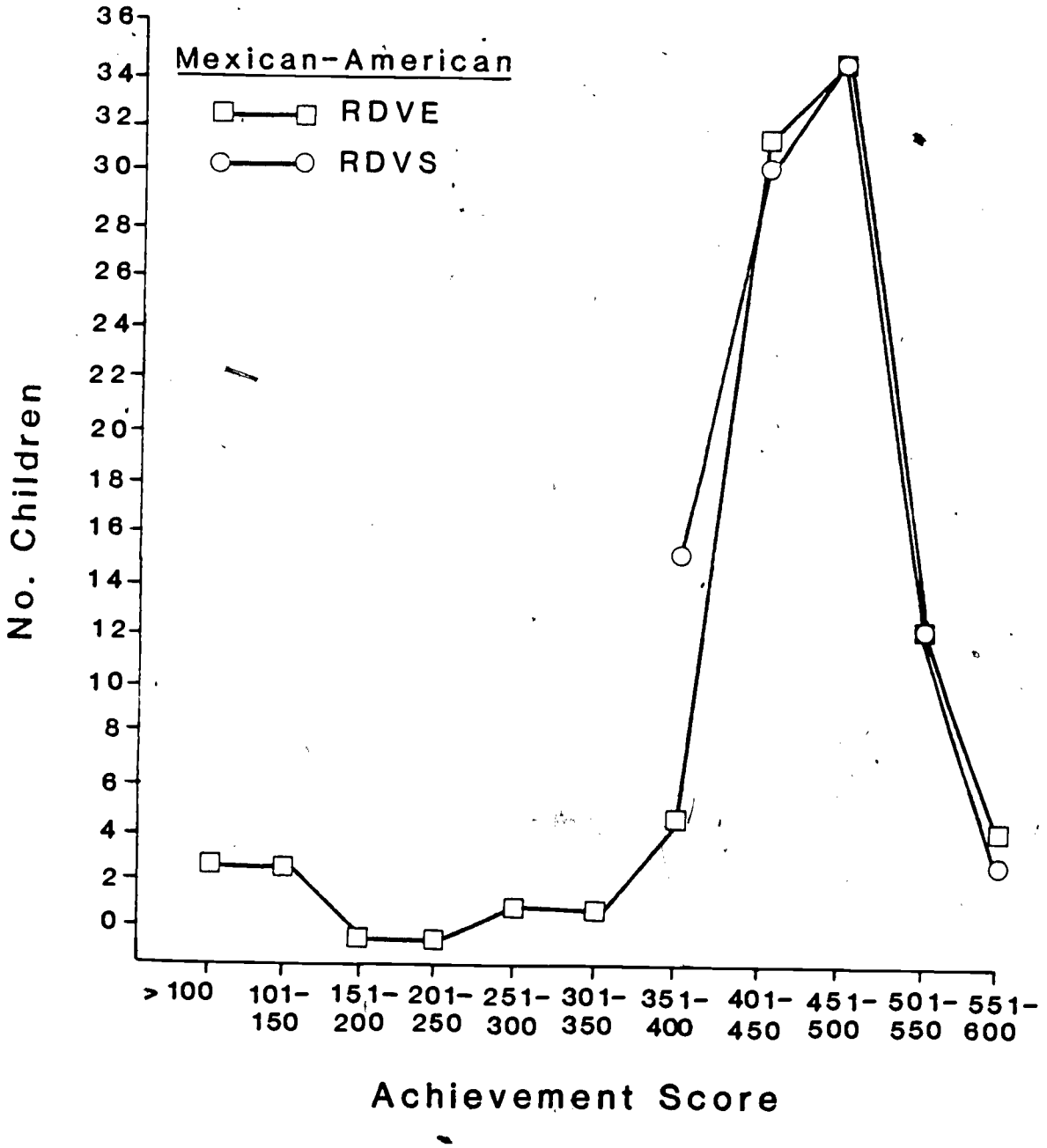

Figure 3

Administration Directions:

Each group of student's is to work independently on the Reading Vocabulary (RV) subtest of the California Achievement Test (CAT). Within each group, one student will be given the task of answering each of the RV questions. The remaining two students will be encouraged to support, encourage, and help this student complete the task, NOT choose the correct answer. At no time are either of the two "supportive" students to influence the selection of the correct answer. Comments such as, "You're doing good," or "Great job," are appropriate; not "B is the answer."

Directions to Students:

SAY: "Please form into groups of three.

"What we are about to do is an important part of an experiment being conducted by Northern Arizona University. Please follow these directions carefully. In each group of students, one person has been chosen to answer 30 vocabulary questions. I will hand these out in a few moments. This person is to answer each question completely on his or her own. The other two students are NOT allowed to help this person in choosing the correct answer. Instead, the two other students should encourage and support the student answering the questions. So, you can say, 'You're doing good,' or 'Good job,' or [Reader may interject appropriate, perhaps humorous, words of encouragement.], but not 'B is the correct answer.' [Pause.] Any questions?

"Now, here is the answer sheet. Before we begin, please fill in your name, grade, and teacher's name; please write clearly on your answer sheet.

"I will read the directions; please follow along. Follow directions through the sample item C.

"If there are no questions, we will begin.

"Remember, you can help the student completing the test by encouraging him, but DO NOT TELL THE ANSWER."

Follow directions on page 22 -- START.

Figure 4

Child Acculturation Scale

(1) Name _____ ID Number ___ ___ ___

(2) Ethnic Background:

Anglo/White _____

Mexican American _____

Black _____

Native American _____

Other _____

(3) Sex: Male _____ Female _____

(4) How religious would you say your family is?

1	2	3	4	5
not religious at all		moderately religious		very religious

(5) How important do your parent(s) think doing well in school is?

1	2	3	4	5
not very important		moderately important		very important

(6) How much do you care about getting good grades in school?

1	2	3	4	5
not at all		somewhat		a great deal

(7) How much schooling would you like to complete?

Junior high _____

High school _____

Vocational/Technical _____

College _____

Graduate school (doctor, lawyer, PhD, etc.) _____

(8) How much schooling do you think your parent(s) want you to have?

Junior high _____

High school _____

Vocational/Technical _____

College _____

Graduate school (doctor, lawyer, PhD, etc.) _____

(9) What language is spoken at home by your family?

English only _____

Mostly English _____

Spanish only (Navajo, Hopi, etc.) _____

Mostly Spanish (Navajo, Hopit, etc.) _____

(10) How would you like the following things in your life to be?

	Completely Mexican-American (Native American)	Mostly Mexican-American (Native American)	Both Mexican-American (Native American) and American	Mostly American	Completely American
Food	1	2	3	4	5
Language	1	2	3	4	5
Music	1	2	3	4	5
T.V. Programs	1	2	3	4	5
Books/Magazines	1	2	3	4	5
Radio Programs	1	2	3	4	5
Ways of Celebrating Holidays and Special Events	1	2	3	4	5