DOCUMENT RESUME

ED 275 414 PS 016 115

AUTHOR Stanley, William B.; And Others
TITLE Assessing Young Children's Social Concept Development.
SPONS AGENCY Louisiana State Univ., Baton Rouge. Coll. of Education.
PUB DATE [85]
NOTE 39p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Age Differences; Blacks; *Concept Formation; Difficulty Level; Elementary School Curriculum; *Elementary School Students; Grade 1; *Individual Differences; Kindergarten Children; Primary Education; Racial Differences; Reliability; Sex Differences; *Social Cognition; Social Studies; Validity; Whites

ABSTRACT

This study investigated a number of questions regarding the nature of social concept development in young children. Subjects were 64 kindergarten children and 65 first grade public school students from lower to upper middle class socioeconomic levels, of whom 66 were male, 63 were female, 78 were Caucasian, and 51 were black. Two assessment instruments were used: (1) a basic concepts assessment task involving simple object sorting, class extension, and cross classification; and (2) a social concepts assessment task designed to measure subjects' knowledge of nine basic social concepts (young/old, urban/rural, family/not family, past/present, rich/poor, war/peace, groups/individuals, houses/other buildings, and people who protect us/other people). Profile analysis was used to compare grade, sex, and racial groups. All three variables had a significant impact on performance. Significant differences in difficulty were found among the nine concepts measured. Three of the most difficult concepts (family/not family, those who protect us/other people, and past/present) are commonly included in the early childhood curriculum. These results suggest the need for consideration of concept development level in planning the social studies curriculum and instruction for young children. (Author/RH)

Assessing Young Children's Social Concept Development

College of Education

William B. Stanley
Rosalind Charlesworth

College of Business Administration

Stephen Looney
Jeffrey Ringuest

Louisiana State University

Running head: ASSESSING SOCIAL CONCEPTS

Footnote

[1] Copies of instructions and coding procedures can be obtained from Charlesworth or Stanley.

Author Notes

This research was partially funded by grants from the Louisiana State University Colleges of Education and Business Administration and the Louisiana State University Council on Research. Appreciation to the East Baton Rouge Parish Public Schools and to Rosemary Brumfield, Sheila Murrill, Tammy Horner, and Robyn Williams. Authors Stanley and Charlesworth's address: College of Education, Louisiana State University, Baton Rouge, LA 70803. Authors Looney and Ringuest's address: College of Business Administration, Louisiana State University, Baton Rouge, LA 70803.

Abstract

This study investigated a number of questions regarding the nature of social concept development in young children. In an earlier study a social concept picture sorting task was developed in order to obtain normative data on young children's social concept development. For this replication study a larger, more heterogeneous sample was used consisting of 64 kindergarten and 65 first grade public school students from lower to upper middle class socioeconomic levels. Profile analysis was used to compare grades, sex, and racial groups. All three variables had a significant impact on performance. Significant differences in difficulty were found among the nine concepts measured. Three of the most difficult concepts (Family/Not Family, Those Who Protect Us, and Past/Present) are commonly included in the early childhood curriculum. These results suggest the need for consideration of concept development level in planning social studies curriculum and instruction for young children.

Assessing Young Children's Social Concept Development

Introduction

An understanding of social studies, like all other subjects, depends on learning and applying its basic concepts. It has been documented that schooling can have a significant impact on a child's concept development (Cole & D'Andrade, 1982). Researchers (Jantz and Klawitter, 1985; McKinney, Larkins, Burts & Davis, 1982) have found that young children can learn some abstract Social Studies concepts if properly taught.

In a review of concept development research, Wellman (1982) concluded that we need more information regarding normative standards for the young child's concept development via instruments designed specifically for young children. For example, Koeller (1981), found there were no satisfactory instruments which measured both economic knowledge and reasoning ability of young children. An additional concern is that we still lack adequate knowledge of the kinds of concepts young children are capable of learning (Jantz and Klawitter, 1985).

Purpose of the Study

This study was designed to address some of the concerns expressed above and to extend the work started in an earlier study (Stanley, Charlesworth and Ringuest, 1985). In the previous study, the authors sought to assess young children's knowledge of nine basic social concepts. Performance was compared on the basis of age, sex and race. In addition, we attempted to determine the relative difficulty (or knowledge of) the nine concepts studied. The results of the previous study indicated that some of the nine concepts were significantly more difficult than others. Although the rankings changed slightly with grade level, certain subtasks were more difficult for both

groups. No significant difference was found in scores between sex or racial groups. While these findings were interesting, they were qualified by certain limitations.

Among the limitations of the previous study was a relatively small sample (n=52) of predominantly Caucasian students from a middle and upper middle class background. They also attended a non-public laboratory school environment with better than average facilities and instructional services. The present study sought to determine if these previous results would be replicated with a larger, more varied sample of subjects drawn from public schools. The 129 students selected for this study represented a wide range of socioeconomic backgrounds, a large minority (black) population, and they attended schools with typical instructional facilities and environments.

An additional concern raised by the first study was the adequacy of the statistical analysis. Profile analysis was employed to assess the results obtained in this study. This method appears to give a clearer indication of the results and is described in more detail below. In addition, a detailed analysis was made of the reliability and validity of the concept assessment instrument, an issue that was not addressed in the previous study. A full discussion of this issue is reported before the section on results.

Method

Subjects

The subjects in this study were 129 students in public elementary schools. There were 64 kindergarten and 65 first graders. Male students numbered 66 and female 63. There were 78 Caucasian students and 51 black students. The same age groups as in the first study were used because, in the Piagetian framework, it is the transition period when most children begin to

move from preoperational to concrete operational thought (Gardner, 1982). Subjects were randomly selected from those whose parents returned consent forms.

Assessment Tasks

Two different assessment instruments were used: A basic concepts assessment task using procedures developed by Brainerd (1979) and a social concepts assessment task developed by Charlesworth and Stanley (1982). Sorting was used as the major mode of response in both instruments, because concept learning in early childhood is dominated by the learning of basic categories (Bruner, Oliver, and Greenfield, 1966; Inhelder and Piaget, 1964; Rosch and Lloyd, 1978; and Isenberg and Jacobs, 1981).

The basic concepts assessment task consists of three activities: (1) simple object sorting (color and shape), (2) class extension (color, shape and number), and (3) cross classification (completing simple matrices). Each task was scored and a composite score (the Basic Concept Score or BCS) was obtained for each subject. This set of scores gives an estimate of the subject's developmental level and was used as a basis to help evaluate performance on the social concept instrument.

The second assessment instrument was designed to measure the subjects knowledge of nine basic social concepts. Earlier research in this area tended to rely mainly on a verbal interview method. When designing the instrument used in this study, we wanted to include activities that offer an opportunity for success to the child who may not be able to express answers in a strictly verbal interview approach. Thus, the social concept assessment task is composed of nine subtasks containing eight pictures each, four of which are examples of the concept to be identified in each sort.[1] The subject's

performance on each social concept subtask was then scored by assigning a numerical value between 0 and 10. The value was determined by the subject's degree of correctness of the sort and the logic and maturity of the subject's justification (Charlesworth, Stanley and Ringuest, 1985). In addition, a total social concept score (TSCS) was calculated by combining the scores of the subtasks. Thus we can compare the relative difficulty of each subtask as well as the general performance of various groups of subjects.

The nine social studies concepts tested were: Young-Old; Urban-Rural; Family-Not Family; Past-Present; Rich-Poor; War-Peace; Groups-Individuals; Houses-Other Buildings, and People Who Protect Us-Other People. These particular concepts were chosen for study for several reasons. First, each represents a basic social concept drawn from history and the social sciences. Second, we wanted to use concepts to which the subjects would have had some exposure prior to kindergarten entrance via the media, first hand experience, books, etc. or school experiences which focused on these concepts. Indeed, most of the concepts (e.g., family, urban-rural, groups, people who protect us, young-old, past-present) are typical of the sort of content found in the early childhood social studies curriculum (Seefeldt, 1984; Walsh, 1980). Finally, we needed concepts which could be presented in picture form, since many of our subjects had limited oral descriptive abilities and only beginning, if any, reading skills. Thus, while others may have chosen a different mix of social concepts, these nine seem to represent a reasonable group to assess.

Procedure

Assessment instruments were administered to individual subjects during a 35 to 45 minute session. One researcher interviewed a subject while the other recorded the subject's specific choices and verbal explanations for the choices. Following a warm-up task, the subjects were first given the basic concept assessment task in which each is asked to sort objects or pictures into groups of "things that belong together", (e.g., color, shape, or number).

Next, the subjects were given the nine social concept assessment tasks. There were two different forms of this task (Form A and B). Each had a different set of pictures to represent the nine concept subtasks. The subjects were randomly assigned to one of the two forms. In addition, the order of presentation of the nine tasks was randomized for each subject to control for order effect. The subject was first asked to sort spontaneously the eight pictures related to a subtask into two piles. If the subject failed to perform the sort successfully, he or she was asked to sort on the basis of a clue (e.g., "Put the pictures of old people here and the pictures of young people there"). Following each sort, the subject was asked to justify his or her choices (e.g., "Why do these pictures go together?"). This procedure was continued until all nine tasks were completed. As noted above, a score was calculated for the basic concept assessment task (i.e., the Basic Concept Score or BCS). Also, the subjects' scores on each social concept were computed and these were summed for a Total Social Concept Score or TSCS.

Reliability and Validity

In the previous study (Stanley, Charlesworth and Ringuest, 1985), a preliminary examination was made of the validity and reliability of the concept assessment instrument. For example, there was a significant

correlation between the BCS and TSCS, and both were positively correlated with age. Furthermore, the average TSCS for first grade subjects was significantly higher than that of kindergratern subjects. Finally, we found no significant difference between the results obtained using the two different forms of the social concept assessment task. Still, a more rigorous analysis of validity and reliability was desirable. Thus, before discussing the results, we turn to a discussion of the validity and reliability of the social concepts assessment task.

Validity

The two types of validity that we examined for the social concepts assessment task were content and construct validity. Criterion validity was not considered because there is no generally accepted measure of concept development that could be used as the criterion measure. As indicated by Nunnally (1978, p. 90), such a criterion is essential to any examination of criterion validity and, without it, criterion validity cannot be addressed.

While there is no agreed upon technique for assessing content validity (Carmines and Zeller, 1979, p. 22), an attempt was made to determine if the nine subtasks of the assessment task provide a representative sample of behaviors from the performance domain associated with concept development. The content validity of the task was judged to be adequate in this regard by a team of experts consisting of faculty members, teachers and graduate assistants. The instrument was also pilot tested with several young children. A few ambiguous items were noted and replaced with others that were clear examples or nonexamples of the concepts involved. As Nunnally suggests (1978, p. 93), such appeals to reason are often the only technique one has for ensuring content validity for an instrument.

The construct of concept development has received considerable attention in the education literature (e.g., Smith & Medin, 1981; Stanley and Mathews, 1985). The cognitive-structural theory underlying this construct suggests the following hypothesis:

Hypothesis 1: First grade students will have higher average scores than kindergarten students on the social concepts assessment task and on each of the nine subtasks that compose it.

In other words, first graders should perform higher on the average on any instrument that is designed to measure concept development. In order to address this issue, the two samples of students mentioned above were broken down by grade and compared on the basis of the average TSCS, using a two sample t test. The means and variances for each of these groups are given in Table 1. For the original sample of 52 students (24 kindergarten and 28 first graders), t (50) = -5.08, $p \leq .000006$ and for the later sample of 129 students (64 kindergarten and 65 first graders), t (127) = -4.47, $p \leq .00002$. These results are clearly consistent with Hypothesis 1. In addition, a profile analysis was performed for the two grades using the sample of 129 students and these results indicate that the kindergarten and first grade students are significantly different across all nine subtasks. (This analysis is discussed in detail in the Results section.) The conclusion to be drawn here is that the nine subtasks that make up the social concepts task measure the same thing; namely, the construct of concept development (e.g., see Nunnally 1978, p. 102).

Insert Table 1 about here

Another hypothesis suggested by the cognitive-structural theory is the

following:

Hypothesis 2: Scores on the social concepts assessment task and each of the nine subtasks that compose it will be correlated with measures of basic concept formation skills (e.g., sorting and classifying) such as Brainerd's (1979) sorting task.

Table 2 contains the values of the Pearson product moment correlation between the BCS and the TSCS, as well as the correlations between the BCAT score and each of the nine subtasks for both samples. Also included are the corresponding values of Spearman's coefficient. Even though these correlations are rather modest, they indicate that the performance of the social concepts task is consistent with Hypothesis 2. In short, the social concepts task appears to satisfy the conditions of construct validity since it produces results that are predictable from theoretical hypotheses concerning the construct (Nunnally, 1978, p. 98).

Insert Table 2 about here

To summarize, the social concept instrument appears to be valid for the purpose of obtaining baseline data and for making preliminary determinations of the level of concept development attained by kindergarten and first grade students.

Reliability

While the relative subjectivity involved in scoring the social concepts task suggests that reliability coefficients based on alternative forms should be used (Nunnally, 1978, p. 232), such procedures are not feasible for use with the instrument due to the high degree of learning that occurs after the

first administration of either form of the instrument. This learning would be especially prevalent in those students who require verbal cues in performing the sorting tasks. These students would probably retain the cues for use with the same sorting tasks when taking the alternative form of the test a short time later (usually two weeks), as required when the alternative form reliability coefficient is used. Therefore, we present coefficient alpha as a measure of reliability.

The reliability of the social concepts assessment task was examined for both samples of students. Table 3 contains a summary of the estimates for the complete samples and for the samples broken down by grade. Estimates of the standard errors of measurement based on these reliability values are also presented. These results indicate that the social concepts assessment task is a reasonably reliable instrument, even though it does not quite meet the suggested minimum acceptable criterion of .80 for widely used scales (Nunnally, 1978, p. 245; Carmines and Zeller, 1979, p. 49). However, since instruments for measuring concept development are still in the early stages of development, a more modest reliability of .70 is acceptable (Nunnally, 1978, p. 245) and the social concepts instrument does meet this criterion. Further research is called for in improving the reliability of the instrument before it can be recommended for general use.

Insert Table 3 about here

Results

Profile analysis (Morrison, 1976, pp. 153-160, 205-216) was used to test the hypotheses of interest for this study. The profile of a particular

student is defined to consist of the nine subtask scores, and hypotheses based on comparisons of certain groups of students can be stated in terms of the population mean profiles of the groups. In particular, the following three hypotheses can be tested using profile analysis:

H_{01}: The population mean profiles are parallel, i.e., the line segments joining the population subtask means are parallel across the groups. This is equivalent to the hypothesis of no subtask x group interaction.

H_{02}: In addition to the population profiles being parallel, they are also coincident, i.e., each population subtask mean is the same across the groups. This is equivalent to the hypothesis of no group effect.

H_{03}: In addition to the population profiles being parallel, they are also level, i.e., the population means are the same across the nine subtasks. This is equivalent to the hypothesis of no subtask effect.

These hypotheses correspond to the ones that are usually tested in a repeated measures design with one repeated factor (subtask) and one nonrepeated factor (group). McCall and Appelbaum (1973) contains a comprehensive review of techniques for testing H_{01}, H_{02}, and H_{03}.

The sample profiles for each of the two grades, two race groups, and two sex groups in the present study are given in Figures 1(a), 1(b), and 1(c),

---

Insert Figures 1(a), (b) & (c) about here

---

respectively. The average score on each subtask is plotted separately for each group and then the points are connected to form a graphical

representation of the average profile for the group. These plots indicate that the population profiles for grade and race are parallel, whereas the profiles for sex are not, but more precise testing is required.

Following the recommendation of Rouanet and Lepine (1970), we test H_{01} using both multivariate and univariate approaches. For the univariate tests, we use the three-step procedure of Greenhouse and Geisser (1959, p. 110) in deciding whether to accept or reject a given hypothesis; namely, if the conservative F test is significant, then the other univariate F tests will also be significant and the hypothesis should be rejected. On the other hand, if the unadjusted F test is not significant, then none of the other F tests will be significant either and the hypothesis should be accepted. If the conservative F test is not significant and the unadjusted F test is significant, then the results of the ε- and ε- adjusted F tests are compared (Maxwell & Arvey 1982). If these two tests agree, then the common conclusion is appropriate. If they disagree, then the average of the adjusted degrees of freedom for the two procedures can be used. Once a decision has been made using the univariate techniques, the results of the Hotelling's T^2 test of H_{01} are used for corroboration.

(We note that both assumptions that underlie repeated measures testing procedures [multivariate normality of the responses and equality of covariance matrices across the groups] appear to be violated for the data in this study. However, recent results [Mendoza, Toothaker, & Nicewander, 1974; Rogan, Keselman, & Mendoza, 1979] indicate that these procedures are sufficiently robust against violations of these assumptions. Typically, these violations tend to increase slightly the Type I error rate of the tests; therefore, lower p-values are required to declare a result significant.)

We first consider the comparison of the two grades. The unadjusted univariate F test of H_{01} yields $\underline{F}(8, 1016) = 1.18$, $\underline{p} \leq .310$, thereby indicating acceptance of the parallelism hypothesis. Furthermore, Hotelling's T^2 yields $\underline{F}(8, 120) = 1.26$, $\underline{p} \leq .269$, which is also an indication that H_{01} should be accepted. Thus, we assume that the profiles for the two grades are parallel and proceed to the test of H_{02}. This hypothesis is tested using a two-sample t test on the composite score of the nine subtasks (Morrison, 1976, p. 156). This test yields $\underline{t}(127) = -4.47$, $\underline{p} \leq .0002$, which is extremely strong evidence against the hypothesis of coincident profiles. Therefore, based on the acceptance of H_{01} and the rejection of H_{02}, it can be concluded that there is a significant difference between the two grades on each of the nine subtasks. The results for the two grades are summarized in the lower half of Table 1.

The three-step univariate procedure was also applied in testing H_{03}. The conservative F test yielded $\underline{F}(1, 127) = 90.48$, $\underline{p} \leq 2 \times 10^{-16}$, overwhelming evidence that H_{03} should be rejected. Hotelling's T^2 was also used to test H_{03}, yielding $\underline{F}(8, 120) = 60.43$, $\underline{p} \leq 2 \times 10^{-16}$. Therefore, both the univariate and multivariate approaches indicate that there are significant differences among the nine subtasks. To determine where these differences lie, union-intersection pairwise comparisons (Roy & Bose, 1953) were performed for every possible pair of subtasks using a familywise error rate of .10. These comparisons indicate that there are several clusters of questions that are not significantly different in terms of difficulty, as indicated in Table 4.

Insert Table 4 about here

Those questions that are of equal difficulty are connected by an unbroken line in the lower portion of the Table. For example, while subtasks 1 and 8 are of equal difficulty, they are both less difficult than subtasks 3, 4, 5, and 6 and more difficult than subtasks 2 and 9.

For the comparison of the two race groups, we conclude that the parallelism hypothesis H_{01} should be rejected: the unadjusted F and conservative F tests disagree, but the ε- and ε- adjusted tests are highly significant: $\underline{F}(6.6, 842) = 3.16$, $\underline{p} \leq .003$, and $\underline{F}(7.1, 900.3) = 3.16$, $\underline{p} \leq .003$, respectively. Hotelling's T^2 also indicates rejection of H_{01}, with $\underline{F}(8, 120) = 3.02$, $\underline{p} \leq .004$. Thus, we conclude that the population profiles are not parallel for the two race groups. Since the two profiles are not parallel, it is not appropriate to test H_{02} in the form given above (Morrison, 1976, p. 157). Instead, as Morrison suggests, the hypothesis of equal group means can be tested separately for each subtask using the two-sample t test, and the hypothesis of equal subtask means can be tested separately within each group using the univariate and multivariate repeated measures techniques described above for testing H_{03}. We suggest using a protected significance level of $\alpha/(p + g)$ for each of these tests, in order to adequately control the familywise error rate. An alternative to testing for equality of group means separately for each subtask is to calculate simultaneous Bonferroni intervals for each pairwise difference between the groups for each subtask. Both procedures were employed in analyzing the data for this study.

The results of the tests of equal means for the two race groups for each subtask are given in Table 5, along with 90% Bonferroni intervals for each

---

Insert Table 5 about here

---

pairwise difference. These results indicate a significant difference between the race groups only for subtasks 3, 4, and 7, using a protected significance level of $\alpha/(p + g) = .10/(9 + 2) = .009$. The tests of equal subtask means within each group are both extremely significant: $F(8, 70) = 26.60$, $p \leq 2 \times 10^{-18}$ for Caucasians and $F(8, 43) = 42.04$, $p \leq 7 \times 10^{-18}$ for blacks. Union-intersection multiple comparisons with a familywise error rate of .10 were performed separately for each race group, yielding the results shown in Table 6. The clusters of subtasks for the two race groups agree, except that

---

Insert Table 6 about here

---

subtask 1 is significantly different from subtasks 2 and 7 for the Caucasian students, but not for the black students.

Finally, we compared the profiles for the two sex groups. The unadjusted and conservative F tests for H_{01} disagree, but the ε- and ε- adjusted F tests both indicate that the parallelism hypothesis H_{01} should be rejected: $F(6.6, 840.1) = 3.76$, $p \leq .001$, and $F(7.1, 898.1) = 3.76$, $p \leq .0004$, respectively. Hotelling's T^2 yields $F(8,120) = 4.03$, $p \leq .0003$, and so we conclude that the parallelism hypothesis is untenable for the two sex groups. As in the analysis for the race groups, we proceed to a separate comparison of the two sexes for each subtask. These results are given in Table 7 and indicate that there is a significant difference between males and females only for subtasks 3 and 7, again using a protected significance level of .009. Separate tests of equal subtask means were also performed for both sex groups, yielding $F(8,58) = 49.53$, $p \leq 4 \times 10^{-23}$, for the males and $F(8, 55) = 23.91$, $p \leq 3 \times 10^{-15}$, for the females. As in the analysis for the two race groups, separate 90% union-intersection multiple comparisons were performed for each sex group,

yielding the results in Table 8. There is some discrepancy in the clusters of subtasks between the two groups; the main difference is that there are only two clusters for the females, whereas there are five for the males.

Insert Tables 7 and 8 about here

Discussion

The social concepts assessment task appears to be a valid and fairly reliable instrument for assessing young children's knowledge of certain basic social concepts that can be expressed in picture form. Given the similar results obtained using two different forms of this task in two separate studies, it seems reasonable to conclude that similar instruments could be developed by other researchers or practicioners who follow the same guidelines.

The significant relationships between the scores for the basic and social concept assessment tasks obtained in the validity analysis indicate parallel development between basic and social concepts. This makes sense, given that category formation skills are basic to most concept formation experiences. Still, it provides another way of assessing general social concept development.

The results of the profile analysis indicate that the subtask scores for kindergarten and first grade subjects are parallel. In addition, the scores of the two groups are not coincident, as the first grade subject's scores were significantly higher on each of the nine subtasks. These results were expected and confirm the general hypothesis that the young child's concept formation skills improve with age. First graders are entering the concrete

operations stage and have one more year of exposure to opportunities for learning about the concepts under study via school, media, family, peers or other sources.

The results also confirm that there are significant differences in difficulty among the nine concept subtasks of the social concepts assessment task. The most difficult concepts for both grade levels, both sexes, and both races seem to be Family, Those Who Protect Us, Past/Present and Rich/Poor. The least difficult were Urban/Rural, War/Peace, and Houses/Other Buildings. Old/Young and Groups/Individuals were only marginally more difficult.

These findings are interesting, but the possible explanations are not completely clear. Certain concepts are less abstract and complex than others; thus, they should be easier to learn or to sort if not previously learned. The easier concepts tested in the subtasks appear to be of this type (e.g., Old/Young; Urban/Rural, etc.). The subjects may also have had relatively more experience with these concepts via the media, schooling or personal experience. For example, media exposure might account for the familiarity of war and peace as few young subjects could be expected to have had direct experience with war. But it is also likely that war-type games remain a common feature of our culture, especially among young males (which may explain their higher scores on this subtask). Such factors could help account for the relative lack of difficulty with this concept.

The concepts tested by the most difficult subtasks all seem to be more abstract and/or complex. Past/Present deals with the concept of time which is highly abstract and quite difficult for young children (Jantz and Klawitter, 1985). The same is probably true of Rich/Poor and People Who Protect Us. The former is a relational concept (see Stanley and Mathews, 1985) which probably

requires significant experience for young children to identify. People Who Protect Us normally requires formal instruction for adequate understanding. In addition, it is somewhat more difficult to display in pictures.

At first glance, one might expect that young children would find Family an easier concept. Certainly most young children have direct and continuous experience with exemplars of the concept. But it is a more complex concept than we generally think. Several different configurations of adults and children might count as an example of family (e.g., a male with a daughter; a female with a son; a male with four daughters, etc.) (Lamb, 1982). However, the young child will tend to see his/her family as the prototype or a best example (and most of the children in our sample lived in traditional families). Finally, young children have little experience (prior to school) with learning about nonexamples of family. Therefore their concept of family is rather narrow and limited. All of these considerations could help account for the subjects' relatively low performance on this subtask.

The results also indicate that Caucasian students had significantly better scores on four sub-tasks: Family, Past/Present, War/Peace and Rich/Poor. However, these results must be viewed with some caution. A preliminary analysis of the data gathered regarding socioeconomic status (SES) suggests that a significantly higher percentage of black students were from low income families. Unfortunately, we were unable to gain access to all the data we needed to do a statistical assessment of the relationship between SES and scores on the social concepts assessment task. Thus there data are not reported here. Still, the data we were able to gather do suggest that the apparent differences between scores of Caucasian and black children is strongly influenced by the variability in social class. There may indeed be differences in the way various groups or sub-cultures (including racial

groups) learn different concepts, but this cannot be confirmed by our data. Still, the results do give one more indication that teachers can often expect a wide range of concept knowledge and recognition skills at the early childhood level.

The differences in scores by gender are also interesting. The male subjects scored significantly higher on the War/Peace subtask; in fact, it was the least difficult subtask for males. Such results do appear to reflect the impact of prevalent sex-role stereotypes in our culture; however, the females also scored well on this subtask. Both groups found family to be very difficult; it was most difficult for males and second most difficult for females. Females found Houses/Other Buildings and males found Urban/Rural easier. Again, these results illustrate the possible variety of concept understanding and recognition skills one might encounter.

Finally, it should be noted that three of the least well developed concepts studied here are among those typically included in the early childhood curriculum, i.e., Family/Not Family; Those Who Protect Us; and Past/Present (see for example, Seefeldt, 1984; Walsh, 1980). Teachers need to be sensitive to the apparent cognitive limitations of young children when teaching these concepts. Their present popularity in the curriculum might tend to obscure these developmental limitations. We are not suggesting that these concepts should not be taught or that young children are not capable of learning the concept at a rudimentary level. In fact, most children appear to arrive at school with some knowledge of the concepts we studied. However, looking at the three most difficult concepts, we can speculate on why they seem so difficult even though the attributes of at least two of them were certainly familiar to the children. For example, in the case of family, as previously mentioned, the childrens' conceptions seem to be rather narrow and

traditional (mother, father, and children) even though today's families come in a variety of configurations (Lamb, 1982). This factor would explain why they would have relatively more difficulty sorting examples and nonexamples that included nontraditional family configurations. Previous research on the concept of family (Powell, Wiltcher, Wedemeyer, & Claypool, 1981) indicated that young children in the past chose mother, father and children most frequently as family and their view does not seem to have broadened with our current social changes. However, this question needs to be looked at in more depth than our task allowed. Taking the second concept Those Who Protect us, young children have certainly had experiences with doctors and nurses and have some knowledge of firefighters, police officers and soldiers, but haven't yet placed them in one category according to their common function. It can be speculated that Past/Present is at a relatively low level of development for a couple of reasons. As children move from preoperational to concrete operational thinking, they are also just beginning to have a clear conceptualization of fantasy and reality. Possibly because they see the past 'live' in movies and on television and because their personal past is relatively short and a great deal of it is not something they can remember firsthand but get from the descriptions and anecdotes of others and from the family photo album, they have difficulty putting past and present into separate categories. Finally, we should also note that our social concept assessment task is not designed to give a full evaluation of a young child's knowledge of the nine concepts studied. This would require the use of a much more extensive interview task. However, the task used in this study does help to determine the level of understanding of each of the concepts in question. When one combines this consideration with the possible range of concept formation skills present in a heterogeneous kindergarten or first

grade, the results of this study illustrate some developmental limitations that need to be considered in planning social studies instruction and point to the need for further research regarding children's knowledge of social concepts and how far this knowledge can be extended through instruction.

References

Brainerd, C. J. (1979) The origins of the number concept. New York: Praeger.

Bruner, J. S., Olver, R. R., and Greenfield, P. M. (1966) Studies in cognitive growth. New York: Wiley.

Carmines, E. G., & Zeller, R. A. (1979). Reliability and validity assessment. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-017. Beverly Hills and London: Sage Publications.

Charlesworth, R. and Stanley, W. B. (1982, November) Assessment of the Social Studies Concepts of Young Children. Presentation at the annual meeting of the National Association of Early Childhood Teacher Educators, Washington, D.C.

Cole, M. and D'Andrade, R. (1982) The influence of schooling on concept formation: Some preliminary conclusions. The Quarterly Newsletter of the Laboratory of Comparative Human Cognition, 4(2), 19-26.

Gardner, H. (1982) Developmental psychology: An introduction (2nd ed.) Boston: Little, Brown and Company.

Greenhouse, S. W., & Geisser, S. (1959) On methods in the analysis of profile data. Psychometrika. 24, 95-112.

Inhelder, B. and Piaget, J. (1965) The early growth of logic in the child. New York: Harper.

Isenberg, J. P. and Jacobs, J. E. (1981) Classification: Something to think about. Childhood Education, May/June, 284-288.

Jantz, R. K. and Klawitter (1985). Early childhood/elementary social studies: A review of recent research. In W. B. Stanley (ed.) Review of research in social studies education. Washington, D.C.: National Council for the Social Studies.

Koeller, S. (1981) Economics education applied to early childhood. Childhood Education, 57, 293-296.

Lamb, M. E. (Ed.) (1982). Nontraditional families: Parenting and child development. Hillsdale, NJ: Erlbaum.

Maxwell, S. E., & Arvey, R. D. (1982) Small sample profile anaylsis with many variables. Psychological Bulletin, 92, 778-785.

McCall, R. B., & Applebaum, M. I. (1973) Bias in the analysis of repeated-measures designs: some alternative approaches. Child Development, 44, 401-415.

McKinney, W. C., Larkins, A. G., Burts, D. C., Davis, J. C. (1982) Teach social studies concepts to first grade students? Research on the Merrill and Tennyson Model. The Social Studies, 73, 235-238.

Mendoza, J. L., Toothaker, L. E., & Nicewander, W. A. (1974) A Monte Carlo comparison of the univariate and multivariate methods for the groups by trials repeated measures designs. Multivariate Behavioral Research, 9, 165-177.

Morrison, D. F. (1976) Multivariate statistical methods (2nd ed.). New York: McGraw-Hill.

Nunnally, J. C. (1978). Psychometric theory (2nd ed.). New York: McGraw-Hill.

Powell, J. A., Wiltcher, B. J., Wedemeyer, N. V. & Claypool, P. (1981). The young child's developing concept of family. Home Economics Research Journal, 73 (4), 137-149.

Rogan, J. C., Keselman, H. J., & Mendoza, J. L. (1979) Analysis of repeated measurements. British Journal of Mathematical and Statistical Psychology, 32, 269-286.

Rosch, R. and Lloyd, B. B. (Eds.) (1978) Cognition and categorization. Hillsdale, New Jersey: Lawrence Erlbaum.

Rouanet, H., & Lepine, D. (1970) Comparison between treatments in a repeated-measures design: ANOVA and multivariate methods. British Journal of Mathematical and Statistical Psychology, 23, 147-163.

Roy, S. N., & Bose, R. C. (1953) Simultaneous confidence interval estimation. Annals of Mathematical Statistics, 24, 513-536.

Seefeldt, C. (1984). Social studies for the preschool-primary c ild, 2nd ed. Columbus, OH: Merrill.

Smith, E. E. and Medin, D. L. (1981) Categories and concept: mbridge, Mass.: Harvard University Press.

Stanley, W. B., Charlesworth, R., & Ringuest, J. L. (1985) Kindergarten and first grade childrens' social concept development. Journal of Social Studies Research. 9, 1-16.

Stanley, W. B. & Mathews, R. C. (1985) Recent research on concept learning: Implications for social education. Theory and Research in Social Education, 12(4), 57-74.

Walsh, H. M. (1980). Introducing the young child to the social world. New York: MacMillan.

Wellman, H. M. (1982) The foundations of knowledge: Concept development in the young child. In S. G. Moore and C. R. Cooper (Eds.) The young child: Reviews of research, Vol. 3. (pp. 115-134) Washington, D. C.: National Association for the Education of Young Children.

Welton, D. A. & Mallan, J. T. (1981) Children and their world. Boston: Houghton Mifflin.

Table 1

Means and Variances of Total Social Concept Scores and Subtask Scores By Grade

Group	Total Social Concept Score	Subtask Score								
		1	2	3	4	5	6	7	8	9
First Sample										
Kindergarten (n=24)										
Mean	49.96	5.83	7.17	3.00	2.63	3.54	3.67	7.71	8.58	7.83
Variance	157.26	8.41	7.54	5.13	5.03	7.39	8.32	11.26	3.91	8.14
First Grade (n=28)										
Mean	65.75	8.54	8.93	5.43	5.46	3.75	6.18	9.14	8.96	9.36
Variance	106.05	4.63	5.25	5.74	4.63	5.82	9.34	3.76	3.07	1.28
Second Sample										
Kindergarten (n=64)										
Mean	59.72	7.14	9.08	4.31	4.70	3.97	5.11	8.70	7.59	9.11
Variance	194.87	8.50	4.87	9.49	10.12	12.22	9.69	6.75	11.13	3.05
First Grade (n=65)										
Mean	69.63	8.57	9.62	4.95	6.09	5.43	6.77	9.54	9.03	9.63
Variance	123.21	3.44	1.77	9.70	10.05	10.78	8.68	2.38	3.66	1.36

Table 2

Correlations of Basic Concept Scores with Total Social Concept and Subtask Scores By Grade

Group	Total Social Concept Score	Subtask Score 1	2	3	4	5	6	7	8	9
First Sample										
Kindergarten (n=24)										
Pearson	.66	.46	.45	.27	.32	.36	.18	.35	.30	.41
Spearman	.66	.43	.55	.24	.40	.35	.19	.33	.41	.45
First Grade (n=28)										
Pearson	.48	.26	.20	.19	.07	.55	.29	.10	.24	.38
Spearman	.48	.40	.22	.14	.04	.49	.21	.25	.30	.33
Total Sample (n=52)										
Pearson	.59	.42	.35	.34	.32	.43	.36	.25	.26	.38
Spearman	.61	.53	.52	.32	.33	.41	.31	.33	.34	.44
Second Sample										
Kindergarten (n=64)										
Pearson	.46	.22	.20	.19	.46	.20	.19	.27	.26	.22
Spearman	.53	.29	.29	.30	.48	.23	.28	.26	.30	.24
First Grade (n=65)										
Pearson	.32	.14	.06	.12	.29	.25	.30	.11	.03	-.01
Spearman	.33	.08	.10	.13	.35	.28	.31	.20	.12	-.09
Total Sample (n=129)										
Pearson	.42	.22	.16	.17	.40	.24	.27	.23	.21	.16
Spearman	.47	.24	.23	.23	.46	.29	.32	.26	.25	.14

Table 3

Reliability Estimates and Standard Errors of Measurement for Social Concept Assessment Instrument By Grade

Group	Coefficient Alpha	Standard Error of Measurement
	First Sample	
Total Sample (n=52)	.762	6.725
Kindergarten (n=24)	.659	7.321
First Grade (n=28)	.663	5.975
	Second Sample	
Total Sample (n=129)	.715	7.205
Kindergarten (n=64)	.687	7.806
First Grade (n=65)	.652	6.550

Table 4

Difficulty Levels of the Nine Social Concept Subtasks for All Students[1]

Clusters				
Most Difficult				Least Difficult
A	B	C	D	E
3.	5.	1.	8.	7.
5.	4.	8.	7.	2.
4.	6.			9.

Continuum

Most Difficult — Least Difficult

3 5 4 6 1 8 7 2 9

[1]Since the parallelism hypothesis was accepted for the population profiles by grade, it is not necessary to present separate results for kindergarten and first grade; both grades have the same clusters and continuum.

Table 5

Means and Test Results for the Nine Social Concept Subtasks by Race

Concept	Caucasian Mean Score	Black Mean Score	Difference	Significance Probability	90% Bonferroni Interval
1. Old/Young	7.92	7.76	.16	.730	(-1.02, 1.34)
2. Urban/Rural	9.51	9.10	.41	.210	(-0.43, 1.26)
3. Family/Not Family	5.28	3.65	1.63	.003	(0.24, 3.03)
4. Past/Present	6.38	3.90	2.48	.000	(1.08, 3.88)
5. Protect/Not Protect	5.22	3.92	1.30	.037	(-0.29, 2.88)
6. Rich/Poor	6.47	5.14	1.33	.017	(-0.09, 2.76)
7. War/Peace	9.62	8.37	1.25	.001	(0.28, 2.21)
8. Groups/Individuals	8.67	7.78	.89	.080	(-0.41, 2.17)
9. Houses/Other Buildings	9.42	9.29	.13	.635	(-0.57, 0.83)

Table 6

Continua for the Difficulty Levels of the Nine Social Concepts Subtasks by Race

Caucasians								
Most Difficult								Least Difficult
5	3	4	6	1	8	9	2	7

Blacks								
Most Difficult								Least Difficult
3	4	5	6	1	8	7	2	9

Table 7

Means and Test Results for the Nine Social Concepts Subtasks by Sex

Concept	Male Mean Score	Female Mean Score	Difference	Significance Probability	90% Bonferroni Interval
1. Old/Young	7.55	8.19	-0.64	.149	(-1.79, 0.50)
2. Urban/Rural	9.65	9.03	0.62	.054	(-0.20, 1.44)
3. Family/Not Family	3.86	5.44	-1.58	.003	(-2.95, -0.21)
4. Past/Present	5.00	5.83	-0.83	.149	(-2.29, 0.64)
5. Protect/Not Protect	4.89	4.51	0.38	.528	(-1.19, 1.96)
6. Rich/Poor	5.79	6.11	-0.32	.560	(-1.75, 1.10)
7. War/Peace	9.67	8.56	1.11	.003	(0.16, 2.06)
8. Groups/Individuals	8.26	8.38	-0.12	.803	(-1.40, 1.15)
9. Houses/Other Buildings	9.21	9.54	-0.33	.216	(-1.01, 0.35)

Table 8

Continua for the Difficulty Levels of the Nine Social Concept Subtasks by Sex

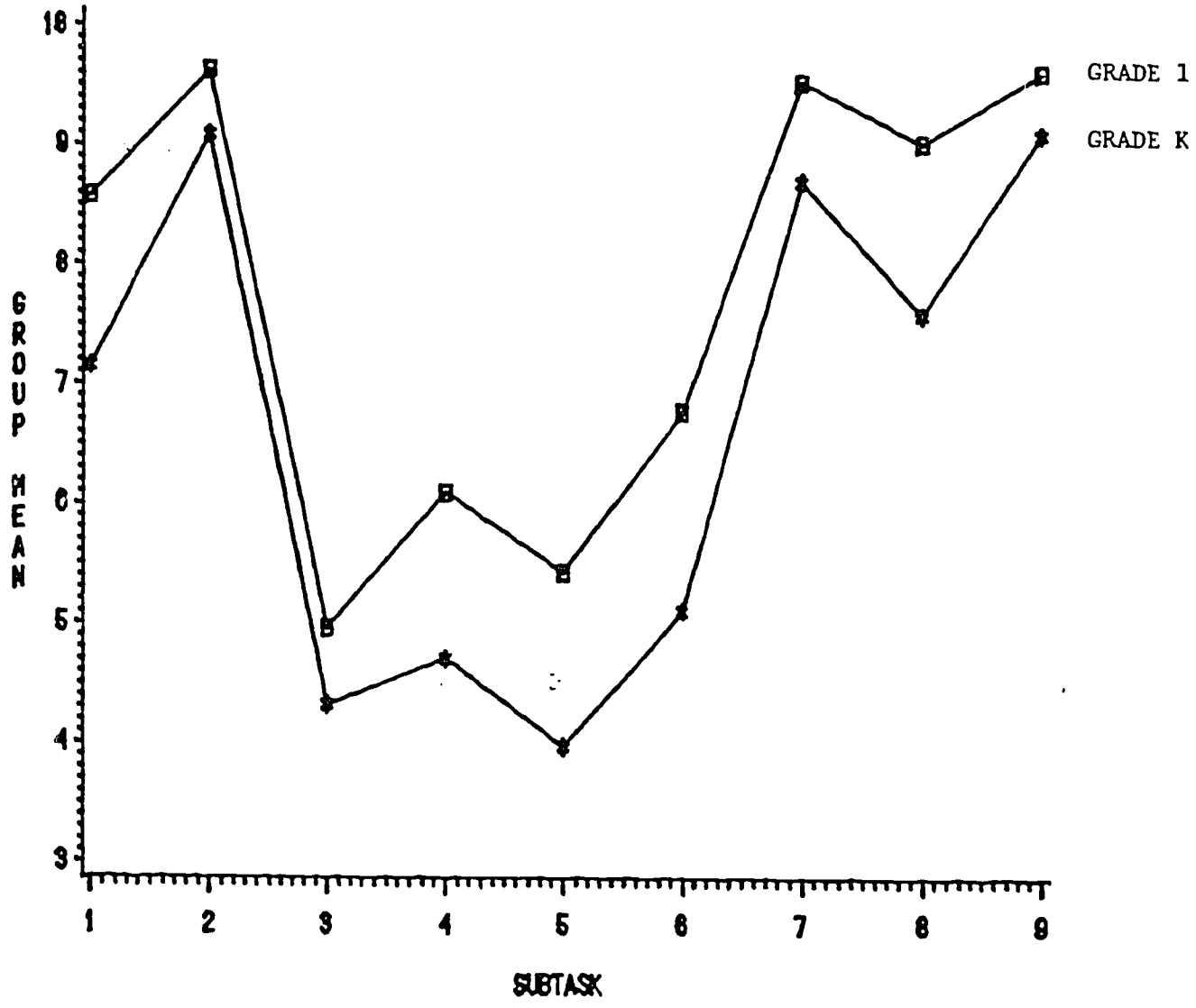
Males								
Most Difficult								Least Difficult
3	5	4	6	1	8	9	2	7[1]

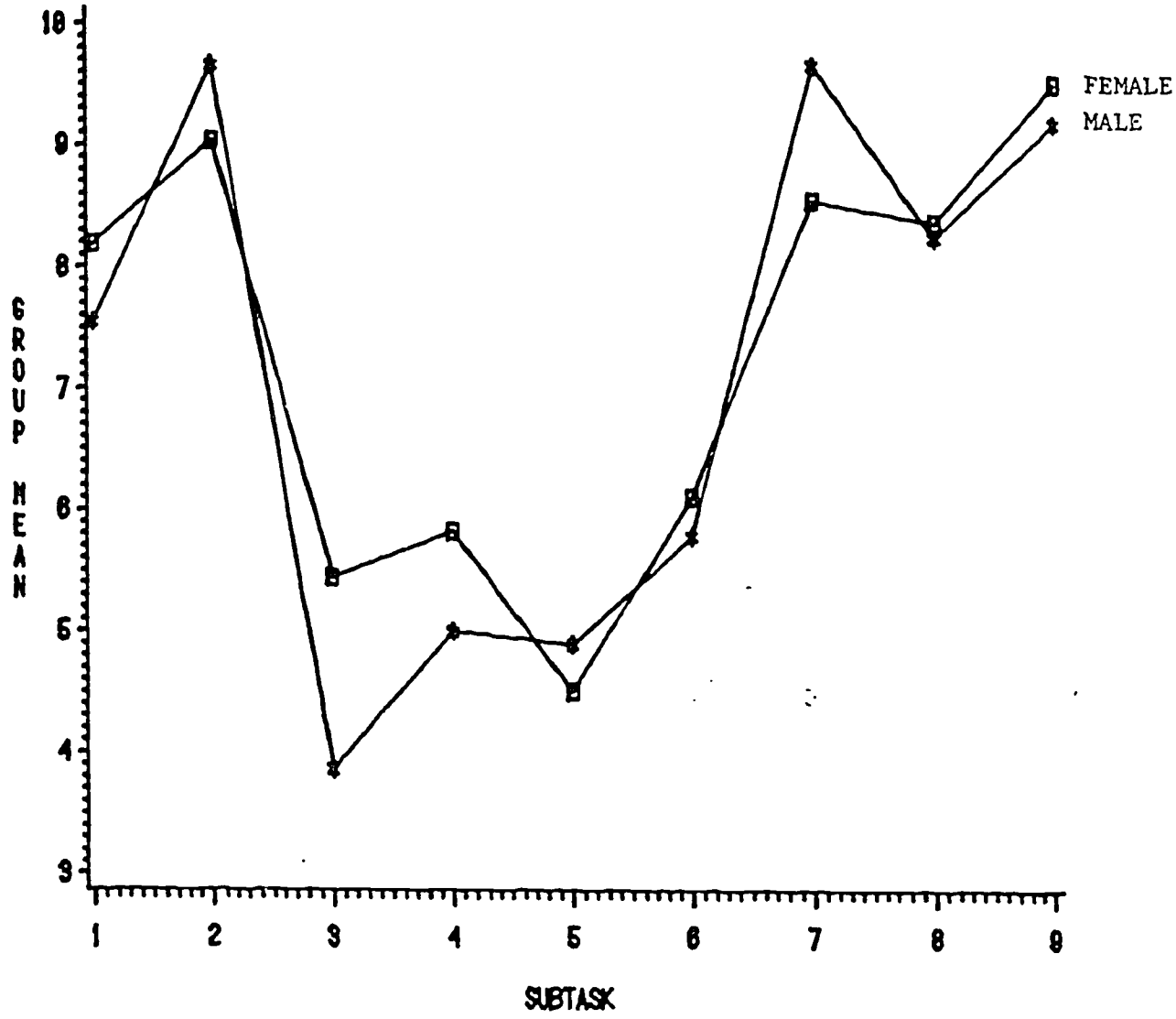
Females								
Most Difficult								Least Difficult
5	3	4	6	1	8	7	2	9

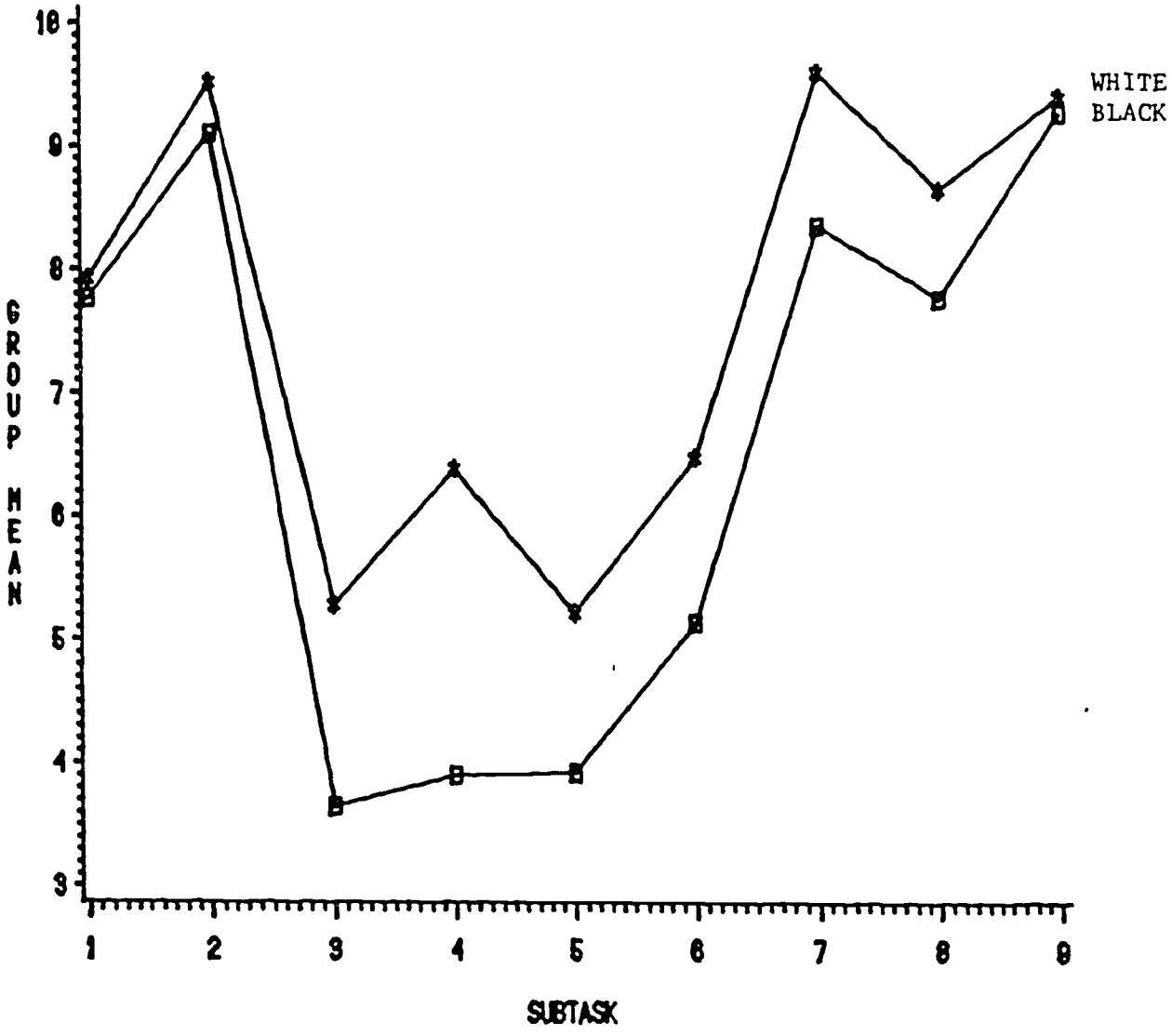
[1]Due to the larger variance for subtask 7, subtasks 8 and 7 were not declared significantly different for the males, whereas subtasks 8 and 2 were.

Figure Caption

Figure 1. Interaction plots for profile analysis of SCAT data by variable. Figure (a) Grade; Figure (b) Sex; Figure (c) Race.


GROUP MEAN
SUBTASK
GRADE 1
GRADE K
10
9
8
7
6
5
4
3
1
2
3
4
5
6
7
8
9


10
9
8
7
6
5
4
3
GROUP MEAN
1
2
3
4
5
6
7
8
9
SUBTASK
FEMALE
MALE


10
9
8
7
6
5
4
3
GROUP MEAN
1
2
3
4
5
6
7
8
9
SUBTASK
WHITE
BLACK