

THINKING AND LEARNING SKILLS

----- Volume 2 -----Research and Open Questions

Edited by SUSAN F. CHIPMAN JUDITH W. SEGAL ROBERT GLASER



3

The Social Context and Socialization Variables as Factors in Thinking and Learning

Anderson J. Franklin The City College and The Graduate School of the City University of New York

Researchers in thinking and learning place importance on models representing general processes of learning. As a result, many studies aim to substantiate the validity of assumptions and hypotheses within these models. I believe that the consequence of this orientation has been the neglect of population characteristics as major variables in model building. Such population characteristics include ethnicity, sex, age, and concommitant social contexts. This phenomenon is illustrated by research reports that treat subject characteristics as incidental to the formulation of hypotheses. It is my contention that another perspective on thinking and learning would emerge from both a more serious theoretical inclusion of population characteristics and the development of models based upon observation of behavior within its natural, social context. Stated simplistically, I am advocating that we develop models of thinking and learning as they are manifested in populations rather than study how populations manifest our models of thinking and learning. Making this conceptual shift is not easy. Likewise, the testing of universal hypotheses and the building of research designs is more problematic in this conceptual framework.

In an effort to move toward this conceptual shift, my research derives materials and hypothesized outcomes from the observation of mnemonic processes operating in the natural daily activities of inner-city adolescents. I selected a social context for analysis that reflects the unique culture and socialization experiences of black inner-city youth. By social context, I mean those social settings in which conditions of interaction consistently yield a prescribed set of behaviors. For me, this became the informal settings for youth such as school hallways, street corners, discos, pizza parlors, etc., where the form of communication entails facility in the use of slang. The use of slang is an essential aspect of socialization

into the peer group. From my observation, the dynamics of using slang manifest many of the issues we are trying to explore with our models of memory. Certainly word association, code switching, encoding, storage, retrieval, short- and longterm memory can be recognized as activities engaged in by youth employing this communicative style. I became interested in exploring how slang, within given social contexts, affects research on fundamental cognitive processes.

In psychology, basic research in memory includes the study of verbal learning processes. The greater ability to use class membership of words during learning is considered representative of acquired higher order learning skills. In his wellpublicized Harvard Educational Review article (1969) and other published papers (e.g., Jensen & Frederiksen, 1973), Jensen has argued that lower class and/or black people "have" less organizing capacity than their middle-class, white counterparts. It is presumed that black children apply relatively more "level I," that is, rote memory, skills to basic recall tasks. Relatively small performance differences between randomly constructed and categorizable lists, plus relatively little semantic clustering in recall by black children, constitute the evidence for this conclusion. Jensen's interpretation of the significance of free recall data has not gone unchallenged. However, the challenges have not been as effective as they should be because they have stayed within the general traditions of standard work in the development of memory. They have not taken account of important issues that enter into intergroup or cross-cultural comparisons (see Cole, Gay, Glick, & Sharp, 1971, Chapters 4 and 7).

In their book on Culture and Thought, Cole and Scribner (1974) note "that a major source of group differences is in the ways of classifying the world that characterize a given cultural group" (p. 99). Central to the verbal-memory task is the group's (or individual's) perception of those attributes and rules that determine class membership. A major unwarranted assumption in all the comparative studies of memory is that category norms elicited from college students are representative of adolescents from social groups not generally found in the college population (Jensen & Frederiksen, 1973; Mensing & Traxler, 1973; Schultz, Charness, & Berman, 1973). In view of the extensive evidence of lexical and syntactic differences among subcultures in the United States (e.g., Cazden, John, & Hymes, 1972; Dillard, 1972; Hall & Freedle, 1973), the use of published norms for comparative studies represents a glaring violation of good research principles. There is abundant evidence to show that word frequency affects recall; yet no provisions are made for assuring equivalence of word frequency across groups, nor for assuring that items within categories are equally related to the category names. In fact, in the absence of any directly relevant data, it seems most reasonable to assume that different groups (e.g., Blacks, Chicanos) do not have the same lexical structure. Before basing conclusions about the differential existence of higher order cognitive capacities on differences between categorizable and noncategorizable lists and on the amount of category clustering, equivalence of materials is an absolutely necessary starting point.

The studies described here are aimed at providing more *valid* data upon which to base statements about ethnic and social class differences in mnemonic skills. The present series of studies was undertaken to first develop category norms for black and white adolescent populations and to then use those norms to form sociolinguistically equivalent lists that could be employed to assess group differences in recall and clustering. When words and categories that are more familiar to black students are included on a recall task, an advantage for blacks over whites in recall and clustering may well result. The first experiment in this series was designed to test this hypothesis by incorporating category instances elicited from black adolescents into a recall list administered to black and white high school students.

We started our investigation by collecting data from 9th and 12th graders on the composition of 25 common categories found in existing category norms (Postman & Keppel, 1970). In this and subsequent studies, subjects were selected from New York area high schools to represent four population groups: black lower class, black middle class, white lower class, and white middle class. Educational opportunity (private vs. public education) also varied.

In our analysis of these data, we were concerned with several points. First, we wanted to establish the group norm for each group—e.g., the relative frequency with which each category member occurs in response to the category name. Second, we were concerned with replicating the work of Jensen and Frederiksen, using stimulus materials tailored to each group. Specifically, we constructed categorizable and noncategorizable lists for each subject group based upon their responses to the category names. By using content variations in materials, we were able to determine whether performance varies correspondingly. Moreover, by sampling different adolescent age levels, the contention that development is stabilized during this period could be empirically tested. These issues represent the major common concern of the verbal-memory experiments described in the following sections. Each experiment provided additional information in this regard by incorporating slight modifications in one or more variables or procedures.

EXPERIMENT 1¹

As mentioned before, the purpose of this experiment was to test the hypothesis that an advantage for blacks over whites in recall and clustering would be found when using words and categories that had been generated (or standardized) by black students.

¹Experiments I and III were also reported in Boykin, A. W., Franklin, A. J., & Yates, J. F. (Eds.), *Research directions of black psychologists*. New York: Russell Sage Foundation/Basic Books, 1979.

Method

Subjects. Subjects in the recall test were 34 adolescents living in the New York metropolitan area. Of these, 17 were black students attending an alternative high school set up in an impoverished area of Brooklyn to accommodate students with poor school performance or unacceptable social behavior in their regular schools. This black sample included 9 females and 8 males with an average age of 18. All black students were placed in the 10th or 11th grade, although the work assigned to them was drawn from the entire range of the normal high school curriculum.

White subjects were drawn from the 11th grade of an integrated parochial school for girls located in central Manhattan. Their average age was 16 years. All white subjects were female.

Materials. To provide a set of category instances that would be specifically tailored to the experiences and categories of urban black adolescents, 75 black teenagers were interviewed in their neighborhoods in central Brooklyn. With each subject in this separate sample, the experimenter went through a list of categories assumed to be high in interest and familiarity for black adolescents and then asked for five instances of each category. The use of slang terms was encouraged. The three categories for which the greatest commonality was obtained for exemplars (dances, soul food, drugs) were selected for use in the recall task. These categories and five commonly cited instances for each composed the Black categories that made up half of the recall list.

The other half of the recall list consisted of the five words most frequently given by college students in response to each of three categories from the Battig and Montague (1969) norms that have been widely used in previous free recall experiments. These category instances, which are referred to as Standard categories, appear along with the Black category items in Table 3.1.

Five randomized orders of the recall list were constructed with the sole restriction that no two items from the same category appeared contiguously in the list. Subjects in Experiment 1 were presented with a different ordering of the list on each recall trial.

Procedure. Subjects were informed that they were participating in a study of memory, that they would be given a list of words to remember and recall in any order they liked, and that the experimenter would go through the list and test their recall a total of five times. The recall list was presented orally at a rate of approximately 2 seconds per word. After the experimenter had read the complete list, the subject verbally recalled as many words as he or she could remember. The procedure was repeated for a total of five trials.

	BLACK CATEGORIES		
I Drugs	II Dances	III Soul Food	
smoke	bump	chicken	
coke	latin	greens corn bread chittlins	
ups	grind		
downs	robot		
acid	truckin'	ribs	
	STANDARD CATEGORI	ES	
IV Tools	V Utensils	VI Clothing	
drill	spoon	shirt	
axe	plate	hat	
saw	cup	socks	
file	glass	pants	
hammer	pan	shoes	
	I Drugs smoke coke ups downs acid IV Tools drill axe saw file hammer	BLACK CATEGORIES I Drugs II Dances smoke bump coke latin ups grind downs robot acid truckin' STANDARD CATEGORIES IV Tools V Utensils drill spoon axe plate saw cup file glass hammer pan	

TABLE 3.1 Conceptual Categories and Member Items

Results

Measures of interest were number of words recalled and level of categorical clustering, as indexed by z-scores (Frankel & Cole, 1971). A preliminary comparison of scores for male and female students within the black sample indicated that the sexes did not differ (at the p < .05 level) in either number of words recalled (t (15) = 0.26) or amount of clustering (t (15) = 0.40). Consequently, scores for black males and females combined were compared to those of white females in all subsequent analyses.

An analysis of variance conducted on the scores for the full 30-item list showed a steady improvement in recall over trials (F(4,128) = 35.50) as displayed graphically in Fig. 3.1. The recall advantage of black students over white students did not attain significance overall (F(1, 32) = 1.66). However, there was a significant interaction between race and trial with black students' recall improving more over trials than that of white students (F(4, 128) = 4.37).

A comparable analysis on the z-scores measuring categorical clustering indicated that black students clustered more than white students (F(1, 32) = 9.63), clustering increased over trials (F(4, 128) = 12.43), and the increase in organization over trials was greater for blacks than whites (F(4, 128) = 3.02). The graphic display of these scores in Fig. 3.2 shows that by the final recall trial, black students were obtaining clustering scores that were more than double those attained by white students.

In order to check whether the deficit of white students in clustering (and, to a lesser extent, in recall on later trials) was confined to the Black categories, ttests were used to compare performance on the two halves of the list (Black



FIG. 3.1 Average number of correct responses per trial.

categories and Standard categories). The mean total recall and clustering score for each type of category by ethnic group are shown in Table 3.2. As is evident from an inspection of the means, white students' level of recall was not better for Standard than for Black categories, (t (32) = 0.92). Although whites' clustering scores appeared somewhat higher for Standard than for Black categories, this difference was not significant (t (32) = 1.56). Similar comparisons conducted on the black students' scores showed no differences between Black and Standard categories in recall (t (32) = 1.23) or clustering (t (32) = 0.71) scores. Sublist score comparisons between races showed that blacks clustered significantly more than whites on the Black (t (32) = 3.58) but not the Standard (t (32) = 1.76) categories.

Correlations between level of recall and amount of clustering showed a strong relationship for blacks in both Black (r = .83) and Standard categories (r = .54). White students' clustering and recall were not positively related for Standard categories (r = .20) or Black categories (r = .08).



FIG. 3.2 Average cluster z-scores per trial.

Discussion

When a list containing items selected for their categorical relatedness for black adolescents was used in a recall task, white adolescents showed a marked deficit in categorical clustering. Applying Jensen's logic to the results of this study

	Black Students n = 17		White Students $n = 17$	
	Mean Recall	Z-Scores	Mean Recall	Z-Scores
Black categories	39.24	3.11	34.00	.29
Standard categories	34.24	2.38	31.00	1.00

TABLE 3.2 Mean Recall and Clustering Over Five Trials, by Race and Category Type

without regard for the list derivation procedure would lead to the conclusion that the white adolescents in the present sample were deficient in terms of conceptual ability. Clustering, the index of conceptual performance (or Level II intelligence) used in Jensen's earlier work, was greater in the recall of black teenagers with poor academic records than in the recall of white parochial school students. However, any conclusions about differences in conceptual ability are clearly unwarranted because many of the list items were probably more familiar and categorizable for blacks. The most important implication of these results is that the outcome of such experiments is going to depend on the particular materials selected and the congruence of those materials with subjects' conceptual categories.

A cautionary note to be kept in mind when considering group differences in clustering scores is that the categorical clustering scores per se do not reveal the particular strategies a group of low clusterers is using. Jensen had interpreted low clustering scores as evidence that subjects are learning a list in a rote or purely associative manner. However, it is also possible that subjects with low categorical clustering scores are organizing the list according to conceptual structures other than those built into the list and measured by the experimenter. In support of this hypothesis, Scribner (1974) has shown that if free recall organization is measured both in terms of categorical clusters and in terms of personal groupings used by individual subjects when asked to sort the recall items, the latter measure often reveals more organization in subjects' recall than does the former. The probably unfamiliarity of many of the Black category instances for the white subjects in the present study raises the question of whether white students were trying to organize their recall along some other lines. In the present case, however, measures of idiosyncratic groupings (Pellegrino's 1971, index of intertrial repetitions) failed to reveal any consistency in the patterns of word groupings in white students' recall over successive trials.

An unanticipated finding was the possibility that the presence of a certain proportion of unfamiliar categories on the list apparently depressed whites' tendency to apply adequately organizational processes even to familiar categories. Previous recall studies using partially categorizable lists (Forrester & King, 1971; Steinmetz & Battig, 1969) had concentrated on demonstrating that recall was greater for the categorizable than for the noncategorizable portion of the list and had not explored the possibility that the presence of some noncategorizable items may impede organization of categorizable ones. Such an effect would be important within the context of evaluating ethnic-group differences in clustering and recall, because the whites' predicament in Experiment 1 may be analogous to that faced by children of nonwhite ethnic groups when they are given lists drawn from standard categories. Many of the list items and categories are probably familiar to them, but some may be unfamiliar, which may result in impaired performance with the list as a whole. Drawing conclusions on the basis of differences in clustering or recall for groups whose relative familiarity with the list items is unknown becomes particularly hazardous in light of the possibility of such "half-categorizable" list effects. Experiment 2 was conducted to obtain clearer evidence for the hypothesis that the presence of some noncategorizable items on a list would impede organization of categorizable items.

EXPERIMENT 2

The performance of the white adolescents in Experiment 1 was consistent with the hypothesis that a certain proportion of items on a recall list that cannot be readily assigned to familiar categories will hinder a subject's performance with those items on the list that he or she *can* easily categorize. However, there was neither direct evidence that the Black category instances were noncategorizable for the white students in Experiment 1 nor the necessary control condition to establish whites' clustering level for an all-categorizable list. To get a clearer test of this "half-categorizable" list hypothesis, standard category norms were used to create a mixed list, half of which was categorizable and half of which was not. Recall on the categorizable portion of the mixed list was expected to be poorer and to show less organization than performance on the same set of items embedded within an all-categorizable list.

Method

Subjects. In order to test for organization effects with a sample similar to that developed with the standard category norms, 40 white college-age students were employed as subjects in the recall task.

Materials. Three 30-item recall lists were developed. The All-Categorizable list contained five frequent responses for each of six Battig and Montague categories (birds, occupations, spices, weapons, geography, and kitchen appliances). Two Mixed lists were constructed, each containing 15 randomly selected items without any category overlap and 15 items from three of the categories included on the All-Categorizable list. (One Mixed list contained the categories *occupations, spices,* and *weapons*; the other included *birds, geography,* and *kitchen appliances*).

Procedure. Subjects received five study-test trials on one of the two types of list using the same procedure as in Experiment 1. Twenty subjects received the All-Categorizable list and then received each of the Mixed lists.

Results

Mean recall and clustering scores over five trials for the categorizable items on the Mixed and All-Categorizable lists are displayed in Table 3.3. (All recall scores in Table 3.3 are means for three categories in order to obtain comparability between the All-Categorizable and Mixed list scores and to allow for easy comparison with the data on Black sublist recall in Table 3.2.) The performance of college students in this experiment parallels that of the white students tested in Experiment 1. The differences in mean recall for categorizable items on the two list types fails to reach significant (t (38) = 1.5, p < .10), but the difference in clustering is significant (t (38) = 2.63 < .01). The mean z-score for categorizable items is almost twice as large when the items are contained within an All-Categorizable list as when they are in a Mixed list. The relationship between clustering and recall is strong when the items are contained within either type of list (r = .82 for both All-Categorizable and Mixed lists). As expected, recall for the noncategorizable items on the Mixed lists was less than that for categorizable items (t (19) = 2.97, p < .01), averaging 6.49 out of 15 for the two Mixed lists.

Discussion

The performance of college-age subjects with a mixed list in Experiment 2 was similar to that of the white adolescents in Experiment 1. It thus seems reasonable to infer that the presence of a large proportion of items for which categories are not readily available will impair the organization of items for which the subject does have convenient categories. This mixed-list effect could stem from the fact that the subjects are less likely to become aware of the categorical nature of these items if the other list items are noncategorizable, either because they decide to devote more study time to the more difficult, noncategorizable items in a mixed list or experience other disruptions in the organization process itself. Regardless of the source of the effect, these results underscore the importance of insuring that *all* items and categories on a free recall list are equally consonant with the experience and conceptual structures of different subject groups if their clustering in recall is to be interpreted as an index of organizational activity.

TAB	E 3.3	
17 (D)	0.0	

Recall, Clustering, and Recall-Clustering Correlations for
Categorizable Items in the All-Categorizable and Mixed-List
Conditions

Group	\overline{X} Recall for Categorizable Items (possible = 15)	Ī	<i>r</i> _{<i>x</i>} , <i>z</i>
All-Categorizable List	8.72	1.54	0.82
Mixed List	7.32	0.66	0.82

EXPERIMENT 3

Evidence for the hypothesis that the sociolinguistic appropriateness of recall items affects recall clustering was obtained in Experiment 1. In that study, half the items on the recall list were category instances elicited from black adolescents, and the other half of the list consisted of items from standard category forms. Black adolescents showed more clustering on this list than white adolescents, and the trend extended even to the standard category items that should have been very familiar to the white students. This finding was interpreted as an indication that the presence of some unfamiliar category items on a recall list (by virtue of being ethnospecific to black adolescents) depressed the whites' attempts to organize even familiar category instances. The present investigation was designed to extend the evidence for effects of ethnospecificity on recall and clustering and to test the generality of the finding that the presence of some unfamiliar items would disrupt performance with familiar ones.

This experiment expands upon the conditions used in the previous study by including several additional permutations in the development of the word lists. Lists were developed from black, white, or standard norms forming lists that were either all "black," "white," or "standard," or "black–white," "black–standard," or "white–standard." The rationale behind this modification was that group by list interactions should generalize to these lists as well if the main hypothesis is correct.

Another minor modification in the recall procedure was incorporated into the design to determine if the expected group differences were a function of storage or retrieval difficulty. After the fifth trial, subjects were given category labels and tested once again for recall on a sixth trial. If no significant improvements could be demonstrated, either storage or retrieval difficulty could be involved, offering no insight into the problem. However, if subjects improved their recall after cuing, retrieval processes would be indicated.

Method

Subjects. Eighty subjects were selected from a parochial school for girls located in New York City. Half the subjects were black and half were white, with each racial group being composed of equal-sized samples drawn from the lower (9th and 10th) and upper (11th and 12th) grades. The mean ages for the two grade-level groups were 14.8 and 17.0 years for the black subjects and 14.9 and 16.9 for the whites.

Materials. In order to develop recall materials expressly tailored for each ethnic group, a separate sample of 109 girls (one-third black) from the same

school were given a questionnaire asking them to name the first five items that came to mind for each of six taxonomic categories. Results of the survey were analyzed separately for blacks and whites, and categories showing the largest degree of group consensus with the five most popular exemplars were identified as Black and White categories, respectively. These categories were used subsequently in composing some of the recall lists. The White recall list (WW) contained the high consensus members of the six categories from the white students' questionnaire responses. The Black list (BB) contained three highconsensus categories from the black students' responses in the present survey and three categories of items elicited from black adolescents in Experiment 1. The Black/White (BW) list contained three of the categories appearing on the Black list and three from the White list. A fourth list was based entirely on items from standard category norms (Battig & Montague, 1969), making this Standard list (SS) similar to those used in many previous recall studies. The final list contained three of these standard categories and three categories from the White list (WS). This White/Standard list (WS) was designed to be analogous to the Black/Standard list used in Experiment 1. This list allowed a test of the hypothesis that, like the white students in the earlier study, black students would show less clustered recall for standard category items when they were part of a mixed list, half of which was tailored to a different ethnic group's sociolinguistic experience. Each recall list contained a total of 30 items, five from each of six categories.

Procedure. Six recall trials on one of the five lists were individually administered to each subject. The word list was read out loud by the experimenter at the rate of about 2 seconds per word and in a different random order on each study trial. After each list presentation, subjects were asked for immediate verbal recall. At the conclusion of the fifth trial, subjects were asked to name all the categories they had noticed on the recall list. The experimenter then gave the subjects the names of all six categories before administering the sixth recall trial.

Results

Recall and Clustering on Trials 1–5: Recall was scored for both number of words recalled and recall organization, as measured by the z-score index of categorizal clustering (Frankel & Cole, 1971). A 5 by 2 by 2 analysis of variance was performed on these data with List, Grade, and Race as between-subjects factors and Trials as a within-subjects factor. The result of the recall and organization analyses closely paralleled each other and are discussed together. Each *F* statistic was evaluated for significance at *p*. less than .05.

A main effect of list was found for both recall (F(4,60) = 3.26) and organization (F(4,60) = 3.80). The basic source of this effect appeared to be the generally inferior performance on the sociolinguistically mixed BW list,

which produced significantly lower recall than the WW (t(14) = 2.89) or the BB list (t(14) = 2.66) and less clustering than the WW (t(14) = 3.53) or WS list (t(14) = 2.50). There was no difference overall between black and white students on either recall or organization (F's (1,60) = 2.83 and 3.52, respectively). Nor did either List by Race interaction attain significance (F's (4,60) = 2.02 and 1.49), contrary to expectation. Older subjects showed more recall and organization than younger ones (F's (1,60) - 13.45 and 8.91 on recall and clustering, respectively) and this age difference did not vary with race (F's (1,60) < 1 and = 2.58). There was a significant Grade by List interaction for both recall and clustering (F's (1,60) = 3.59 and 3.77). Considerable improvement with grade occurred for the SS list, whereas none was found for the BB or BW list. Performance on the other lists showed intermediate levels of improvement with age. There was also a significant Grade by List by Race interaction in recall (F(4,60) = 2.70) but not organization (F(4,60) = 1.32), reflecting somewhat different patterns of list improvement with grade for the two races.

Performance improved over the five recall trials in terms of both number of words recalled and clustering (F(4,240) = 87.21 and 24.66). The extent to which organization increased over trials varied with lists (F(16,240) = 1.75), with the BW list showing little improvement over five trials. The increase in number of words recalled did not vary with list (F(16,240) = 1.33).

Sublist Analysis. Previous research had suggested that the effect of having sociolinguistically tailored categories on a recall list would depend on the nature of the total list. Specifically, the amount of organization subjects display on the list may depend on whether the entire list is familiar and easily categorizable for them or whether some portion of it is relatively unfamiliar and difficult to categorize (e.g., tailored to a different cultural group's experience, cf. Franklin & Fulani, 1979). To see if the present data contained any such effects, separate analyses were performed on those categories that appeared on both the "mixed" BW list and on one of the sociolinguistically "pure" lists (either BB or WW). The BW list was composed of the categories black leaders, soul food, and black *musicians* from the BB list and the categories *national leaders*, *musicians*, and dances from the WW list. A Duncan Multiple Range Test was employed at the p < .05 level for all group comparisons. White students did better in recall on the common white categories when they were contained within the pure WW list than when they were part of the mixed BW list. Black students, on the other hand, did equally well on the common black categories whether they were part of the BB or the BW list. Looking at both types of common categories, black students recalled more words from black categories than from white ones, whereas white students showed the opposite performance pattern. Hence, this analysis of a portion of the recall data provides some support for the hypothesis that list recall and compatibility of items with subjects' sociolinguistic experience are related.

Similar comparisons can be made for the standard categories contained on both the SS and the WS lists (tools, utensils, clothes). The hypothesized outcome was that black students' organization and recall of the standard category items would be better when they were embedded in the SS list than when they were part of the WS list, which presumably contained some category items that were relatively less familiar for blacks. As predicted, blacks showed recall on the common standard category items when they were part of the SS list. White students did not show this effect, as expected, in light of the fact that both white and standard categories should be highly familiar for them. The clustering results for blacks were inconsistent with predictions, however; blacks' clustering of standard categories was not reduced in the WS condition.

Effects of Providing Category Labels. The effect that providing subjects with the category labels prior to the sixth study trial had on their clustering and recall was assessed in separate analyses of variance, with List, Grade, Race, and Trials (Trial 5 or Trial 6) as factors in the design. Clustering and recall were higher overall on trial 6 (F(1,60) = 54.08) for organization and (F(1,60) = 30.60) for recall with the improvement between trials 5 and 6 varying by list in clustering only (F(4,60) = 5.75), and with the SS list showing the greatest increase. Of course, some improvement in recall and clustering would be expected after an additional study trial even if category labels had not been provided; nevertheless, the slope of the function relating clustering to trial clearly changed after category labels were provided. The pattern of recall results was similar. When subjects were interviewed about how they remembered the words, 51% of all subjects reported using categories; 37% of the black subjects claimed using categories in contrast to 65% of the white subjects.

Discussion

The study's results offered mixed support to the original research hypotheses. It seems clear that lists designed to be differentially ethnospecific will elicit varying recall and clustering levels. However, these effects do not consistently vary with race in any straightforward, simple manner. The analysis of common categories in the Black, White, and Black/White lists gave some support for the ethnospecificity hypothesis. Black students did better on black categories and white students on white ones. However, white students did as well as black students on black category items when they were part of the Black list, and analysis of the full set of study results failed to confirm the presence of a List by Race interaction.

The results were likewise equivocal in supporting the hypothesis that a group's recall of familiar categories would be impaired if those categories were part of a list that contained relatively unfamiliar items also. The recall performance of white students on white categories contained within the Black/White list was

poorer than their performance on the White list; similarly, the black students indicated poorer recall for standard categories within a Standard/White as compared to a Standard list. These findings conform to the hypothesis. However, black students did equally well with black categories whether they were part of a Black/White or a homogeneous Black list.

Perhaps the lack of clear-cut results stems from the fact that categories for each cultural group were chosen for their degree of within-group consensus rather than on the basis of the size of group *differences* in familiarity with the items. The underlying assumption was that the most familiar and sharply defined categories for one group would be less familiar to the other. This outlook may be too simplistic. Ethnospecificity is a relative concept rather than an absolute one. Clearly, cultural *sharing* exists among black and white adolescents just as culturespecific experiences do. Some evidence of such sharing was demonstrated by a few items common to both the black and white high-consensus categories (specifically, *Walt Frazier* in the athlete and black athlete categories; *chicken* in the food and soul food categories; *ups* and *downs* in both groups' drug category; and *bump* in both dance categories). Moreover, black students are familiar with the national leaders contained in the white students' leader category, and white adolescents are familiar with the black musicians included on the Black recall list.

The generally poor performance of both racial groups on the Black/White list could be attributed either to its sociolinguistically "mixed" character or to its many proper names (potentially causing interference, or organization into fewer categories, either of which could account for reduced recall).

An unexpected finding was the poor performance on the Standard list in the lower grades. One could argue that this list was simply less unusual or less related to personal interests and experience than the others. However, it did produce the best performance in the upper grades. The marked improvement with grade on performance with the Standard list was more striking among black students but occurred among whites as well. Although inclination to organize items into taxonomic categories in such tasks has been associated with formal education (Cole, et al., 1971; Greenfield & Bruner, 1969), previous research led to the expectation that this tendency would be well established by the ninth grade. Determination of whether or not this apparent increase in the use of taxonomic categories in recall over the high school years is a genuine phenomenon requires further research.

EXPERIMENT 4

The purpose of this study was to replicate the basic study using improved controls for elicitation with fewer list conditions. The elicitation procedures were made sensitive to Grade by Race differences in order to more fully reflect experimental

group characteristics. Also, the elicitation techniques used were more comparable to techniques employed in other elicitation studies, which used questionnairetype formats. Because most classification normative data is collected from college students and word frequency lists, such as Thorndike and Lorge's (1944), which involve word frequencies for primary grade children and adults, there is very little data on high school populations. With the many variables influencing category clustering, the question of whether an adolescent population would generate different classification norms from previous studies, and whether there would be any differentiation in norms across ages or between black and white adolescents became the basis for this study. The empirical question was whether developmental and sociocultural differences would emerge in responses to standard and nonstandard taxonomic categories.

Method

Subjects. The subjects were 108 high school students: 27 white 12th graders, 27 black 12th graders, 27 white 9th graders, 27 black 9th graders. The white subjects were from a parochial school that was 99% white in population. The black subjects were from a city public school with a 99% black population.

Materials. Some of the categories selected were taken from previous norm studies such as Shapiro and Palermo (1970) and Battig and Montague (1969). Some of the categories chosen were ones the experimenter felt would generate cultural differences, e.g., types of dinner foods and types of dances. Some of the categories were universal, generating high-frequency norms, e.g., types of animals and kinds of tools.

Procedure. Each subject was given a booklet containing 29 categories; two categories were listed on each page, and all pages in each booklet were randomized. The following instructions were read aloud by the experimenter: "Name five different things that belong in the following categories. Use one word only. Example, 'A beverage: coffee, tea, milk, soda, and beer." The task was self-paced but had to be completed within a 45-minute class period.

Results

Responses were tallied separately for each group. No attempt was made to separate different forms of the same word (e.g., singular or plural). Synonymous words were tallied the same way as any other response. An index of commonality (IC) was calculated by dividing the five most frequent instances by the total number of responses for all categories. This calculation shows the degree of consensus for all categories.

There were no major developmental differences between ninth and 12th graders or between racial groups on the top five instances for each category. Many categories have a few word instances that differ from group to group, but this was more a function of level of commonality index rather than an absence from subjects' lexicon.

Discussion

The consistency in responses for both adolescent populations by grade and by ethnic group strengthens the assumption of developmental uniformity of instances for conceptual categories. This typing phenomena starts early, with no significant modification during the adolescent phase of development.

Norms generated in this study were consistent with both groups and uniform with previous normative studies, in spite of contrasts between school environments. The white students attended an upper middle-class, highly structured parochial school, whereas the black students attended an overcrowded, loosely structured and frequently noisy public school.

Further inquiry suggests that additional research is needed on the associative strength of classification norms for different populations and its impact on the organization of memory. If the task had been conducted outside the school environment, would there be any differentiation in norms generated between ethnic groups? Did the environment stimulate the use of standard norms instead of norms that might be culturally influenced and verbalized in everyday speech? Moreover, in producing words for the experimental task, black subjects were impeded in part by their intrusive, noisy school environment, spelling deficiencies, and low task motivation. This brings into question the efficacy of group survey as a method for a task of this nature.

EXPERIMENT 5

The fifth experiment was capitalized on the norms collected in Experiment 4, which were used to develop free recall lists for black and white adolescents that were equated in terms of appropriateness and specifically tailored to each group's best defined categories. The design included giving separate samples from each racial group the other group's list so that the performance of different groups on the same list (with appropriateness varying) could be compared as well as performance on different lists with appropriateness controlled.

Method

Subjects. Sixty 12th graders enrolled in the same high schools participating in Experiment 4 were used in Experiment 5. Of these, 20 were girls attending the white parochial school, 20 were girls in the regular curriculum at the black

public school, and 20 were girls participating in the college-bound program at the same black public school. This latter group was included because their educational achievement and curriculum was more comparable to that of the white students attending parochial school, yet they presumably shared the same cultural background as their public school peers.

Materials. Two 30-word lists, one derived from the norms of 12th graders in the black public school and one from 12th graders in the white parochial school, were devised for use in the recall task. The five most frequently cited instances from six categories with very high IC ratings were placed on each group's list with an effort to equate the mean IC for the two lists. The mean IC for the categories selected for the Black list (based on the data of black 12th graders) was .63 and the corresponding index for the White list was .62. Table 3.4 contains a listing of the categories and items selected for the two recall lists based upon each group's respective achieved level of commonality index.

It should be noted that the criteria used in selecting list items yielded quite similar lists from the black and white norms. Four of the six categories are the same for the two lists, and within these common categories, 16 of 24 words are identical. Overall then, over half the words on each recall list also appeared on the other. However, lists were based upon the six categories with the highest commonality index for each ethnic group separately, under the assumption that this list structure was primary in the lexicon of the group. Five randomized word orders were generated to be used on the five recall trials, with the restriction that no two words from the same category appear successively.

Procedure. The recall task was individually administered to each subject. Students were told that they were participating in a study on memory and that the experimenter was interested in how many words they could remember from a list that would be read to them a total of five times. The experimenter read the words at a rate of about 2 seconds per word and asked for verbal recall after each complete reading of the list. Each subject received a random sequence of the five word orders. Recall was recorded on tape for later transcription.

Results

The total number of words recalled over five trials and the average z score indicating degree of categorical clustering were computed for each subject. A 2 by 3 by 5 analysis of variance, with List and Student Group as between-subjects factors and Trials as a within-subjects factor, was run on each of the response measures. The results of these analyses included significant main effects for Student Group and Trials on both recall and clustering (F(2,54) = 6.75 and 8.73, respectively). The significant group effect stemmed from the relatively poor performance of the black students in the regular school program. Recall

			WHITE LIST		
Fruits	Instruments	Liquors	Religions	School Subjects	Relatives
orange	guitar	whiskey	Protestant	math	mother
pear	flute	rum	Moslem	English	sister
apple	piano	scotch	Catholic	religion	aunt
grape	drum	vodka	Hindu	history	cousin
banana	organ	gin	Jewish	science	uncle
			BLACK LIST		
Fruits	Instruments	Tools	Religions	Sports	Relatives
banana	piano	pliers	Christian	basketball	sister
apple	drum	wrench	Moslem	football	mother
orange	trumpet	screwdriver	Catholic	tennis	cousin
strawberry	guitar	hammer	Baptist	baseball	uncle
grape	flute	saw	Jewish	track	aunt

 TABLE 3.4

 Categories and Items from the Two Recall Lists

and clustering increased over trials (F(4,216) = 15.75 and 3.68), and the extent of this increase did not vary by Student Group list. Neither the main effect of list nor any of the interactions attained significance.

Discussion

Contrary to expectation, black public school students did not show as much recall and clustering as white parochial school students when given materials designed to be as appropriate for them as the white students' were for that group. Black public school students in the college-bound program did, however, recall as proficiently as white students. One interpretation of these findings is that the educational experiences offered in the regular track of a crowded black urban high school do not foster the memory skills needed to apply the structures in one's semantic memory to a recall task in the same way that experiences in a disciplined white parochial school or a special college-bound program.

Before accepting the recall and clustering differences as genuine evidence for group differences in propensity to use conceptual structures in a recall task, however, one should critically examine the assumption that the black list used in the present experiment was valid as reflection of the conceptual structures of the black students in the regular track as the white list was for parochial school students. One cause for reflection is the similarity in content of the lists generated from the category norms despite obvious differences in everyday lexicon, interests, and cultural experiences between students at the two schools. A more objective cause for concern grows out of the experimenter's observations during the category elicitation study of Experiment 4. Black high school students did not appear to be comfortable with and fully engaged in the task at hand. Despite urging from the experimenter, some students appeared to be unmotivated, and others expressed a reluctance to write items they were not sure they could spell correctly. The reactions of black students in this relatively formal, school-like elicitation procedure contrasted sharply with those of adolescents from similar backgrounds in an earlier investigation (Experiment 1). In the earlier study, black teenagers were interviewed informally and individually in their neighborhoods and gave responses verbally rather than in writing. They appearaed more willing to think of items in the informal interview, responded more spontaneously, and gave more slang or "street" terms as responses. It is the difference between the items elicited from black adolescents by these two procedures that raises the most serious question about whether the more formal elicitation procedure used in the present experiment and in previous research with college students is really tapping the dominant conceptual structures of black teenagers.

The differences between the types of items elicited under the two procedures can be most unequivocally demonstrated by examining the most frequently given items for the two categories that were included in both elicitation studies dances and drugs. When adolescents were interviewed individually in their neighborhoods, the most frequently cited dances were bump, robot, truckin', grind, and twist. Black 12th graders asked to write instances of dances in a classroom setting, on the other hand, gave such "high-brow" dances as waltz and *ballet* among their five most frequent responses. Similarly, whereas teenagers interviewed on the street gave smoke, acid, and dope as their most frequent responses in the drug category, black students tested in school gave the more "proper" terms for the same drugs, marijuana, LSD, and heroin instead. Category elicitation may be yet another area of behavior where black students appear to have a certain understanding of what is accepted or expected in school and to modify their behavior accordingly. There appear to be more differences between the responses given by black adolescents in the two different contexts than there are between blacks and whites tested in the same context. Concern that the categories in the Black list used in Experiment 5 might not really reflect the semantic memory content of black adolescents led to a follow-up study in which a recall list containing some of these informally derived black "street" category items was administered.

EXPERIMENT 6

Method

Subjects. From the three school populations used in Experiment 4—regular black public school, black public school college bound, and white parochial school—three groups of ten 12th graders served as subjects. All were female.

Materials. The recall list (Black Street list) consisted of six items each from five categories developed from informal interviews with black adolescents conducted in conjunction with an earlier comparative recall study (Franklin & Fulani, 1979).

Procedure. The recall task was individually administered in the same manner as in Experiment 5.

Results

The recall and clustering scores for the Black Street list were combined with those from the previous experiment and one-way analyses of variance performed on the recall and clustering scores for the groups defined by the various combinations of student type and list. (The combined analysis appeared justifiable because the length and structure of the Black Street list was isomorphic with those of the other two lists, the subject samples were from the same populations, and the testing was carried out within the same academic term.) The major result of interest from this analysis was that black public school, regular track students

with the black street list recalled as many words as white parochial school students with the white list. Black regular track students also appeared to cluster somewhat more with the black street list than with the black list but the difference was not significant. Black students' clustering on the Black Street list still fell below the level of white students on the White list.

Discussion

The results of Experiment 6 suggest that efforts to develop a maximally appropriate list for black regular track high school students through the norming procedure of Experiment 5 were not entirely successful. The fact that a list developed from informal interviews produced superior recall for those students suggests that the formal elicitation procedure is not optimal for tapping the content of their semantic memory. The formal procedure may not be optimal for white adolescents either, and the question of whether racial groups would differ in recall on group-tailored lists generated through some other procedure remains open.

CONCLUSIONS

The most noteworthy finding revealed in these studies is the relationship between familiarity with content material and performance in recall and categorical clustering. White subjects' recall and clustering appeared less prominent when items selected for their conceptual relatedness for black adolescents were included in word lists. Furthermore, recall and clustering scores for black subjects increase as their familiarity with the items on the list increases. This is illustrated by their performance improvements from Experiment 5 to Experiment 6, where word lists were more distinguishable for black and white groups. The analyses within and across experiments are consistent in that list recall improves when the items are compatible with the subjects' sociolinguistic experience.

Race or sex per se did not appear related to recall, clustering, or item generation. Also age differences were not found in the elicitation task. However, grade and educational background was associated with recall and clustering. In several instances, college-bound blacks and white subjects had equal recall on White lists. Furthermore, on all tasks, performance of black and white subjects enrolled in the integrated parochial school was more often equivalent than that of subjects enrolled in segregated schools. These findings indicate that equivalence in educational background leads to greater equivalence in performance. The obvious advantage of integrated schooling and equitable exposure is made more obvious by these findings.

Recall and clustering differences were more apparent between ethnic groups when items were derived from individualized elicitation procedures than when the formal (booklet) group elicitation procedures were used. It is assumed that the informal surveys were more reflective of environmental differences because subjects were more relaxed and cooperative in the informal task, and because a one-on-one format may be more conducive than booklets to itemizing slang terms.

Finally, the Jensen hypothesis was clearly questioned by the results of these studies. The demonstration of a relationship between Level II thinking (clustering) and familiarity with items easily draws his conclusions into doubt because word familiarity was a variable Jensen neglected to control.

Implications: The Social Context as the Arena for Basic Research

This series of studies on the sociolinguistic relevancy of verbal learning tasks is only one attempt at formulating research by originating hypotheses that are based on the experiences of the subject populations. There are a number of investigators who are advocating this approach in the study of learning. The October 1979 Special Issue of the American Psychologist on Psychology and Children contains noteworthy articles on this point. When we consider the multitude of social situations in which people make daily transactions, numerous factors are uncovered that contribute to the shaping of their behavior. In considering the processes of learning and thinking, I believe that these social factors are major agents in mediating cognitive development. At present, there is no coherent theory that adequately takes into account the impact of variables within the social context on learning and thinking. Such a theory would explicate how thinking and learning are shaped by a socialization process within each cultural group. I review briefly five primary social contexts that are major settings for socialization experiences and that in my judgment, have an affect on how learning and thinking develop. Each of these domains is a potential area for empirical studies. These suggestions are offered as recommendations for research and funding, particularly when the goals are to comprehend the school performance of adolescents. They are the family, peer groups, school, community, and the individual.

In my opinion, we have not fully understood the role of the family in the development of learning and thinking. Some of the major questions that come to mind are: What is the effect of parent-child interaction over school work on learning and thinking? What role does parental motivation and behavior have in fostering achievement? What is the influence of role-model identification on educational motivation and achievement? What is the influence of sibling interaction and identification on motivation and achievement? What are the roles of extended family networks and family traditions? Certainly in contemporary times one needs to understand the impact of socioeconomic status (SES), which remains dubious and elusive as a research variable because in practice SES

classification is based upon family income and educational level, with little regard for the influence of family values and behavior.

Peer groups provide another major social context in which experiences can shape the processes of learning and thinking. In my research I attempted to use only one aspect of the peer context; that is, slang, which is an important mode of verbal communication. However there are many other ways in which peer groups might affect learning and thinking processes. For example, there is still insufficient information on how peer interaction affects educational motivation and achievement. During adolescence, peer-group membership becomes a prime objective in that it helps to fulfill social interests and to develop identity. Honor clubs in school stimulate positive peer identification. They are intended—at least implicitly—to indirectly influence the learning and thinking processes and, subsequently, achievement. It would be interesting to have a series of in-depth systematic studies examining the effect of peer-group socialization on individual learning strategies.

The school is a third major setting for peer-group interactions and it is a primary social context from which many studies have evolved. No other social context can equal it in research attention about learning and thinking processes. By definition, the school setting is the logical place to explore the development of cognitive processes, but it is not the only context.

The community is an ambiguous concept but it must not be ignored in attempts to describe how the learning and thinking processes develop. There is certainly evidence to indicate that students from high-income communities with considerable school resources have a better performance record than the students of the average inner-city school with limited resources. On the other hand, studies for the New York City Board of Education have shown that effectively administrated schools in the inner city can have a better than expected achievement record. This suggests that more research should be directed to examining the characteristics of achieving and nonachieving schools. For example, what are the characteristics of the community or of other significant support systems of the achieving school? Although this is a rather broad area of proposed research, I still contend that understanding of the most basic elements of the learning and thinking processes should be based on the analysis of such a social context, at least to a greater extent than it currently is.

I also propose the *individual* as a fifth "context" that mediates these critical cognitive processes. The literature has not definitively explained the linkage between motivation and learning processes. In school settings we are repeatedly confronted with the way the socioemotional state of the individual students affects learning and performance. More and more programs are beginning to adopt a psychoeducational approach to the management of students' school performance. We are still ignorant of the way many other individual psychological processes affect the process of learning and thinking.

In conclusion, it is my opinion that the background factors comprising the social contexts are insufficiently considered in research projects on learning and thinking processes. The development of learning and thinking in the individual are shaped by an elaborate network of social learning contingencies; ability and performance must be understood within that framework. The process of socialization contains many variables whose consequences for basic cognitive processes are still unknown. Cross-cultural research in cognition suggests that different ethnic groups perceive and use information according to the traditions of learning for that specific society. Some of the same research issues can be raised for the culturally pluralistic society of the United States. What complicates such investigations into the relationship between socialization and cognitive processes in this country is the factor of acculturation. For example, the inconsistent findings in the studies reported here probably resulted from the fact that the backgrounds and social experiences of the black subjects varied considerably, reflecting the complexity of the acculturation process. However, if we propose to study the learning and thinking processes of inner-city students-many of whom are black and Hispanic or from a diverse collection of immigrant groups-and if we propose to develop educational intervention programs to facilitate and enhance learning, we should formulate our research and intervention programs by considering the characteristics of learning and thinking as manifested by the indigenous populations.

ACKNOWLEDGMENTS

This research was supported by a a grant from the Ford Foundation and the facilities of the Rockefeller University with research assistance from Michael Cole, Lenora Fulani, and Linda DeJesus.

REFERENCES

- Battig, W. F., & Montague, W. E. (1969). Category norms for the verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph*, 80 (3, Pt. 2).
- Cazden, C., John, V., & Hymes, D. (Eds.) (1972). Functions of language in the classroom. New York: Teacher's College.

Cole, M., Gay, J., Glick, J., & Sharp, D. (1971). *The cultural context of learning and thinking*. New York: Basic Books.

Cole, M., & Scribner, S. (1974). Culture and thought. New York: Wiley.

Dillard, J. L. (1972). *Black English: Its history and usage in the United States*. New York: Random House.

Forrester, W. E., & King, D. J. (1971). Effects of semantic and acoustic relatedness on free recall and clustering. *Journal of Experimental Psychology*, 88, 16–19.

- Frankel, F., & Cole, M. (1971). Measures of category clustering in free recall. *Psychological Bulletin*, 76, 39–44.
- Franklin, A. J., & Fulani, L. (1979). Cultural content of materials and ethnic group performance in categorized recall. In A. W. Boykin, A. J. Franklin, & J. F. Yates (Eds.) *Research directions* of black psychologists. New York: The Russell Sage Foundation/Basic Books.
- Greenfield, P. M., & Bruner, J. S. (1969). Culture and cognitive growth. In D. Goslin (Ed.), *Handbook of socialization theory and research*. New York: Rand McNally.
- Hall, W. S., & Freedle, R. (1973). Towards the identification of cognitive operations in standard and nonstandard English usage (ETS RB 73-18). Princeton, NJ: Educational Testing Service.
- Jensen, A. R. (1969) .How much can we boost I.Q. and scholastic achievement? *Harvard Educational Review*, 39, 1–123.
- Jensen, A. R., and Frederiksen, J. (1973). Free recall of categorized and uncategorized lists: A test of the Jensen hypothesis. *Journal of Educational Psychology*, 65, 304–312.
- Mensing, P. M., & Traxler, A. J. (1973). Social class differences in free recall of categorized and uncategorized lists in black children, *Journal of Educational Psychology*, 65, 378–382.
- Pellegrino, J. W. (1971). A general measure of organization free recall for variable unit size and internal sequential consistency. *Behavior Research Methods and Instrumentation*, 3, 241–246.

Postman, L., & Keppel, G. (1970). Norms of word association. New York: Academic Press.

- Schultz, T. R., Charness, M., & Berman, S. (1973). Effects of age, social class, and suggestion to cluster on free recall. *Developmental Psychology*, 8, 57–61.
- Scribner, S. (1974). Developmental aspects of categorized recall in West Africa. Cognitive Psychology, 6, 475–494.
- Shapiro, S. I., & Palermo, D. S. (1970). Conceptual organization and class membership: Normative data for representatives of 100 categories. *Psychonomic Monograph Supplements*, 3, (11, Whole No. 43).
- Steinmetz, J. I., & Battig, W. R. (1969). Clustering and priority of free recall of newly learned items in children. *Developmental Psychology*, 1, 503–507.
- Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York: Bureau of Publication, Teachers College, Columbia University.