DOCUMENT RESUME

ED 052 213 24 TE 002 512

AUTHOR	Williams, Frederick; And Others
TITLE	Attitudinal Correlates of Children's Speech Characteristics. Final Report.
INSTITUTION	Texas Univ., Austin. Center for Communication Research.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-0-0336
PUB DATE	Mar 71
GRANT	OEG-0-70-7868(508)
NOTE	106p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$6.58
DESCRIPTORS	*Dialects, Language Skills, *Speech Evaluation, Statistical Analysis, *Stereotypes, *Student Teacher Relationship, *Teacher Attitudes

ABSTRACT

This research was a series of experiments focused upon teachers' evaluations of videotaped samples of children's speech. The theoretical framework drew mainly from current ideas about dialect features and the associations with social stereotypes which they appear to prompt in listeners. The practical focus was how this process relates to teachers' expectations of children's academic performances. The results of five experiments led to the following generalizations: (1) Teachers' evaluations of speech samples involved judgments along two relatively independent dimensions of confidence-eagerness and ethnicity-nonstandardness. These had generality across video-only, audio-only, and audio-video conditions of stimulus presentation and generality across samples of Anglo, Black, and Mexican-American children's speech. (2) The judgmental process appeared to combine stereotyping behaviors with evaluations of the language samples. (3) Teacher ethnicity interacted with child ethnicity in speech evaluations. (4) Speech evaluations could be used to predict teachers' expectations of children's academic performance in language arts classes. These results suggest that speech evaluation and the attitudinal correlates of dialect characteristics be incorporated into teacher training programs. (Author/DB)

CCR CENTER FOR COMMUNICATION RESEARCH
School of Communication / The University of Texas at Austin / Austin, Texas 78712 (512/471-1095)

FINAL REPORT
Project No. 0-0336
Grant No. OEG-0-70-7868(508)

ATTITUDINAL CORRELATES OF CHILDREN'S SPEECH CHARACTERISTICS

Frederick Williams
Jack L. Whitehead
Leslie M. Miller

March 1971

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE
Office of Education, Bureau of Research

FINAL REPORT
Project No. 0-0336
Grant No. OEG-0-70-2868(508)

ATTITUDINAL CORRELATES OF CHILDREN'S SPEECH CHARACTERISTICS

Frederick Williams
Jack L. Whitehead
Leslie M. Miller

Center for Communication Research

Austin, Texas

March 1971

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE
Office of Education, Bureau of Research

T A B L E O F C O N T E N T S

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The present research was a series of experiments focused upon teachers' evaluations of videotaped samples of children's speech. The theoretical framework for such research drew mainly from current ideas about dialect features and the associations with social stereotypes which they appear to prompt in listeners. The practical focus was how this process relates to teachers' expectations of children's academic performances. The results of five experiments led to the following generalizations:

1. Teachers' evaluations of speech samples involved judgments along two relatively independent dimensions of confidence-eagerness and ethnicity-nonstandardness. These had generality across video-only, audio-only, and audio-video conditions of stimulus presentation and generality across samples of Anglo, Black, and Mexican-American children's speech.

2. The judgmental process appeared to combine stereotyping behaviors with evaluations of the language samples. Given minimal cues for evaluation, it was speculated that stereotyping is relied upon more heavily in making evaluations than when there is the opportunity to experience a relatively detailed series of cues pertaining to a child's language performance.

3. Teacher ethnicity interacted with child ethnicity in speech evaluations. Both Anglo and Black teachers rated Anglo children as the least ethnic-nonstandard, but Black teachers rated Black and Mexican-American children as less ethnic-nonstandard than did Anglo teachers. Mexican-American teachers tended to differentiate only the status of Black children in terms of the judgmental dimensions of confidence-eagerness and ethnicity-nonstandardness. A teacher's amount of experience did not interact with speech evaluations.

4. Speech evaluations could be used to predict, to a substantial degree, teachers' expectations of children's academic performance in language arts classes (e.g., reading, composition), and to some degree performances in classes such as mathematics and social studies.

These results were taken as a basis for suggesting that speech evaluation and the attitudinal correlates of dialect characteristics be incorporated into teacher training programs.

INTRODUCTION AND PROBLEM

Background

The aim in this project was to undertake a series of studies which focused upon relations between children's speech characteristics and teachers' attitudes toward the children. Both the practical and theoretical arguments for the studies centered upon the thesis that a child's speech characteristics furnish important cues for the formation of teacher attitudes toward a child, and such attitudes may weigh heavily in the way the teacher deals with the child.

The first part of this thesis was based mainly upon prior research into speech characteristics and attitudes conducted by the principal investigator (Williams, 1970) as well as upon the early theorizing by others (Lambert, 1963) about the relation between speech and social stereotypes. The second part of this thesis can be defended mainly upon the basis of the research by Rosenthal and Jacobson (1968) which indicated that the experimental manipulations of teacher attitudes had not only effects upon teachers' behaviors toward children, but upon the educational performance of the children as well. Since this latter research is well known, no more of it will be described here.

In a prior experiment (Williams, 1970) it was found that teachers' attitudes about a child's being "disadvantaged" could be described within the context of a two-factor judgmental model relating to the child's speech. This model was developed in a series of steps ranging from the construction of a set of semantic differential scales[1] (based upon interview data from teachers), to the subsequent factor analytic study of teachers' uses of these scales in rating audio-taped speech samples of children (fifth, sixth grade) of varying sex, social status, and ethnicity (Black, Anglo[2]). Such

---

[1]For example: THE CHILD SOUNDS: confident__:__:__:__:__:__:__unsure.

[2]The term "Anglo" as used here refers to Caucasian children; it was employed throughout the present studies because of its use by the populations involved in the project.

analyses revealed two gross dimensions of judgment which were labeled confidence-eagerness and ethnicity-nonstandardness. Based upon interpretations of the individual judgmental scales which entered into these factors, confidence-eagerness appeared to be judgments related to a child's lack of reticence in speech, his fluency, and his tendency to maintain a conversation. Ethnicity-nonstandardness seemed to be the recognition by the teacher that the child was nonwhite and (in this study at least) spoke other than the so-called "standard" version of English. Teachers in this study were sampled from the tapes of the Detroit Dialect Study (Shuy, Wolfram, and Riley, 1967).

Subsequent analyses of the children's speech samples and the prediction of teacher ratings (by multiple regression analyses) based upon characteristics of these samples bore out the above interpretations. The best predictor of confidence-eagerness ratings, for example, was the infrequency of hesitation phenomena in a child's speech. The best predictors of ethnicity-nonstandardness were, as expected, certain recurring nonstandard features in the child's speech (e.g., 'd' for 'th' substitutions, lack of noun plural 's,' pronominal apposition, etc.). Separate ratings of a child's being "disadvantaged" could be directly predicted from these same variables, although Anglo teachers seemed more likely to equate nonstandardness with being disadvantaged than did Black teachers. In short, not only child differences, but teacher differences emerged in the results.

Many aspects of this research were suggestive of the theorizing reported in various places by Lambert (1963) and his colleagues. In essence, their position is that whenever one provides attitudinal associations for speech characteristics, he is revealing his personal stereotype of the people whom he thinks "speak that way." This theorizing, it was thought, provided a potential linkage between the study described above and the types of attitudes which are pertinent in how a teacher treats a child--in particular, the attitudes studied by Rosenthal and Jacobsen (1968). Thus, given the evidence that speech characteristics can be a reliable predictor of a teacher's attitude about a child being "disadvantaged," then there is the potential to carry the prediction one step further--that her expectations about a child's academic performance might also be predicted. Research in this area could provide practical insights into teacher behavior, as well as implications for teacher training. Such insights and implications could be particularly useful for cases where teacher and pupil are widely separated by ethnicity, social class, or both.

Problem

The above reasoning prompted a number of suggestions for further study, including, above all, the use of videotaped speech samples instead of audio-tapes, the attempt to sample varieties of pupils and teachers, and the attempt to relate speech attitudes to social stereotypes, and both to academic expectations. Accordingly, a series of experiments was designed to provide answers to the following overall questions:

1. Can the two-factor judgmental model of confidence-eagerness and ethnicity-nonstandardness be replicated with the use of videotapes and with other teacher and pupil populations?

2. Does the judgmental process tend to reflect a social stereotype elicited quickly and by only a few cues, or does it reflect detailed perception of whatever is presented as representative of a child?

3. What are the relationships of teacher characteristics to judgments of children's speech?

4. What type of preliminary evidence might be obtained of the relation between judgments of speech and a teacher's expectations of academic performances?

The Research Plan

Answers to the above questions about attitudes and children's speech characteristics were sought from five studies, some of which dealt individually with one of the above questions, others of which related to more than one of these questions. The aims of the individual studies are summarized as follows:

Study I: To determine whether the two-factor judgmental model would be found when speech samples were presented in an audio-visual mode, and whether such factors (or any factors that were derived) were pertinent to the differentations of Black, Mexican-American, and Anglo children from families of middle and low social status.

Study II: To assess the latencies of judgments on the two dimensions of confidence-eagerness and ethnicity-nonstandardness with reference to judgments of Black, Anglo, and Mexican-American children.

Study III: To assess the degree to which teacher-Ss' ratings of videotape samples of children's speech would correspond to ratings of speech that teachers might expect to hear from stereotyped descriptions of children from Black, Anglo, and Mexican-American low or middle social status groups.

Study IV: To assess the degree to which the video image of a child's ethnicity affects ratings of standard English speech samples.

Study V: To assess the relations of teacher ethnicity and experiences with speech evaluations relative to videotapes and stereotypes, and to assess the relation between speech evaluations and expectations of pupils' academic performances.

The design and results of these five studies are described in subsequent individual sections of this report. Their interpretations relative to the overall questions of the project appear in a concluding section.

BACKGROUND LITERATURE

Overview

The idea that listeners will make evaluations about a speaker's personality, ethnicity, education, intelligence, or even appearance has received considerable support in the research literature of the social sciences. In experiments concerning such evaluations, the independent variable has been stimuli characteristic of cues ranging all the way from speech in a particular language to the details of particular dialects or styles. The dependent variables have been equally diverse, ranging from personality assessments to details about individual behaviors or traits.

Although there is a variety of studies available in this area, it is difficult to relate them in terms of any systematic scheme. One can, however, arrange the studies in terms of the relative increase in the specificity of the cues (independent variables) serving in the designs.

Relations Between Language and Attitudes

Variations in the Language Spoken

Several studies have involved manipulation of the language which is spoken, as an independent variable, and have sought relations with personality traits, as a dependent variable. Perhaps the best known study of this type was conducted by Lambert and his colleagues (Lambert, Hodgson, Gardner, and Fillenbaum, 1960). In this study, French Canadians and English Canadian college students were presented tapes of bilingual speakers of French and English and asked to rate the speakers on fourteen traits such as leadership, sense of humor, self-confidence, height, and so on. Ss, unaware that the same speaker was heard under two guises, judged the speaker using only the language characteristics as cues for the personality judgments. Surprisingly, both French and English Canadians judged the speakers using the French guise less favorably than the same speaker using the English guise. These responses were interpreted by the researchers as evidence of a community-wide stereotype in which English Canadians were viewed more favorably. Thus it seems reasonable to conclude that the "matched guise" speech samples elicit in the listener attitudes which he associates with the language group of which the speaker is a member.

A study by Preston (1963) demonstrated the effect of sex differences of both speaker and listener on the above type

personality judgments. In this instance English Canadian Ss evaluated French Canadian female speakers as generally more "confident," "intelligent," and "ambitious" than their English Canadian counterparts. However, the same judges showed a negative bias toward French Canadian males although not as predominantly as in the 1960 study. On the other hand, French Canadian female listeners preferred the English Canadian female guises, but found the French Canadian male guise more appealing than the English counterpart. Male French Canadians tended to rate both the male and female English speakers more positively than the French speakers.

Several studies have examined the evaluational reactions of children to determine the age at which stereotypes of different ethnic-language groups are acquired. Again using the matched guise technique, Anisfeld and Lambert (1964) presented monolingual and bilingual French Canadian 10-year-olds with taped speech samples. The children rated each speaker's personality on 15 traits. Results indicated that the monolingual (French-Speaking) children upgraded the French voices whereas the bilingual children showed less differentiation in their ratings of the French and English voices. The researchers suggest that at age ten the child has not been influenced by the cultural stereotype prevaling in the adult population and thus the ratings of the children were not similar to those obtained in the adult French population (as in Lambert, et al., 1960).

Pursuing the question of when devaluation of their own linguistic community becomes apparent in French Canadian children, Lambert, Frankel, and Tucker (1966) asked French Canadian girls ranging in age from 9 to 18 to evaluate the personality of French and English speakers presented in several passages. A comparison of the ratings of the matched guises revealed that a definite preference for the English guise began to emerge at age twelve, especially among bilinguals attending private schools.

An additional study (Lambert, Anisfeld, and Yeni-Komshian, 1965) involving entire differences between languages was conducted using the matched guise approach in the comparison of Hebrew and Arabic speech. Adolescents from Arab Israeli and Jewish Israeli groups rated both Hebrew and Arabic language samples. In rating their own and the other ethnic language samples, Ss did not follow the French and English Canadian patterns of stereotyping. That is, one linguistic community was not rated consistently less favorably by both groups. Rather, Hebrew listeners and Arabic listeners responded to representatives of the opposite group in a mutually unfavorable manner. For instance, both samples of Ss judged their own cultural group as more "reliable," "better looking," more "friendly," and the like.

Variations in Dialect of the Speaker

A number of studies have involved variation in the dialect spoken, but most of these have involved variations in the speakers rather than the use of the matched guise technique. One of the first such studies (Putnam and O'Hern, 1955) involved both the description of dialect differences and a test of the importance of these differences in judging the social status of the speaker. The speech of members of a lower class Negro urban area in Washington, D.C., was analyzed for the range of phonemic and syntactic deviations from "standard" English. Such linguistic phenomena as "weakened consonant articulation," the inclusion of "aberrant vowel and dipthong allophones" and "unsophisticated vocabulary and sentence structure" marked the major differences of this dialect group. Three speakers from this lower class community and nine other speakers of varying higher social status were recorded after they had been read the fable of "The Lion and the Mouse" and were asked to retell it in their own words. These twelve speech samples were then arranged in random order and played to seventy untrained judges who were asked to rate the speakers' social status. Correlation between the judges' ratings and an objective index (Warner's Index of Social Status) was .80. The fact that respondents produced such accurate judgments of the speakers' status on the basis of short speech samples emphasizes the importance of dialect cues in social class identifications.

Harms (1963) using the same speech samples obtained by Putnam and O'Hern in a Negro area of Washington, D.C., elicited responses from midwestern college students regarding the social status of the twelve speakers. Although the midwestern group, unlike the untrained judges in the East, were unaware that the speakers were Negro, the status differentiations were much the same. That is, status dialects were recognized regardless of the race of the speaker. Similarly, although the judges were from two different regions of the country their identification of social status of the speaker was highly correlated with an objective index of status. It becomes clear that dialect differences and the social prestige which they may possess are sometimes not restricted to any particular region of the country.

In another study, Harms (1961) varied not only the status of the speakers but also the status of the listeners. Ss in the Columbus, Ohio, area were classed into three groups using the Hollingshead Two-Factor Index of Status Position. One-minute speech samples from three speakers in each of the three status groups (high, middle, low) served as stimuli. Non-college adult listeners from each of the three status groups rated the speakers' status and credibility. Results

indicated that listeners of all statuses rated speakers in accordance with the objective index, and that listeners, regardless of status, judged high social status speakers to be more credible than low status speakers. The consistency of judgments across status boundaries provides additional evidence of the significance of dialect differences.

Similar to the earlier study of the relationship of language variation and personality characteristics, Tucker and Lambert (1969) assessed relations between different American dialect groups and attitudes. Taped speech samples representative of six American-English dialect groups (Network, Educated White, Educated southern Negro, Mississippi Peer, Howard University, New York Alumni) were played to three groups of college students (northern White, southern White, southern Negro). The students were asked to evaluate the speech samples using an adjective check list describing personality characteristics the speaker might possess and also to rank the dialects in order of their preference. Evaluators of both regions and races rated a dialect typical of "network newscasters" most favorably. Educated southern Negro was rated next most favorably by northern White and southern Black listeners. The southern White students, in contrast, rated educated southern White as the second most favorable. Northern and southern White evaluators agreed in their ratings of the Mississippi Peer dialect as the least desirable; the Negro students, however, described the educated southern White as the least favorable. These findings tended to support the social stereotyping thesis suggested in earlier research (Lambert, et al., 1960) in which ratings of speech samples elicit attitudes that different ethnic groups hold toward one another.

A study by Buck (1968) involved racial identification of speakers, in addition to general attitudinal ratings of "competence" and "trustworthiness." College students were asked to listen to tapes of White and Negro New York speech judged to be variations of "standard" and "nonstandard" English. On several attitude measures the students expressed a preference for the speakers of "standard" dialect. Additionally, the Negro speaker of standard dialect was judged by 24 out of 26 subjects as being White. However, there was no confusion in distinguishing the Negro "nonstandard" speech from the White "nonstandard" speech. Ratings of trustworthiness and competence were also elicited from the listeners as a measure of the speakers' overall credibility. Competence ratings favored standard speakers and showed no race differences. Trustworthiness ratings showed a mixed pattern. Standard dialect speakers of both races were judged more favorably, but the Negro nonstandard speaker was rated more trustworthy than his White counterpart.

A further study (Anisfeld, Bogo, and Lambert, 1962) of this type involved the matched guise technique where the variation was between Jewish accented and non-accented speech samples. Bidialectical speakers were recorded once in standard (Canadian) English and again in Jewish accented English. Results indicated that gentile listeners responded less favorably to a speaker when he was using Jewish accented speech. However, Jewish listeners evidenced variation in their evaluations of persons with Jewish accents. On certain personality traits they were judged more favorably while on other characteristics they were devaluated.

Variations in Individual Linguistic Cues

Recent urban language research in the United States has sometimes involved the study of associations between language characteristics that have been found to be socially stratified and the kinds of attitudes that respondents will relate to such characteristics. Such studies tend to examine two sides of the same coin regarding the phenomena of linguistic stratification. If the social or stylistic variations of a linguistic feature are correlated with the social stratification of its users, then it seems reasonable that such features may be cues when listeners judge the social status of a speaker.

Labov (1966), in a study of New York City speech, sought to determine which variations of pronunciation in a linguistic community were socially significant. In field studies, five phonological variables (/r/, /æ/, /ɔ/, /θ/, /ð/) were systematically examined to determine their appearance when stylistic and contextual situations were held constant.

On the basis of preliminary data, Labov (1966, p. 64) predicted that: "if any two subgroups of New York City speakers are ranked in a scale of social stratification, then they will be ranked in the same order by their differential use of /r/." To test this hypothesis, samples of the pronunciation of /r/ were elicited from sales personnel in three large Manhattan department stores representative of the top, middle, and bottom of the fashion and price scales. The assumption underlying this choice was that the employees of a store attempt to share in the status of the clientele and therefore persons working in the highest ranked store will emulate the speech of its customers rather than the speech of New York sales personnel in general. In a subtle manner, the interviewer asked 264 clerks at random for the location of an item which he knew to be located on the fourth floor, thus eliciting the utterance "fourth floor" in a natural conversational style. He followed their reply with the question,

"Excuse me?" to elicit the same reply in a stressed utterance. The four occurrences of /r/ in the utterances "fourth floor" served as the dependent variable, which was then rated on the basis of a previously devised binary scale. An (r-1) was entered for each case of a plainly constricted pronunciation; (r-0) for each case of unconstricted schwa, lengthening vowel, or no representation. In general, the results revealed a stratification of the pronunciation of /r/ similar to the stratification of the department stores, that is, the employees of the highest ranked store also were ranked highest in their use of (r-1) in both casual and emphatic situations. This evidence tends to confirm the thesis that certain detailed variations in phonemes (in this particular instance /r/) are indicators of status.

In a second field study, Labov (1966) dealt with stylistic variation. Again, the five phonological variables were elicited from speakers in several contexts (e.g., "casual speech," "careful speech," "word lists") and their stylistic variation was analyzed. This analysis revealed that certain phonological features varied consistently across social class lines. Labov postulated that these same variables might serve as cues for listeners in determining a speaker's social status. The task then became one of presenting listeners with samples of speech containing different variations of each phonological variable and asking them to rate the speaker on an occupational scale. Listeners, after hearing sentences containing variants, rated the speaker according to his probable occupation ranging from "television personality" to "factory worker." This test of the subjective evaluation of phonological variables supported the correlation between the appearance of a phonological variant in a certain social strata and the rating of a speaker exhibiting that variant in a similar social strata. Labov concluded that when listeners are given speech samples containing these socially significant variables, they are able to determine the social class of the speaker.

Shuy, Baratz, and Wolfram's (1969) research added an additional variable to this line of investigation. They were primarily concerned with the effect that race and socio-economic status of the listener has on the identification of the race and socio-economic status of the speaker. Speech samples, 20 to 30 seconds in length, which exhibited dialect and pronunciation cues typical of male Negro and White speakers in the Detroit area served as stimuli in this study. The speakers represented four social classes: upper-middle, lower-middle, upper-working, and lower-working. Listeners were drawn from a wide spectrum of Detroit residents, including Negro and White sixth graders, eleventh graders, and adults of both sexes from four social strata. The

response instrument included judgments of the speaker's race and a rating of the speaker's educational/occupational level in addition to several semantic differential items.

Regarding the effect of socio-economic status, results indicated that middle class listeners were more accurate as judges of social class than were lower-class respondents. However, considering the speaker's socio-economic status, it appears that the lower the status of the speaker the more accurately he is identified by listeners regardless of their socio-economic status. In other words, Ss had little difficulty in recognizing the speech of a lower status speaker.

The variable of race tends to follow a somewhat similar pattern, that is, the lower the socio-economic status of Black speakers the more accurate were the listeners' identification of race. In general, Blacks were identified correctly 80% of the time and Whites 81% of the time on the basis of only 30 seconds of speech. From the above data, the researchers concluded that the most outstanding fact in the differentiation of social dialects in Detroit is the presence of stigmatized grammatical and phonological features which are usually present in the speech of lower class and ethnic groups. The speech of the middle class is typified by the absence of these features. In other words, lower class speakers provide more salient cues from which the listener, (regardless of his race or socio-economic status) can differentiate, and these cues are tied to the education/occupation level of the speaker.

Children's Speech and Teachers' Attitudes

Current Research

Williams (1970) undertook the task of uniting two approaches--namely, to simultaneously investigate whatever speech characteristics provided salient cues for attitudinal judgments of listeners and to examine the stereotyping behavior in listeners in response to these salient cues. In short, he attempted to discover the relation between linguistic cues of children's speech and the kinds of psychological stereotypes they prompt in teachers.

The initial phase of this research involved the construction of a set of semantic differential scales on the basis of interview data provided from eight urban teachers. The teachers were asked to describe the speech and language differences of children of varying social status and ethnicity. Such adjectives as "nonstandard," "disadvantaged," "Negro-like," and so on resulted from their descriptions.

In order to determine the relevance of all of the adjectives elicited, a second phase of investigation required teachers to evaluate actual speech samples of children of varying social status and ethnicity. A testing procedure was devised in which audio tapes of fifth and sixth grade, Negro and White, low and high status children were played to a sample of inner-city Chicago teachers. Stimulus tapes had been selected from samples of the Detroit Dialect Study (Shuy et al., 1967). A factor analysis of 33 teachers' semantic differential responses to the speech samples revealed two gross dimensions of judgments which were labeled confidence-eagerness and ethnicity-nonstandardness. Within these two dimensions, scales relating to status judgments were incorporated mainly within the dimension of ethnicity-nonstandardness and to a lesser degree within the dimension of confidence-eagerness. In general, judgments of confidence-eagerness appeared to be related to the child's lack of reticence in speech, his fluency, and his tendency to maintain a conversation. Ethnicity-nonstandardness seemed to be the recognition by the teacher that the child was nonwhite and spoke other than the standard version of English. Scales such as: "The child seems reticent to speak-eager to speak;" "the child sounds unsure-confident," reflected those characteristics associated with confidence-eagerness judgments. Ethnicity-nonstandardness judgments were revealed in such scales as: "The language shows a standard American style-marked ethnic style;" "the child's speech indicates a good educational background-a poor one."

Results of a third phase of analysis speak to the issue of mathematically predicting judgmental responses from specific characteristics of a child's speech. From previous research (Williams and Naremore, 1969a,b) involving these same language samples, detailed and quantified measures of such features as sentence length, silent pauses, verb construction, and /d/ for /th/ substitutions were available. Multiple regression equations utilized these measures as predictors of teachers' judgments. From these equations it was revealed that the best predictors of confidence-eagerness ratings were characteristics such as the infrequency of hesitation phenomena in a child's speech. The best predictors of ethnicity-nonstandardness were certain recurring nonstandard features in the child's speech (e.g., /d/ for /th/ substitutions, pronominal apposition, etc.). Separate ratings of the child's being "disadvantaged" could be directly predicted from these same variables, although White teachers seemed more to equate nonstandardness with being disadvantaged than did Black teachers. Additionally, it was found that some individual teacher-raters were quite consistent with themselves in terms of their ratings of all Black children or all White children, regardless of the actual speech

characteristics of a child. This was evidence that some teachers may have been reporting a stereotype of a child of a particular race or status rather than making differentiations on the basis of the language sample.

In short, this study (1) provided a set of semantic differential scales which were found to be relevant to judgments teachers make regarding children's speech, (2) provided a two-factor judgmental model which was predictable on the basis of objective measures of speech characteristics, and (3) suggested a social stereotyping process as reflected in teachers' judgmental behavior.

This latter suggestion of teachers' judgmental behavior prompted Naremore (1969; in press) to examine the extent to which teachers could be grouped in terms of the commonality in their judgmental responses. Using the same data, this study attempted to determine whether underlying the gross picture provided in the earlier analysis there might be a more detailed and accurate picture of specific types of teachers as defined by the commonality of their rating behavior. If so, on what basis could they be compared or contrasted (e.g., child characteristics, rating scales characteristics, and so on). A factor analysis of the 12 most relevant scales (as determined by the previous study, Williams, 1970) was undertaken in order to group the teachers on the basis of their responses to the speech samples. This particular analysis provided a picture of the rating behavior of four teacher types. The most striking difference in each of the four types was related to the race of the teacher. All teachers under type 1 were White; type 2-4 teachers contained both races, but were split within the groups on the basis of race. For example, of the type 2 teachers, White teachers fell into the positive loading segment, which related to their low ratings of children on the pronunciation standardness scales. Black teachers in type 2 tended to rate children higher on the pronunciation standardness scales and therefore fell into the negative loading segment. This seems a plausible phenomena since it is not uncommon perhaps for teachers of a particular ethnic group to be more familiar with the children of that ethnic group than would a White teacher. Additionally, this exposure and familiarity of Black teachers with the speech of other types of children would perhaps increase their sensitivity to differences in linguistic features. Similarly, the infrequent contact on the part of White teachers with Black children may account for their tendency to rate Black children regardless of status as ethnic sounding and low in social status.

A second analysis in this study examined the differences in the objective correlates of the subjective ratings

assigned by the teacher types. That is, 18 objective measures (e.g., silent pauses, /s/-/z/ deviations, /t/-/d/ deviations) were correlated with the subjective ratings of the children's speech (12 semantic differential scales). The results of these correlations lead the researcher to conclude that there are several different types of teachers in terms of what they attend to in the speech sample. For instance, one type might be characterized as a "detail oriented rater," another as a "communication oriented rater," which implies a tendency to engage in the totality of the situation and the topic rather than dwell on specifics of pronunciation or dialect. In general, White teachers tend to belong to the class of detail oriented raters.

The manner in which teachers form attitudes concerning children was the central focus in a recent study by Seligman, Tucker, and Lambert (unpublished research report). A variety of independent variables--photographs, speech samples, drawings, and compositions of grade III boys--were examined for their role in attitude formation. Samples of these variables were obtained from 36 children from low class and middle class families in the Montreal area. Each of these samples was then evaluated by student-teachers using 7-point scales appropriate to the matter being evaluated. For example, when evaluating the children's compositions the scales included an item such as: "The plot is: simple-complex;" when evaluating the voices, items such as: "Speed of speech is: quick-slow," "Pronunciation is: articulate-inarticulate," were included. On the basis of the student-teachers' ratings, eight drawings and eight compositions were chosen, four from those considered to be poor in quality and four considered good in quality. The selection of eight speech samples was based on the degree to which they were rated as sounding or looking "intelligent" or "unintelligent." Then all combinations of "poor" and "good" characteristics were made, with the exception that a good drawing was always paired with a good composition and together these were taken as one unit of evaluation.

In the main experiment, each combination of characteristics was taken to represent a different hypothetical child. Eight children in all were devised. Nineteen education majors, serving as Ss, were then presented the eight children's photographs, speech samples, drawings, and compositions for their examination and evaluation. Ss were asked to form an overall impression of each child and to rate him on a set of semantic differential scales similar to those used previously (e.g., "The child seems: intelligent-unintelligent"). Analysis of the scores on each rating scale revealed that those children who were rated as sounding intelligent also were rated significantly more favorably on

other dimensions such as being "friendlier," "happier," "more enthusiastic," and so on. Similarly, those students judged from photographs as looking more intelligent were also thought to be significantly more "privileged," "happier," "more self-confident," and better students than those who were rated as looking unintelligent. Also, the effect of a good composition and drawing was shown in those students being rated significantly more intelligent, better students, and more enthusiastic than those boys with poor drawings and compositions.

Based on the results of Williams (1970) the importance of a teacher's rating of intelligence, status, and ability were thought to be affected by the child's voice sample. In the present study, those students with good voices in combination with good drawings and compositions, and in combination with good photographs were rated significantly more intelligent than those children with poor voices, poor drawings plus compositions and photographs. Also on the disadvantaged-advantaged scale good voices were rated significantly higher, as were good photographs. On the scale self-confident-not self-confident (which may be compared to Williams' confidence-eagerness dimension), results revealed the influence of voice. Students with a good voice but a poor photograph were rated as high on self-confidence. In other words, regardless of the quality ascribed to the photograph, a child with a good voice was rated higher on confidence than a child with a good photograph, but a poor voice. In general, speech style was an important factor in the teachers' evaluations on the majority of scales. However, other characteristics were also shown to have their effects on the teachers' ratings, but to a lesser degree. On dimensions such as intelligence, which are crucial to teachers' attitudes toward a child, Ss relied heavily on voice for their judgments. Also, when judging a student's ability physical appearance and voice were the most influential.

In short, it seems that teachers make serious judgments about a child's ability and intelligence on factors such as speech and appearance rather than on information more directly relevant to these judgments (compositions and art work). The effect these judgments have on the child in determining his future scholastic success has yet to be determined.

Implications for Further Research

One question of the Williams (1970) study was whether the same two-factor judgmental model would apply when

teachers could see as well as hear the child. The Seligman et al. study strongly suggests that not only speech attitudes but academic expectations may be closely tied with the types of stereotypes teachers have of pupils. Of course, there is also the question of the generality of the judgmental dimensions of confidence-eagerness and ethnicity-nonstandardness for different tea[illegible]r and pupil populations.

A more theoretical series of implications pertains to the nature of the judgmental process. As demonstrated in several of the studies just reviewed, judgments could be made upon the basis of a very brief exposure to the stimulus materials. Here it would seem that perhaps exposure to only a cue or two would be sufficient to elicit a stereotype in the mind of the listener. It could be that evaluations of this type might better represent the report of a social stereotype than the careful evaluation of a series of cues. Accordingly, a key question for research is the degree to which speech or language judgments involve social stereotyping.

In only a few of the studies reviewed here was there an attempt to study differences among persons doing the speech or language evaluations. Thus, for example, it was important to note differences between French and English Canadians or Black and White teachers as both groups rated speakers representative of their own populations. As shown in the Williams (1970) and Naremore (1969; in press) studies, teacher race is a significant variable in the use of the two-factor model in rating Black and White children. Presumably, there are other teacher characteristics, such as experience which may interact with speech and language judgments.

Finally, there is the intriguing possibility of a fit between the processes of stereotype judgments based upon speech cues and Rosenthal and Jacobson's (1968) thesis about self-fulfilling prophecies in the classroom. In the latter well-known study, as already mentioned in the first section of this report, it was found that manipulation of teachers' attitudes could significantly affect how the child was treated by the teacher as well as the child's overall progress in school. An immediate question is whether teacher expectations about pupils' academic performances can be related to the two-factor model. Is there an overall positive or negative set of academic expectations or might ratings of a child's language be mainly tied to expectations in language arts couses?

The present series of experiments was addressed to the above questions.

THE LANGUAGE SAMPLES

Background

Key to the present research was that videotape samples be obtained of children in the three ethnic and two status groups described earlier. Several somewhat arbitrary decisions were necessary in obtaining these samples.

Child Characteristics

Mainly for the purposes of being able to compare the present series of studies with earlier work (Williams, 1970), it was decided to sample speech from fifth and sixth grade children. For purposes of economy in the present project, and since no differences were observed in the ratings of males and females in the previous study, only boys were recorded.

Status was defined largely upon the basis of the child's neighborhood which in all cases corresponded also to the location of his school. Ethnicity was identified by the researchers. Altogether, 41 children were contacted by liaison with schools in or near Austin, Texas. Of these, the status and ethnic characteristics were 6 Black, low; 6 Black, middle; 6 Anglo, low; 7 Anglo, middle; 9 Mexican-American, low; 7 Mexican-American, middle.

The status identifications of the children's families can be additionally described in terms of the father's occupation. Representative occupations included:

	Low Status	Middle Status
Anglo	Laborer for Gas Company	Architect
Black	Laundry Worker	Professor of Sociology
Mexican-American	Service Station Employee	Painting Contractor

Speech Style

Rather than sample different speech styles, it was decided to elicit the type of speech expected between child and teacher (presuming a relatively friendly teacher). In terms of contemporary studies of social dialects, this would

be a semi-formal speech style. Questions asked of the child were the same as used in the free-speech portions of the Detroit Dialect Study (Shuy, Wolfram, and Riley, 1967):

(1) "What kinds of games do you play?"

(2) "What are some of your favorite TV programs?"

Given responses to these opening questions, the child was encouraged to engage in free conversation (e.g., "How do you play baseball?" "What happened on Gunsmoke?").

Circumstances of the Interview

Permission was sought from families to interview the children on weekends, holidays, or after school in a living-room-like atmosphere arranged in facilities of the Center for Communication Research on the University of Texas at Austin campus. The children were transported to the campus in small groups ranging from three to six at a time, and usually were from the same ethnic and status groups. Prior to individual interviews, the children were entertained in a small conference room by being given plastic car models to assemble and provided with snacks of various kinds. At most times, a college undergraduate student of the same ethnic group as the children supervised the groups. Children were interviewed individually in the room mentioned earlier. Each child was seated on a bench near the interviewer's upright chair. The interviewer, an Anglo female in her mid-twenties, immediately struck up a casual conversation by asking the child about his family. Usually within several minutes, the atmosphere was sufficiently relaxed so that the interviewer could proceed with the scheduled questions. Each child was engaged in approximately 8 minutes of conversation, all of which was videotaped.

All videotape recordings were done on an Ampex 6000 recorder using an Electro-Voice 644 dynamic microphone. These recordings comprised the master files of the project.

Tape Editing

For the studies described in the subsequent sections of this report, segments were selected from the master file of videotapes. Segments were typically taken from selections on the master tapes at similar points in the interview, such as of children responding to the same question. Edited segments were dubbed from the master tapes first to a tape on another Ampex 6000 recorder and then to a one-half inch

test tape using a Sony AV 3600 recorder. The Ampex 6000 recorder was employed in Studies I, II, and III. Because portability was needed in the latter phases of the project, the Sony AV 3600 was used in Studies IV and V.

STUDY I:
SEMANTIC DIFFERENTIAL SCALING OF AUDIOVISUAL RECORDINGS OF CHILDREN'S SPEECH SAMPLES[1]

Purpose

Previous research (Williams, 1970) had indicated that teachers' semantic differential evaluations of audio-taped samples of children's speech resulted in the identification of two major judgmental factors: confidence-eagerness and ethnicity-nonstandardness. The goal of this experiment was to determine whether such a judgmental model would be found when speech samples were presented in an audio-visual mode, and whether such factors (or any factors that were derived) were pertinent to the differentiation of Black (B), Mexican-American (M), and Anglo (A)[2] children from families of middle and low social status.

Method

Subjects

Ss were 102 undergraduate female students enrolled in a course in speech for prospective teachers. Most were freshmen and sophomores, and Anglo.

Materials

Stimulus tapes. A set of six two-minute videotape stimuli was prepared, one for each of six ethno-status groups, Black-Middle (BM) and Lower (BL); Mexican-American-Middle (MM) and Lower (ML); and Anglo-Middle (AM) and Lower (AL). The stimuli were edited from the black-and-white videotaped interviews described in the preceding section.

[1]This study was conducted by Frederick Williams, Jack L. Whitehead, and Jane Traupmann.

[2]This is the ethnic label most often used regionally for Whites or Caucasians.

Semantic differential scales. A 59-item semantic differential was prepared by (1) having some 50 teacher-candidates write short paragraphs describing their reactions to a speech sample after seeing a tape in each of the above ethno-status categories, (2) collating adjectives from these paragraphs, combining them with adjectives from an earlier study (Williams, 1970), then (3) pilot testing scales with teacher-candidates (N=30). The final experimental scales represented an edited (removing redundancies, etc.) version of the pilot scales. The 59-item instrument incorporated seven-step bipolar scales, and each scale gave Ss the further opportunity to check "NR" (not relevant). The 59 items are listed in Table I-1.

Procedures

Testing involved the presentation of stimulus tapes to teachers (for semantic differential responses) in variations of presentation mode--viz. audio (Aud.) only, visual (Vis.) only, and audiovisual (AV) conditions. These three conditions, combined with three ethnic categories (B, M, and A), and two status levels (M and L), were administered to teacher-Ss in a Lindquist (1953) Type IV testing design. Within this design, each S saw three stimulus tapes; these were either all of M or L-status; one was in each of the Aud., Vis., and AV conditions; one each was representative of B, M, and A ethnic groups. Testing was accomplished in small groups (five to 10 Ss) in a conference-type room (Approx. 10 by 20 feet) with tapes played on a 21" TV monitor. Ss were given standard instructions for the semantic differential (with the NR option) and were told that the task was to enable us (Es) to see how the children would differ in ratings. Testing for each subgroup took approximately 25 minutes.

Results

Generality of the Two-Factor Model

Factor analysis. The first major inquiry as to factor structure of the judgments was undertaken as a factor analysis of the intercorrelations among the 59 semantic differential scales, where replicates in the correlations were teacher-Ss by stimuli (N=306). Unities were placed in the

[3]The NR scale option, when taken, was treated as missing data.

TABLE I-1

Adjectival pairs in the 59 scale instrument.

	CHILD PROBABLY IS: eager to learn--"unmotivated" in school*
	CHILD IS: *ignored--listened to BY HIS PARENTS
	WORD USAGES ARE: *consistently-incorrect--consistently correct
	PARENTS SPEND: much--little time WITH THE CHILD*
	THE CHILD IS: highly-fluent--highly-disfluent*
	PARENTS PROBABLY ARE: interested--not interested IN CHILD*
	THE MEANING OF THE MESSAGE IS: *very-unclear--very clear
	CHILD LOOKS: *poorly dressed--neatly dressed
	CHILD LOOKS: attractive--unattractive*
1.#	CHILD SEEMS: *tense--relaxed
	IN THE CLASS THE CHILD PROBABLY: *does--does not DRAW ATTENTION TO HIMSELF
2.	PRONUNCIATION IS: *nonstandard--standard
	CHILD IS: liked--disliked BY HIS PEERS*
	CHILD IS: included--not included BY HIS PEERS*
	SENTENCES ARE: complex-elaborated--simple-unelaborated*
	CHILD'S LANGUAGE IS GENERALLY: *difficult--easy to UNDERSTAND
	CHILD IS: admired--ridiculed BY HIS PEERS*
	CHILD IS MOST OFTEN A: *follower--leader WITH HIS PEERS
	THE CHILD USES LANGUAGE: effectively--ineffectively*
	CHILD SEEMS: competitive--non-competitive*
	CHILD SEEMS TO: like--be indifferent to OTHER CHILDREN*
3.	THE CHILD'S FAMILY IS PROBABLY: *low-social-status--high-social-status
	THE AGE OF THE CHILD IS: seven, eight, nine, ten, eleven, twelve, thirteen, fourteen
	CHILD SEEMS: alert--indifferent*
4.	CHILD SEEMS: *hesitant--enthusiastic
5.	CHILD SEEMS: *shy--talkative

*The asterisks define the pole of the scale assigned a value of 1.0 in the quantification scheme. The asterisks did not appear on the actual instrument.

#The number next to each scale is the scale index for the reduced factor analyses (Tables 2-4).

6. THE CHILD'S SPEECH INDICATES: *a-poor educational-background--a-good-one
   CHILD'S HOME LIFE IS: very similar--very different FROM YOURS WHEN YOU WERE HIS AGE*
   VOCABULARY IS: *unsophisticated--sophisticated
7. CHILD SEEMS TO BE: interested--uninterested IN HIS ENVIRONMENT*
   CHILD FEELS PARENTS: care--don't care ABOUT HIM*
   CHILD WOULD PROBABLY BE: *hostile--accepting TO A TEACHER LIKE YOU
   THE MESSAGE PERSPECTIVE IS: seldom-tied-to-speaker--solely-tied-to-him*
   USUALLY THE CHILD: succeeds--fails IN WHAT HE TRIES TO DO*
   CHILD SEEMS: intelligent--unintelligent*
   CHILD'S RELATIONSHIP WITH HIS PARENTS PROBABLY IS: close--distant*
8. CHILD SEEMS TO BE: observant--not observant*
   THE OVERALL MESSAGE IS: *disorganized--organized
   CHILD PROBABLY SPENDS: large--small AMOUNT OF TIME AWAY FROM HOME*
9. LANGUAGE SPOKEN IN THIS CHILD'S HOME IS PROBABLY: standard American style--marked ethnic style*
   CHILD WOULD PROBABLY BE: *withdrawn--outgoing WITH A TEACHER LIKE YOU
   PARENTS ENCOURAGE CHILD TO DO WELL IN SCHOOL: a great deal--not at all*
10. CHILD SEEMS TO: enjoy--dislike TALKING*
    SENTENCES ARE: *fragmentary--complete
    CHILD WOULD PROBABLY: modify his behavior--refuse to modify his behavior in RESPONSE TO DISCIPLINARY ACTIONS FROM A TEACHER LIKE YOU*
11. THE CHILD SEEMS CULTURALLY: *disadvantaged--advantaged
    CHILD FEELS HE: has much--has little CHANCE TO MAKE GOOD IN THE WORLD*
    CHILD WOULD: respect--disrespect A TEACHER LIKE YOU AFTER HE WAS JUSTIFIABLY DISCIPLINED*
12. CHILD IS: happy--sad*
    CHILD IS: determined--not determined IN SCHOOL*
13. THE CHILD SEEMS: *reticent-to-speak--eager-to-speak
    THE MESSAGE IS: rich-in-detail--sparse-in-detail*
14. CHILD IS: active--passive*
    CHILD IS: *slow--quick
15. THE CHILD SEEMS: confident--unsure*
16. CHILD IS: *dull--alert
    PRONUNCIATION IS: *unclear-indistinct--clear-distinct
17. THE LANGUAGE SHOWS A: standard-American-style--marked-ethnic-style*
    THE GRAMMAR IS: *quite-bad--quite-good

diagonals of the correlation matrix, and factors with latent roots greater than one were rotated with Varimax criteria.

Results indicated the extraction of nine factors accounting for approximately 71% of the variance. A factor corresponding roughly to confidence-eagerness was dominant with 24% of the total variance, followed by a second factor identifiable as ethnicity-nonstandardness with 12% of the total variance. The remaining factors were minor and generally irrelevant to the present research interests.

Since evidence of the two expected factors was found in the first analysis, a more refined analysis was undertaken by reducing the variables to only those scales which had loaded relatively highly on the above two factors in the first analysis. This second analysis resulted in two factors accounting for 70% of the total variance. The rotated factor matrix is presented in Table I-2. As can be seen in this table, the confidence-eagerness and ethnicity-nonstandardness factors are clearly identifiable.

Factor structures by modes and ethnicities. Although the anticipated two-factor structure was found in the overall analysis, there was still the question of its relevance to each of the stimulus subsets. Accordingly, an additional factor analysis of intercorrelations among the reduced set of scale variables was run for each of the Aud., Vis., and AV conditions, and for each of the B, M, and A ethnic subsets.

Results of these analyses are summarized in Tables I-3 and I-4. Without exception, the same scales load highly on Factors I and II in the different mode conditions as well as in the different ethnicity conditions. Even the relative dominance of factors remains similar across the subsets, Factor I accounting on the average for 45% of the total variance and Factor II accounting for 25% of the total variance.

An economical way to describe the almost exact correspondence of the factors involves selecting pairs of analysis (e.g., Vis. and AV) and computing a matrix of cosines among the factor vectors for each pair (Veldman, 1967). These cosines represent the relationship between the factors after the two sets of scale vectors have been aligned for maximum contiguity. The Vis. and Aud. were paired with the AV condition and the B and M conditions were paired with the A condition using the above approach. Results presented in Tables I-5 and I-6 clearly show the high factor correspondence as was observed subjectively in Tables I-3 and I-4.

TABLE I-2

Rotated factor matrix of teacher responses to 17 semantic differential scales.

Variables	Factors I	Factors II
1. relaxed	.77	.11
2. pronun. std.	.22	.70
3. fam. high soc. status	.23	.80
4. enthusiastic	.84	.16
5. talkative	.83	.18
6. ed. bkg. good	.46	.73
7. interested	.76	.19
8. observant	.76	.29
9. home lang. std. Am.	.07	.90
10. enjoys talking	.85	.18
11. cult. advant.	.35	.80
12. happy	.73	.24
13. eager-to-speak	.83	.22
14. active	.86	.18
15. confident	.81	.32
16. alert	.78	.24
17. lang. std. Am.	.09	.90
(Percentage total variance)	44%	26%

TABLE I-3

Factor analyses of 17 scales in each of three mode conditions.

Scales	AV		Aud.		Vis.	
Factor I	I	II	I	II	I	II
1. relaxed	.73	.16	.73	.07	.86	-.01
4. enthusiastic	.86	.07	.84	.24	.85	.12
5. talkative	.77	.27	.86	.14	.88	.05
7. interested	.72	.32	.79	.12	.84	.19
8. observant	.76	.29	.73	.29	.82	.29
10. enjoys talking	.87	.17	.83	.12	.90	.15
12. happy	.58	.32	.80	.34	.81	.09
13. eager-to-speak	.87	.21	.76	.22	.88	.09
14. active	.84	.20	.84	.20	.88	.11
15. confident	.78	.33	.80	.39	.86	.18
16. alert	.72	.31	.83	.25	.82	.11
Factor II						
2. pronun. std.	.18	.74	.29	.65	.05	.73
3. fam. high soc. status	.35	.72	.23	.84	.19	.81
6. ed. bkg. good	.44	.76	.54	.67	-.19	.89
9. home lang. std. Am.	.15	.91	.04	.87	.10	.88
11. cult. advant.	.40	.78	.37	.83	.37	.79
17. lang. std. Am.	.08	.91	.04	.89	.30	.99
(Percentage total variance)	42%	27%	45%	26%	49%	27%

TABLE I-4

Factor analyses of 17 scales in each of ethnicity subset conditions.

Scales	B		M		A	
Factor I	I	II	I	II	I	II
1. relaxed	.75	.26	.76	.05	.71	.11
4. enthusiastic	.82	.04	.90	.11	.82	.23
5. talkative	.83	.18	.82	.14	.81	.21
7. interested	.86	.18	.73	.29	.61	.33
8. observant	.83	.17	.72	.38	.63	.41
10. enjoys talking	.89	.06	.87	.16	.73	.37
12. happy	.79	.20	.64	.28	.78	.20
13. eager-to-speak	.90	.13	.77	.15	.83	.20
14. active	.86	.20	.87	.22	.81	.23
15. confident	.81	.34	.79	.31	.86	.27
16. alert	.81	.26	.80	.13	.72	.42
Factor II						
2. pronun. std.	.19	.63	.14	.73	.47	.49
3. fam. high soc. status	.26	.60	.23	.75	.13	.82
6. ed. bkg. good	.48	.58	.48	.73	.52	.72
9. home lang. std. Am.	-.05	.82	.02	.89	.22	.84
11. cult. advant.	.35	.68	.37	.76	.35	.79
17. lang. std. Am.	-.02	.82	.09	.91	.15	.85
(Percentage total variance)	48%	20%	43%	26%	42%	26%

TABLE I-5

Factor correlations of I and II in AV condition with similar factors in Aud. and Vis. conditions.

		Vis.		Aud.	
		I	II	I	II
AV	I	.9986	-.0538	.9996	.0288
	II	.0538	.9986	-.0288	.9996

TABLE I-6

Factor correlations of I and II in A condition with similar factors in B and M conditions.

		B		M	
		I	II	I	II
A	I	.9989	.0478	.9999	.0152
	II	-.0478	.9989	-.0152	.9999

Factor reliability. As an estimate of factor reliability, a procedure was undertaken involving the calculation of intraclass correlations for selected scales on each factor, conversion to z-scores, averaging for each factor, then conversion back to a correlation coefficient for an estimate of reliability for each factor. These reliability coefficents were .821 for Factor I and .797 for Factor II. The scales incorporated in this analysis were five highest loading for each factor as identified in the results of the reduced scale variable analysis (Table I-2). (A return to the overall analysis was thought justified upon the basis of finding very high similarity among the factor structures calculated for the data subsets, that is, the results shown in Tables I-3 and I-4.)

Scale rejection. It may be recalled that Ss had the option of rejecting certain scales as being not relevant to their differentiation of a given stimulus. Although the foregoing factor analyses are evidence of the relevance of the two factors to the mode and ethnicity conditions, it was nevertheless observed that there were occasional scale rejections. The question prompted by such rejections was whether they would reveal a pattern, say, where a given factor would tend to have a concentrated incidence of scale rejection under given conditions of stimulus ethnicity or mode. To answer this question the incidence of scale rejection among the reduced scale variable set was calculated for Factor I and Factor II so as to serve as a dependent variable in an analysis of variance incorporating stimulus mode, status, and ethnicity as independent variables. One such univariate analysis was conducted for Factor I and one for Factor II.

Results for Factor I indicated a significant main effect for mode, $F(2,180)=4.8, p<.01$: a significant mode-by-ethnicity interaction, $F(2,180)=5.5, p<.01$; and a significant three way interaction, $F(2,180)=4.1, p<.05$, involving mode-by-ethnicity-by-status. Interpretations were drawn from the three-way interaction. Considering the count of scale rejection as an irrelevance index, for confidence-eagerness ratings the highest rejection mean (.813) involved MM children being rated in the Vis. condition. The other most salient feature was that the AV condition has the lowest rejection index (.031) as compared with the Aud. (.208) and Vis. (.271) conditions. These mean figures represent rejections out of a possible 10.0, thus although there was some variation in rejection across the conditions of the experiment, the incidence of rejection was relatively low. If anything, there was only a slight tendency to have greater than average rejections in the Vis. condition, and this seemed to apply mostly to the MM child.

Results for Factor II rejections indicated only a significant main effect, $F(2,180)=342.9, p<.001$, for the mode variable. Inspection of the means indicated that scales on the ethnicity-nonstandardness factor were far more likely to be rejected in the Vis. condition (2.50), than in the Aud. (.188) or AV (.042) conditions. With a possible total of seven scales, this represents an approximate one-third rejection rate. Thus, if there is a question of the irrelevance of either factor to any of the conditions, the point of most potential irrelevance is for ratings of ethnicity-nonstandardness in the Vis. mode.

Differentiation of Children by Ethnicity and Status

To assess for differentiation of the stimulus conditions, factor scores were first calculated as a weighted combination of scales based upon factor loadings in the reduced variable analysis (Veldman, 1967); these scores were then subjected to a univariate analysis of variance for each factor.

Factor I, confidence-eagerness. All sources of variance except the mode main effect were significant in the analysis of variance. Most pertinent to interpretation was the significant, $F(2,180)=9.4, p<.01$, interaction of mode-by-ethnicity-by-status. The cell means for this interaction are presented in Table I-7. As found in earlier use of the two-factor model (Williams, 1970), the expectation was that children who had been selected *a priori* to represent middle and low status families would be differentiated by ratings on each of the two factors. With some exception--with M children in the Vis. and AV mode--children from the middle status group were consistently rated more favorable on confidence-eagerness than children from the low status group. As can be seen in Table I-7, however, some of the differences are relatively small. Since the present experiment represented only one child in each category, it was not deemed worthwhile to undertake statistical generalizations based upon multiple mean comparisons in this table. The only gross generalization claimed is that a status differentiation was, for the most part, revealed.

Factor II, ethnicity-nonstandardness. The analysis of variance results did not reveal another three-way interaction in the case of Factor I, but did identify two two-way interactions. These were between mode and ethnicity, $F(2,180)=12.5, p<.01$, and in mode-by-status $F(2,180)=17.4, p<.01$. In order to compare these results with those shown in Table I-6 and because the two two-way interactions can be seen in this table, Table I-8 presents the cell means for the mode-by-ethnicity-by-status conditions. The most marked feature in

TABLE I-7

Cell means of Factor I scores in mode-by-ethnicity-by-status interaction.

Ethnicity:	B		M		A	
Status:	M	L	M	L	M	L
Aud. Mode:	.84	-.26	-.33	-.47	.62	.36
Vis. Mode:	.99	-1.56	.09	.19	.26	-.24
AV Mode:	.73	-.52	-.88	-.02	.60	-.11

TABLE I-8

Cell means of Factor II scores in mode-by-ethnicity-by-status conditions.

Ethnicity:	B		M		A	
Status:	M	L	M	L	M	L
Aud. Mode:	-.50	-1.26	.55	.07	.82	.34
Vis. Mode:	-.57	-.12	-.27	-.07	.53	.53
AV Mode:	-.57	-1.27	.52	-.35	1.06	.19

the pattern of the cell means is the lack of anticipated status differentiations in the Vis. mode. The mode-by-ethnicity interaction seems generally a function of a pattern whereby B children were rated lowest in the Aud. and AV conditions, as against A children being rated high in all three. The mode-by-status interaction seems generally a function of the lack of status differentiations in the Vis. mode, as against the presence of such differentiations in the other two modes.

In all, the results point to the generalization that Factor II or ethnicity-nonstandardness will not result in anticipated status differentiations when the child is seen but not heard. Although this may seem to belabor the obvious, earlier speculation (Williams, 1970) as well as the emergence of the two-factor model in the present factor analyses of the Vis. condition suggests that teacher-Ss will use Factor II scales such as relate directly to speech even though they only see and do not hear the child. Such use, however, does not seem to result in the anticipated differentiation.

Discussion

Primary implications of the present study refer to the reliability and validity of the two-factor model in teachers' evaluations of children's speech as presented on videotapes. Reliability was assessed in terms of the consistent emergence of the two-factor model as well as in terms of indirect estimates based upon intraclass correlations. Validity was implied by interpretable status differentiations of the children. The results provide a basis for using this two-factor model in audiovisual studies of children's speech ratings, particularly where the AV stimulus mode is to be used.

STUDY II:
LATENCY OF TEACHERS' SEMANTIC DIFFERENTIAL RATINGS OF CHILDREN'S SPEECH[1]

Purpose

The aim in this experiment was to assess the latencies of judgments on the two dimensions, confidence-eagerness and ethnicity-nonstandardness. Casual observation (Study I) had indicated, and speculation suggested, that ratings of ethnicity-nonstandardness required less exposure to the stimulus in real time than ratings of confidence-eagerness. The main reasoning here was that visual cues of ethnicity and the high frequency of cues pertinent to nonstandardness would make the former dimension more immediately relevant to a rater than the latter.

Method

Subjects

Ss were 15 teachers and prospective teachers from three upper division and graduate summer courses at the University of Texas. All but two of the Ss were experienced teachers. Fourteen of the Ss were Anglo, one was Black. There were five males and ten females in the group.

Materials

Stimulus tapes. A set of six two-minute videotape stimuli was prepared, one for each of six ethno-status groups, Black-Middle (BM) and Lower (BL); Mexican-American-Middle (MM) and Lower (ML); and Anglo-Middle (AM) and Lower (AL). The stimuli were edited from the black-and-white videotaped interviews described in a preceding section.

Semantic differential. Scales selected for use were derived from Study I. The judgmental model of confidence-eagerness and ethnicity-nonstandardness was indexed by scales 1 to 5 and 6 to 10 respectively in Table II-1. Scales 11 to 15 were filler items. These scales were individually printed on Hollerith data cards which were prepunched to facilitate subsequent collation and scoring.

[1]This study was conducted by Frederick Williams, Jack L. Whitehead, and Jane Traupmann.

Procedure

Ss were tested individually while seated at a small table approximately four feet from a 21-inch television monitor. A footswitch located under the table allowed the S to signal E who was located in an adjacent room where the television monitor was controlled. S was instructed to depress the footswitch for any duration up to two minutes of viewing the videotape stimulus. This footswitch also controlled an event marker, making it possible for E to record the incidence and duration that S had signalled for tape playback. For each videotape presentation, S was given a shuffled deck of 15 scale cards and was instructed to distribute these cards scale-side up on the table in front of him. S was then told to begin watching the tape (signalling by footswitch) and to complete the scales in any order he wished. S was instructed to stop the stimulus tape and mark the scales as soon as he was able to make a judgment. As S completed each scale he was instructed to place the card in a nearby box. By use of a one-way glass E observed and recorded the time at which each card in sequence had been completed by S. By keeping the response cards in the order that S had stacked them, it was subsequently possible to identify the individual scales involved in the recorded response times. Altogether, testing involved the presentation of six stimulus tapes, the order of which was randomized for each S. Ss were given standard instructions for use of the semantic differential scales. A brief practice session involving a sample stimulus tape and one set of 15 scale cards was undertaken prior to testing.

By interpretation of the event recordings it was possible to calculate for each S the amount of viewing time taken relative to each scale completion. This involved the identification and summation of times that the footswitch had been depressed prior to a given scale's completion.

Results

Latency Variation by Scale and Stimulus

The most direct inquiries were whether the average time measured for individual scale completion would vary, first, across the scales; and secondly, whether the stimulus sets themselves would show latency differences. To answer these inquiries, a two-by-three-by-ten analysis of variance was undertaken with response latency as the dependent variable.

Relative to the above order of inquiries, there was a significant main effect, $F(9,126)=5.6, p<.01$ across the scale

TABLE II-1

Scales used to index the two-factor model and fillers.

1. THE CHILD SEEMS: *reticent-to-speak--eager-to-speak
2. THE CHILD SEEMS: *hesitant--enthusiastic
3. THE CHILD IS: active--passive*
4. THE CHILD SEEMS TO: enjoy--dislike TALKING*
5. THE CHILD SEEMS: *unsure--confident
6. THE CHILD SEEMS CULTURALLY: *disadvantaged--advantaged
7. THE CHILD SOUNDS: Anglo-like--does not sound Anglo-like*
8. LANGUAGE SPOKEN IN THIS CHILD'S HOME IS PROBABLY: standard American style--marked ethnic style*
9. THE CHILD'S HOME LIFE IS: very similar--very different FROM YOURS WHEN YOU WERE HIS AGE*
10. THE CHILD'S FAMILY IS PROBABLY: *low-social-status--high-social-status
11. THE CHILD SEEMS TO BE: interested--uninterested IN HIS ENVIRONMENT*
12. THE CHILD SEEMS: intelligent--unintelligent*
13. THE CHILD SEEMS: *non-competitive--competitive
14. THE CHILD PROBABLY SPENDS: large--small AMOUNT OF TIME AWAY FROM HOME*
15. THE CHILD IS: determined--not determined IN SCHOOL*

*The asterisks define the pole of the scale assigned a value of 1.0 in the quantification scheme. The asterisks did not appear on the actual instrument.

TABLE II-2

Ranked mean latencies of the 10 response scales.

Variable	$\bar{X}$ Latency
lang. std. Am.	74.7a*
Anglo-like	74.9a
eager-to-speak	85.8ab
enjoys talking	86.6ab
enthusiastic	91.0bc
fam. high soc. status	93.4bcd
confident	94.4bcd
active	99.1cd
home similar to yours	103.6d
cult. advant.	104.7d

*Means of a common subscript are not significantly different ($p<.05$) from one another.

variable. The means for the levels of this variable are presented in Table II-2, as are the results of a Duncan (1955) based multiple mean comparison. It may be recalled that the anticipation was that scales pertinent to ethnicity-nonstandardness ratings would be used prior to those pertinent to confidence-eagerness. As can be seen from Table II-2, this anticipation was not realized in the data. Although there was a trend in this direction, the differences were not significant between the first cluster of two scales from the ethnicity-nonstandardness dimension and the second cluster of two scales from the confidence-eagerness dimension. Thus, while there is some variation in individual scales, it may be concluded that the dimensions of the two-factor judgmental model do not differ markedly in terms of response latency.

Relative to the second inquiry--differences due to ethnicity or status of the child--there were significant sources of variation. One was a significant main effect, $F(2,28)=3.3, p<.05$, involving the ethnicity variable. The order of latencies was as follows: B (96.5 sec.), M (91.1 sec.), and A (84.8 sec.). There was a significant ethnicity-by-status interaction, $F(2,28)=5.6, p<.01$. The means pertinent to this interaction are summarized in Table II-3. Here the source of the interaction is that the aforementioned latency differences by ethnicity only appear for the middle status children, the lower status children's means being roughly equivalent. Although the present design does not provide a basis for objectively interpreting reasons for this pattern, one speculation is as follows: The average latency--here, say, about 91 sec.--would be observed when Ss are responding to stimuli that generally "fit" their stereotype expectations. If that fit is exceptionally good--that is, if the videotape of the child is in direct accord with expectations--the latency may be even less than average. On the other hand, when the expectation may run counter to a stereotype, the latency may be longer than average. The present pattern could fit the average stereotypes held by a group of young predominantly Anglo teacher-Ss. To be sure, however, this is a speculation in need of further research.

Latency Variation by Magnitude of Rating

Another possible factor related to latency variation was the magnitude of individual scale ratings--that is, the degree to which a child was rated as favorable (or unfavorable) on an individual scale. Dual arguments could be made for having either a significant positive or negative correlation between latency and actual scale markings. Thus, for example, there

could conceivably be a bias for Ss to mark unfavorable characteristics first or vice-versa. Accordingly, correlational analyses were undertaken for each scale relative to its corresponding latencies. Each of the ten correlation coefficients were tested for significance against a null hypothesis of zero correlation. No correlations were statistically significant. In fact, the average (by z-transformation) correlation between scale marking and latency was negligible (.080).

Discussion

The primary finding of the present study was that both ethnicity-nonstandardness and confidence-eagerness were relatively close to one another in terms of average rating latency. In brief, the implication was that the dimensions of the two-factor judgmental model do not tend to precede one another markedly in response time. Further, from S's own control of the videotape stimuli, it was found that the average response latency across all stimuli was roughly on the order of one and one-half minutes. Some latency differences were found as a function of child ethnicity and status, however, the patterns of these differences could only be interpreted speculatively. Finally, no relation was observed between the magnitude of ratings and the latency of same.

Theoretical implications of the study relate mainly to the temporal similarity of the two dimensions of the judgmental model, and the generality of this similarity across child ethnicity and status, as well as across magnitudes of scale rating. A practical implication is that testing designs of the present type can safely be used with two minutes of stimulus presentation, and that temporal interactions with judgmental factors should remain nil. These implications depend, of course, upon having children and Ss of the same general category as used in the present study.

TABLE II-3

Mean latencies in seconds for the ethnicity-by-status interaction.

		Ethnicity		
		B	M	A
Status	M	102.3	90.0	74.8
	L	90.6	92.3	94.8

STUDY III:
CORRESPONDENCE BETWEEN SEMANTIC DIFFERENTIAL RATINGS OF CHILDREN'S SPEECH AND SPEECH ANTICIPATED UPON THE BASIS OF STEREOTYPE[1]

Purpose

The goal of this study was to assess the degree to which teacher-Ss' ratings of videotape samples of children's speech would correspond to ratings of the speech that teachers might expect to hear from children of certain ethnic and social status groups. Previous research (Naremore, 1969) suggested that teachers, in being highly consistent with themselves in their ratings of the speech of different children, may have been reporting their stereotypes rather than detailed judgments of those children. Elsewhere (Williams, 1970) this speculation has been extended to the thesis that teachers' judgments of children's speech can reveal the social stereotypes associated with children of different ethnic and social status groups. In the present research, it was assumed that stereotype judgments could be obtained by having teachers rate the speech they "expected" from a child whose social status and ethnicity was described for them. The question was whether ratings obtained under such conditions would correspond to ratings obtained when the teacher judged the videotaped speech sample of a child selected a priori to represent a specified ethnic and status group.

Method

Subjects

Ss were 15 teachers and prospective teachers from three upper division and graduate summer courses at the University of Texas. These were the same Ss as in Study II.

[1]This study was conducted by Frederick Williams, Jack L. Whitehead, and Jane Traupmann.

Materials

Stimulus tapes. A set of six two-minute videotape stimuli was prepared, one for each of six ethno-status groups, Black-Middle (BM) and Lower (BL); Mexican-American-Middle (MM) and Lower (ML); and Anglo-Middle (AM) and Lower (AL). The stimuli were edited from the black-and-white videotaped interviews described in a preceding section.

Semantic differential. Scales selected for use were derived from Study I. The judgmental model of confidence-eagerness and ethnicity-nonstandardness was indexed by scales 1 to 5 and 6 to 10 respectively in Table III-1. Scales 11 to 15 were filler items. These scales were individually printed on Hollerith data cards which were prepunched to facilitate subsequent collation and scoring. Whereas the individually printed scale cards were used in a condition where the S saw videotaped stimuli, the scales used with stereotype descriptions were printed in randomized order on an 8½ by 11 inch page.

Stereotype descriptions. Six brief descriptions for the ethnic and status groups of children represented in the videotapes were composed for use in eliciting Ss' stereotypes. These are presented in Table III-2. All six descriptions were printed on a single page of a booklet and followed by six pages of semantic differential scales.

Procedure

In an average of one week's time before participation in the videotape presentation, Ss were visited in their classes and were requested to fill in the stereotype response booklets described above. This was also done between three and five days after the experiment. No mention was made of stereotyping and the like when Ss were requested to complete the response instrument. The explanation was simply that the researchers were attempting to study the degree to which teacher-Ss would agree in terms of the speech associated with types of children.

Videotape testing was undertaken individually with the Ss where each S was given the opportunity to control the stimulus presentation and the order in which he chose to fill in the 15 cards. This procedure was undertaken because of dual testing for another experiment (Study II).

Scale quantification was undertaken by assigning a one to seven value beginning with the adjective for each scale indicated in Table III-1. Factor scores were calculated as the

TABLE III-1

Scales used to index the two-factor model and fillers.

1. THE CHILD SEEMS: *reticent-to-speak--eager-to-speak
2. THE CHILD SEEMS: *hesitant--enthusiastic
3. THE CHILD IS: active--passive*
4. THE CHILD SEEMS TO: enjoy--dislike TALKING*
5. THE CHILD SEEMS: *unsure--confident
6. THE CHILD SEEMS CULTURALLY: *disadvantaged--advantaged
7. THE CHILD SOUNDS: Anglo-like--does not sound Anglo-like*
8. LANGUAGE SPOKEN IN THIS CHILD'S HOME IS PROBABLY: standard American style--marked ethnic style*
9. THE CHILD'S HOME LIFE IS: very similar--very different FROM YOURS WHEN YOU WERE HIS AGE*
10. THE CHILD'S FAMILY IS PROBABLY: *low-social-status--high-social-status
11. THE CHILD SEEMS TO BE: interested--uninterested IN HIS ENVIRONMENT*
12. THE CHILD SEEMS: intelligent--unintelligent*
13. THE CHILD SEEMS: *non-competitive--competitive
14. THE CHILD PROBABLY SPENDS: large--small AMOUNT OF TIME AWAY FROM HOME*
15. THE CHILD IS: determined--not determined IN SCHOOL*

*The asterisks define the pole of the scale assigned a value of 1.0 in the quantification scheme. The asterisks did not appear on the actual instrument.

TABLE III-2

Ethno-status stereotype descriptions.

Student A:	He is a Mexican-American boy who comes from a family of 10. His father is a gas station attendant. He lives in a lower class neighborhood.
Student B:	He is Anglo and lives with his mother who is a laundress and his 3 brothers and 1 sister in a lower class neighborhood.
Student C:	He is Anglo and lives in an upper middle class neighborhood. He is the son of a doctor and has two older sisters.
Student D:	He is Black and lives in an inner city ghetto area. His father is a truck driver and his mother is a factory worker. There are 8 children in the family.
Student E:	He is Mexican-American living in an upper middle class neighborhood. His father is a professor at the University and his mother is a high school teacher.
Student F:	He is Black and the son of a professor of sociology at the University. He lives in an upper middle class neighborhood.

sum of scales 1-5 for Factor I, confidence-eagerness, and 6-10 for Factor II, ethnicity-nonstandardness. The data were subjected to a two-by-three-by-three-by-two analysis of variance with the above summation score as the dependent variable.

Results

All four of the three-way interactions in the analysis of variance results were statistically significant. These results prompted a direct interpretation of the means involved in these interactions that would be pertinent to the purpose of the study. That is, what would be the differences, if any, between the responses to stereotype conditions and the videotape stimulus condition, and how would these differences (or lack of same) vary as a function of child ethnicity, child status, and judgmental factor? Figures III-1, III-2, and III-3 display the mean comparisons used to answer the above questions. Multiple mean comparisons incorporated protection levels based upon a modification of the Duncan Multiple Range procedure (Duncan, 1955), and significance was interpreted at the $p<.05$ level.

Black Children (Figure III-1)

For one thing, the middle and low status groups of children (videotape and stereotype stimuli) were rated significantly different from each other, and in the anticipated direction. Among the stimuli representative of low status conditions, there were no statistically significant differences among videotape and the two stereotype stimuli. In the case of the middle status cluster, however, the Factor II rating of the videotape of the middle class Black child was significantly different from the two stereotype ratings, which were not different from each other. This difference was in the direction of the videotape being rated as more ethnic and nonstandard than the stereotypes.

Mexican-American Children (Figure III-2)

Again, the two status groups were rated significantly different from one another and in the anticipated direction. In these results, however, both the middle status and lower status videotapes were rated significantly different from their stereotype counterparts. In the case of the middle status cluster, the videotape stimulus was rated lower in confidence-eagerness (Factor I) than the pre-exposure stereotype. In the lower status cluster, the videotape

Figure III-1

Two-factor display of mean ratings of label and stimulus conditions of lower and middle status Black children.

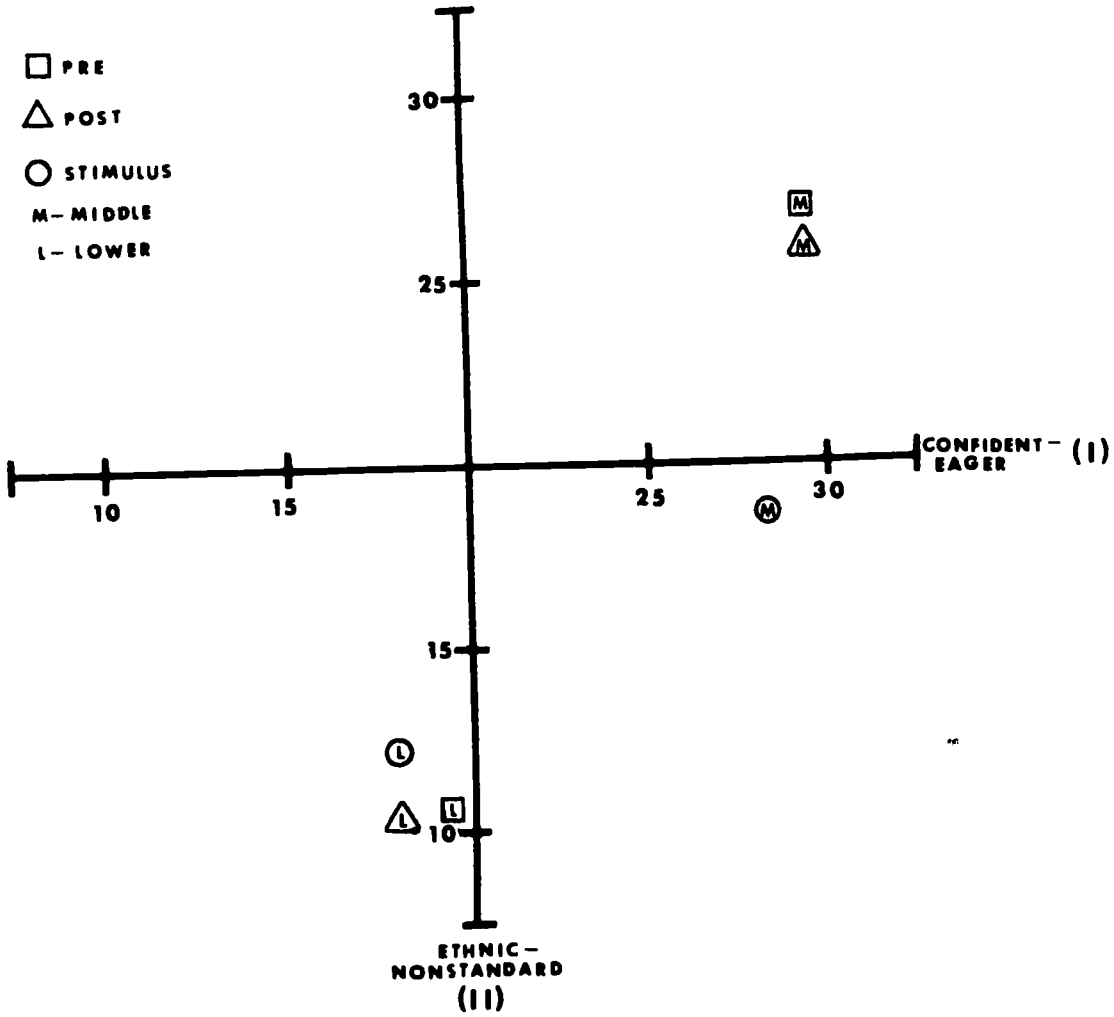
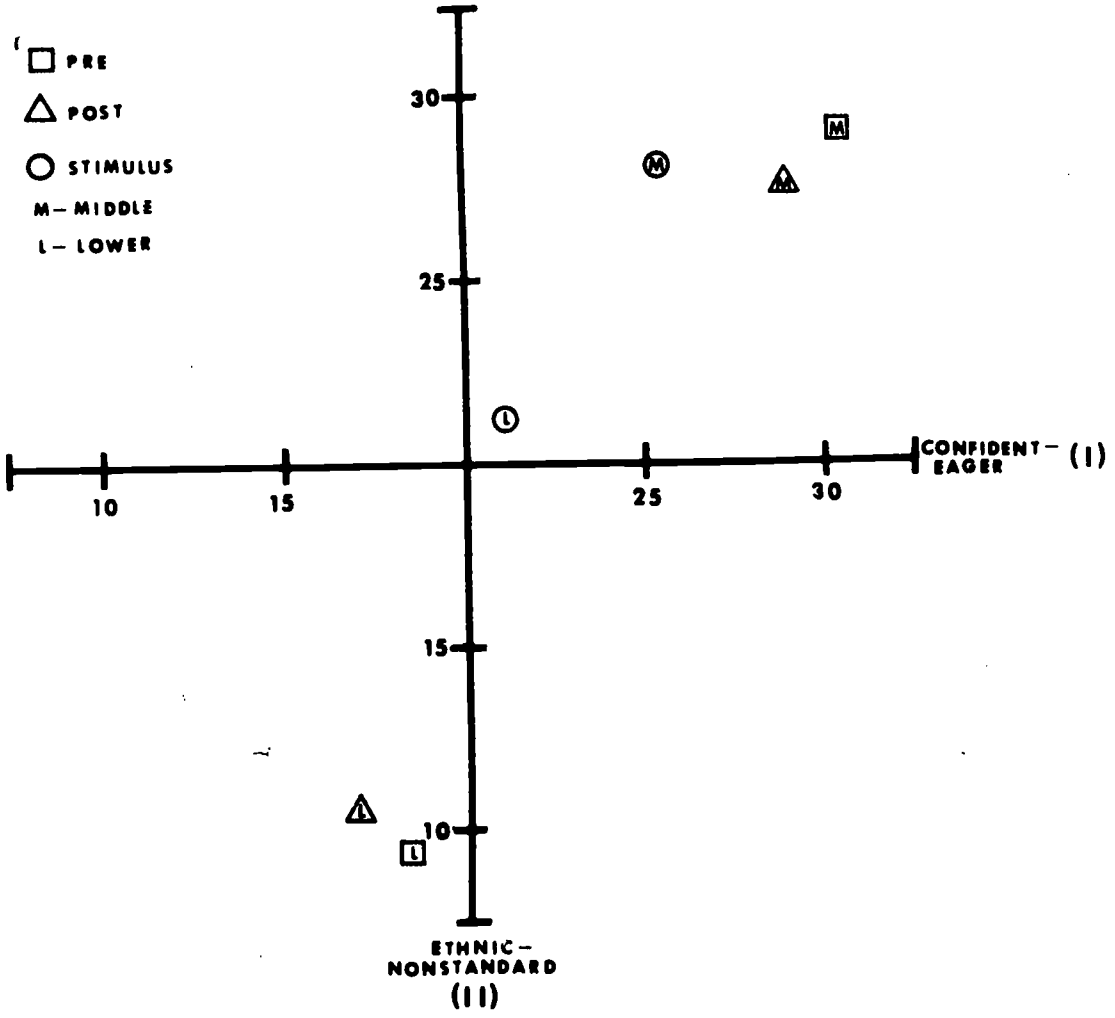


Figure III-2

Two-factor display of mean ratings of label and stimulus conditions of lower and middle status Mexican-American children.



stimulus was rated as less ethnic and nonstandard (Factor II) than the stereotype stimuli. In both status clusters, the two stereotype stimuli were not significantly different from one another.

Anglo Children (Figure III-3)

As in the above two analyses, the two status groups were differentiated in the anticipated direction. There were no statistically significant differences among the middle status stereotype and videotape stimuli. However, the videotape stimulus in the lower status cluster was rated significantly less ethnic and nonstandard (Factor II) than the two stereotype stimuli.

Some Overall Generalizations

In overview, across all three ethnic groups, status differentiations were significant, were in the anticipated direction, and had generality relative to the stereotype and videotape stimulus conditions. In all analyses, there were no significant differences between ratings of the pre- and post-stereotype stimuli.

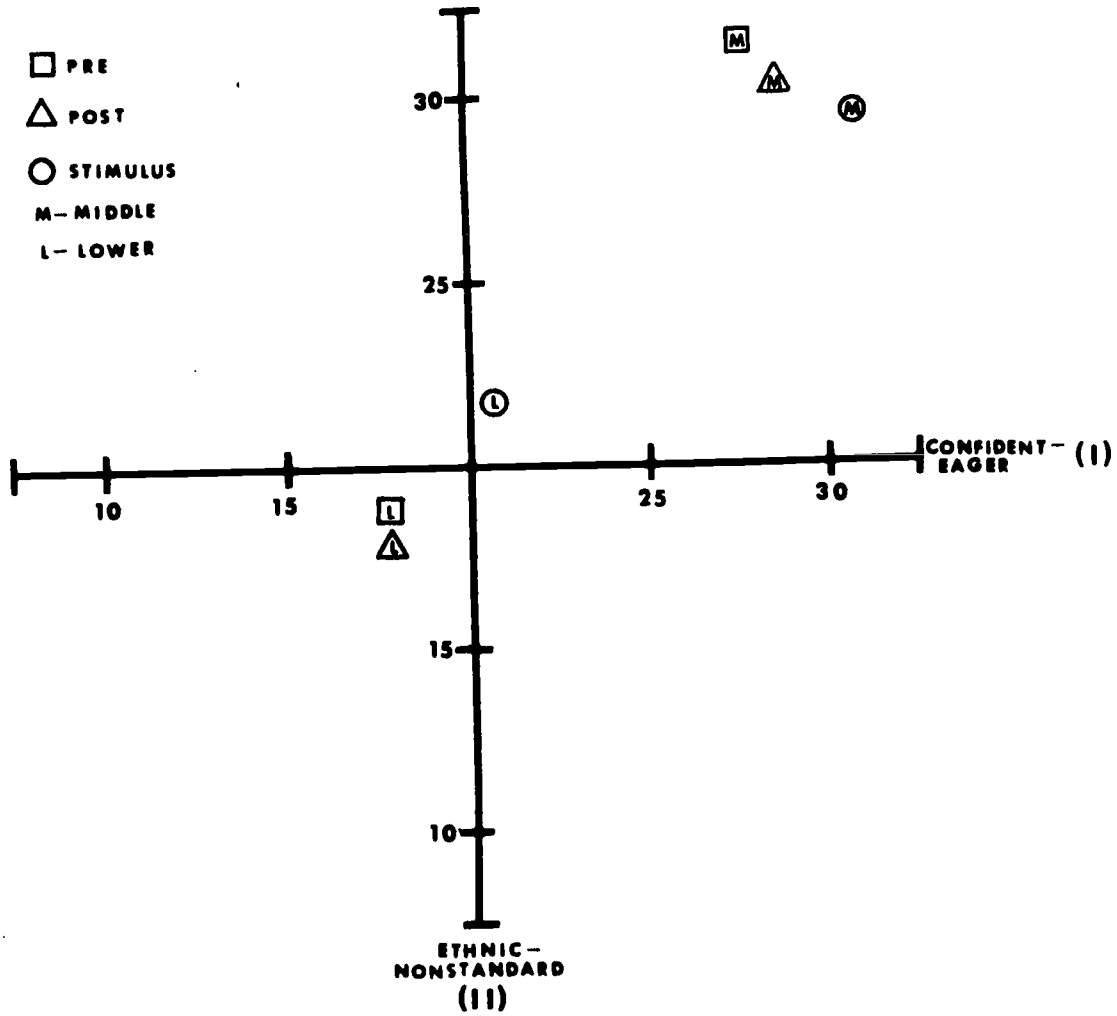
Discussion

The primary finding of this research was that with some exception, ratings of stereotyped descriptions of children and their videotape speech samples often corresponded in terms of the two judgmental dimensions, confidence-eagerness and ethnicity-nonstandardness. Although some differences between stereotype and videotape responses were found, this never involved more than one judgmental dimension in a particular comparison. In fact, assuming that the stereotypes were essentially the same between the pre- and post-conditions, this would make a total of 12 comparisons between videotape and stereotype, four of which showed significant differences. On the other hand, however, even when stereotype and videotape stimulus were different, both still showed significant contrasts in all cases between the middle and low status stimulus groups. In short, even though the videotape stimulus was sometimes discrepant from the stereotype stimulus, status differentiations were still maintained.

A secondary finding was the high constancy of an almost exact correspondence between pre- and post-stereotype ratings. This is evidence of the stability of such ratings across time (up to two weeks) and even in the face of

Figure III-3

Two-factor display of mean ratings of label and stimulus conditions of lower and middle status Anglo children.



intervening experience with the videotapes. An implication of this finding is that the stereotype may be quite resistant to change. However, use of the present two-factor model would provide a means for measuring such change as it might be prompted by the manipulation of experiences with videotapes or even the actual children.

STUDY IV: EFFECTS OF VISUAL CUES OF ETHNICITY UPON SPEECH RATINGS[1]

Purpose

Study III indicated a relation between teachers' ratings of speech imagined for written stereotypes of children and teachers' ratings of the videotaped speech of children who might fit those stereotypes. These findings suggested the presence of an ethnic stereotyping behavior on the part of teachers which may bias their perception of a child's speech. In the present research this implication was tested by pairing the same standard English speech recordings with videotapes of children from Anglo, Black, and Mexican-American families. Teachers' ratings of the speech samples were then assessed for variation due to the ethnicity of the child whom they were led to believe was doing the speaking. By also obtaining ratings of types of children, it was possible to see how stereotyping was related to biases when rating children from different ethnic groups.

Method

Subjects

Ss were 44 undergraduate education majors (42 female, 2 male) enrolled in a speech course for prospective elementary teachers.

Materials

Stimulus tapes. Fifth and sixth grade male children representative of three ethnic groups--Black, Mexican-American, and Anglo--from the Austin, Texas area were videotaped while in the process of assembling a plastic model car. Each child was asked to describe what he was doing or what he intended to do with his new car. A sound recording was made of each child's verbal response. From the black-and-white-videotapes,

[1]Conducted by Frederick Williams, Jack L. Whitehead, and Leslie M. Miller.

a 90-second segment of a child of each ethnicity was reproduced. Similarly, 90-second audio segments were reproduced of two Anglo children, one Black child, and one Mexican-American child. Slight lip movement in the video segments was detectable, but was insufficient for lip-reading on the part of the teacher-evaluators. In a testing design, teachers used a 15-scale semantic differential for rating videotapes of a Black child, a Mexican-American child, and an Anglo child. The video image of the Anglo child was always paired with the audio dub of an Anglo child speaking standard English. However, teachers also heard one of two Anglo tapes (necessary for the testing design) paired with the visual image of either a Mexican-American or a Black child. Eventually it was possible to make direct comparisons of the ratings of standard English passages as they were paired with an Anglo child, a Mexican-American child, or a Black child.

Response instrument. The set of semantic differential scales developed in Study I was used as the rating instrument. These scales were listed in Table III-1. After factor analyses substantiated the presence of the two judgmental dimensions, factor scores were calculated as the sum of scales 1-5 for confidence-eagerness, and 6-10 for ethnicity-nonstandardness.

Procedure

Two weeks prior to the videotape presentation, Ss completed a pretest response booklet consisting of (1) an explanation of the session under the guise of a two-part experiment designed to determine the correspondence of teachers' evaluations of children's speech with the prediction made by a standardized test battery of the child's intellectual ability; (2) instruction regarding the completion of the semantic differential; and (3) three evaluation forms requesting imagined ratings of a "Black child," "Anglo child," and "Mexican-American child."

For the videotape ratings, Ss were tested in groups of five or six persons in a small conference-type room, utilizing an 11-inch television monitor. Each group viewed (1) a Black or Mexican-American child whose nonstandard speech had been replaced by dubbing-in the speech of a standard English speaking child, (2) a Black or Mexican-American child speaking nonstandard English, and (3) an Anglo child speaking standard English. Teachers saw either the Black or Mexican-American child in the standard English version, but not both.

Data Analysis

Two types of analysis procedures were employed in the current study. First, to test whether the scale usages reflected the two dimensional model found in prior research, two factor analyses of the responses were performed. One was for scale usage with the label stimuli, the other was for usage with the videotape stimuli. In both, unities were placed in the diagonals of the correlation matrix, and factors with latent roots greater than one were rotated according to Varimax criteria. Second, to test for effects of the different ethnicity conditions, the data for each factor were subjected to two-by-three analyses of variance. Dimensions corresponded to the three ethnicity conditions and the two standard English speech samples which had been used. Dependent variables for the above analyses of variance consisted of summation scores for Factor I, confidence-eagerness and Factor II, ethnicity-nonstandardness.

Results

Judgmental Dimensions

Label judgments. Results of the factor analysis of scales used in judging labels are presented in Table IV-1. The ethnicity-nonstandardness factor was dominant, accounting for 33% of the total variance, while the confidence-eagerness factor accounted for 28%. In short the usual two dimensional model was found.

Videotape judgments. The second factor analysis results are shown in Table IV-2. Here a larger portion of the total variance is accounted for by the two-factor model, with the factor of confidence-eagerness accounting for 37% of the variance and ethnicity-nonstandardness for 31%. Again, the two dimensional model was found to obtain.

Stimulus Differentiation

Confidence-eagerness ratings. The analysis of variance of the dependent variable of confidence-eagerness revealed a significant main effect on the ethnicity variable, $F(2,82)= 9.52, p<.001$, and a significant main effect for the two different standard English passages used as stimuli, $F(1,82)= 11.62, p<.001$. There was no significant interaction.

Results of comparisons among the means of the three levels of the ethnic factor revealed that Mexican-American children were rated lower on confidence-eagerness than were

TABLE IV-1

Rotated factor matrix of teachers' responses to ethnicity labels.

Variables	Factors* I	II
1. unsure	.57	.53
2. passive	.12	.77
3. reticent	.19	.77
4. hesitant	.32	.79
5. dislike talking	.21	.70
6. lang. marked ethnic style	.75	.36
7. non-Anglo like	.79	.28
8. home life unlike yours	.89	.19
9. low social status	.88	.24
10. disadvantaged	.86	.27
(Percentage of total variance)	33%	28%

*Factor I = ethnicity-nonstandardness

Factor II = confidence-eagerness

TABLE IV-2

Rotated factor matrix of teachers' responses to stimuli.

Variables	Factors* I	II
1. unsure	.65	.37
2. passive	.80	.28
3. reticent	.89	.18
4. hesitant	.82	.30
5. dislike talking	.84	.23
6. lang. marked ethnic style	.25	.86
7. non-Anglo like	.25	.86
8. home life unlike yours	.37	.78
9. low social status	.35	.79
10. disadvantaged	.38	.83
(Percentage of total variance)	37%	31%

*Factor I = confidence-eagerness

Factor II = ethnicity-nonstandardness

Anglos or Blacks whose ratings were not significantly different from each other. The confidence-eagerness means (different alphabetical subscripts indicate significant differences at the $p<.05$ level) were:

Anglo = 23.1^{a}

Black = 22.2^{a}

M.-Am. = 18.7^{b}

Differences on the language sample variable indicated that one Anglo tape was generally rated higher (22.8) than the other (19.8). This difference had generality across the three ethnic categories (hence, the lack of a significant interaction).

Ethnicity-nonstandardness ratings. Results of the analysis of variance of ethnicity-nonstandardness ratings revealed a significant main effect on the ethnicity dimension, $F(2,82)= 31.8, p<.001$. No significant effects due to the two Anglo tapes nor an interaction were found.

Results of individual comparisons of the ethnic factor means indicated that Mexican-American and Black children were rated as more ethnic-nonstandard than the Anglos. The pertinent means (different alphabetical subscripts indicate significant, $p<.05$ differences) were:

Anglo = 30.1^{a}

Black = 21.8^{b}

M.-Am. = 21.3^{b}

Generalizations. In all, the findings indicated that the visual cues of ethnicity did bias teachers' evaluations of the same standard English language samples. This was in the direction of rating the Mexican-American child as being less confident and eager than the Black and Anglo children, and rating the Black and Mexican-American children as more ethnic-nonstandard than the Anglo child.

Stereotyping

Although the directions of the above biases seemed compatible with the findings of Studies I, II, and III, the interpretation that the biases were a function of stereotyping behavior on the part of the teachers was more an implication than a conclusion of this phase of the research. However,

there was a way to seek further support for this implication. This was done by considering the teachers' ratings of the children imagined for labels as a stereotyped response, then examining the degree to which variations in these stereotypes could be related to biases in rating the standard English samples as they were paired with the Black and Mexican-American children.

For each teacher-evaluator a difference score was calculated for confidence-eagerness and ethnicity-nonstandardness between ratings of the child imagined for the Anglo label and the children imagined for the Black and Mexican-American labels. These difference scores were assumed to constitute measures of stereotyping variations that the evaluators held for the three ethnic groups. The statistical question then was the degree to which these variations would predict the ratings of the Black and the Mexican-American children whose dubbed speech was standard English.[2] The correlations associated with four regression equations of the videotaped ratings upon the stereotype variations are summarized as follows:

	Black	Mexican-American
confidence-eagerness	-.01	-.46
ethnicity-nonstandardness	-.42	-.31

The above coefficients suggest that a small but reliable[3] relation exists between the stereotype biases and the ratings

[2]Presumably the prediction of difference scores between the ratings of the videotapes of the Anglo and Black children or the Anglo and Mexican-American children would be a more precise answer to the question. However, in the present testing design a given teacher's rating of the Anglo videotape and the minority group child with dubbed speech necessarily involved the use of different Anglo speech samples. Because some differences were found between these two samples it was reasoned that prediction of this type of difference score might be hampered by an additional source of variance.

[3]Because each teacher only saw one of two minority group children with a dubbed standard English audio tape, the number of replicates in the regression analyses was 22 rather than the total of 44 teachers. For d.f.=20, a correlation coefficient of .42 or greater could be interpreted as grounds to reject the null hypotheses that the correlation equaled zero. Negative coefficients were expected since generally the greater the difference in ratings between the Anglo and minority group children (as labeled), the lower the rating on the two judgmental dimensions when rating videotapes.

of ethnicity-nonstandardness for Black children and confidence-eagerness for Mexican-American children. Although the correlation was less, the same interpretation could tentatively be advanced for rating of ethnicity-nonstandardness of the Mexican-American children. No relation was found for predicting confidence-eagerness ratings of the Black children. This latter finding could be expected since it was found earlier that there were not significant differences on the confidence-eagerness dimension between the Anglo and Black children's videotapes.

In sum, visual cues of ethnicity did affect ratings of standard English speech samples, and to some degree this variation could be associated with stereotyping on the part of the teacher-evaluators.

STUDY V:
RELATIONS OF TEACHERS' ETHNICITY AND EXPERIENCE WITH SPEECH ATTITUDES AND EXPECTATIONS OF PUPIL PERFORMANCE[1]

Purpose

Previous findings in this series of studies indicated: (Study I) that a two-factor judgmental model was applicable to the evaluation of videotaped speech samples of Anglo, Black, and Mexican-American children, (Study II) that the two judgmental factors were exercised within a common time span, and (Study III) that the judgmental model could be applied to stereotyped written descriptions of children. Given these findings regarding the evaluation instrument, the present research shifted the emphasis to the relations of teacher ethnicity and experience to making speech evaluations, and the relations between speech attitudes and teachers' expectations of pupils' academic performances.

In overview, the research involved having in-service teachers evaluate videotapes taken to represent Anglo, Black, and Mexican-American children from the Central Texas area. Teachers were also asked to provide speech evaluations of children based upon their average experiences with Anglo, Black, and Mexican-American pupil groups. Here, teachers were given only a label ("Black") instead of a videotape as a stimulus. These label ratings were intended to index teachers' speech stereotypes of children in the three ethnic groups. For both videotape and label ratings teachers were also asked to assign a child to a particular class level (1=remedial, 2=below average ... 5=well above average) in the following subjects: art, grammar, physical education, social studies, mathematics, spelling, music, composition, and reading. These assignments were taken as an index of teachers' expectations of pupils' performance in the various subject matter areas. Given this body of rating data, specific questions of research were as follows:

[1]This study was conducted by Frederick Williams, Jack L. Whitehead, and Leslie M. Miller.

1. What are the relations of teacher experience and ethnicity to the differentiation of videotape speech samples of Anglo, Black, and Mexican-American children?

2. What are the relations of teacher experience and ethnicity to the differentiation of stereotype labels of Anglo, Black, and Mexican-American children?

3. To what degree can speech ratings based upon stereotypes be used to predict ratings of videotaped samples?

4. To what degree can speech ratings based upon either videotaped samples or stereotype labels be used to predict teachers' expectations of pupils' academic performances?

Method

Subjects

Evaluators in this research were in-service teachers who were surveyed in field work in the Central Texas area during winter, 1970-71. Of the total number of teachers included in the fieldwork, 193 were self-identified as Anglo, 77 as Black, and 18 as Mexican-American. Teachers also ranged in self-reports of experience ranging from less than one year to over thirty. A promise of anonymity prevents the identification of the teachers' schools, but 69 percent of the teachers were from towns of 35,000 and under, as against 31 percent from larger areas. Teachers were tested within the context of an approximately one-hour experimental in-service training program on the topic of "Language Differences in Children." The instructions for testing and the testing itself comprised the first segment of this period, and these were followed by instructional materials which were unrelated to the present research. Teachers' evaluation of the videotapes as well as the stereotypes were thus not confounded by the other materials of the in-service training activities. Such activities were conducted in 15 schools. Some 85% of the teachers were instructors in the primary grades, with a remainder being secondary school teachers.

Materials

Videotape samples. Language samples used in the field research were drawn from the collection of videotapes

described in the previous section of this report. In a given testing session, a group of teachers saw six different videotapes, each approximately two minutes in duration. These six videotapes represented a randomized sequence of Anglo, Black, and Mexican-American children each from the two status levels. Four different test sequences, each involving different children but representing the same ethnic and status categories, were prepared for the field-work. By systematically rotating the tapes for different test groups, data for the study represented a nearly equal frequency of administration of the four different tapes.

Evaluation form. Figure V-1 is an exact replica of the evaluation sheet that each teacher used for evaluating a videotape or for responding to a stereotype label. Briefly, this response sheet contains 15 speech evaluation scales, 10 of which are applicable to the two-factor model, and 5 of which are simply inserted as filler items. Scales 1, 8, 9, 12, and 14 were consistently identified in prior studies as indexing a confidence-eagerness factor. Scales 2, 6, 7, 10, and 13 have been identified as a factor of ethnicity-nonstandardness. The second part of the evaluation form contains the scales for assignments of a child to graded classes. As can be seen in Figure V-1, nine different academic subjects are listed. Each S had the opportunity to assign a child to graded classes of 1-5 in each of these subjects. Response materials were reproduced on Digitex forms which provided for automatic keypunching.

Ss' test booklets contained a first sheet for demographic data, a second sheet which explained the guise of the experiment, and the third sheet which explained how to use the semantic differential scales. Following these instructional forms were three evaluation forms, each with an ethnic label in the upper right-hand corner. The remaining six pages of the booklet were numbered 1-6 to correspond with the six videotapes that were to be presented to the Ss.

Procedures. Teachers were tested in subgroups ranging from as few as 4 Ss to as many as 30 in a given group. Testing typically took place in the schools and most often in space set aside in school libraries. Typically, two researchers conducted a testing session and were introduced to the teachers as the staff for the in-service training session. The videotapes and the booklets were introduced to the teachers as materials for the in-service training. After the staff was introduced, each session typically began with having the teachers fill in the information on the first pages of the test booklet, then having them read pages two- and three. After approximately five minutes, teachers were then asked if they had any questions about the instructions.

Figure V-1

EVALUATION FORM

1. Please give ratings of the child on the following scales:

[1] THE CHILD SEEMS: unsure ▯ ▯ ▯ ▯ ▯ ▯ ▯ confident

[2] THE LANGUAGE OF THE CHILD'S HOME IS PROBABLY: standard American ▯ ▯ ▯ ▯ ▯ ▯ ▯ marked ethnic style

[3] THE CHILD SEEMS: intelligent ▯ ▯ ▯ ▯ ▯ ▯ ▯ unintelligent

[4] THE CHILD PROBABLY SPENDS: little ▯ ▯ ▯ ▯ ▯ ▯ ▯ much time away from home

[5] THE CHILD SEEMS TO BE: sad ▯ ▯ ▯ ▯ ▯ ▯ ▯ happy

[6] THE CHILD SOUNDS: Anglo-like ▯ ▯ ▯ ▯ ▯ ▯ ▯ non-Anglo like

[7] THE CHILD'S HOME LIFE IS PROBABLY: like ▯ ▯ ▯ ▯ ▯ ▯ ▯ unlike yours

[8] THE CHILD IS: active ▯ ▯ ▯ ▯ ▯ ▯ ▯ passive

[9] THE CHILD SEEMS: reticent ▯ ▯ ▯ ▯ ▯ ▯ ▯ eager to speak

[10] THE CHILD'S FAMILY IS PROBABLY: low ▯ ▯ ▯ ▯ ▯ ▯ ▯ high social status

[11] THE CHILD IS: determined ▯ ▯ ▯ ▯ ▯ ▯ ▯ not determined in school

[12] THE CHILD SEEMS: hesitant ▯ ▯ ▯ ▯ ▯ ▯ ▯ enthusiastic

[13] THE CHILD SEEMS CULTURALLY: disadvantaged ▯ ▯ ▯ ▯ ▯ ▯ ▯ advantaged

[14] THE CHILD SEEMS TO: like ▯ ▯ ▯ ▯ ▯ ▯ ▯ dislike talking

[15] THE CHILD SEEMS: non-competitive ▯ ▯ ▯ ▯ ▯ ▯ ▯ competitive

▯ ▯ ▯ ▯ ▯ ▯ ▯

▯ ▯ ▯ ▯ ▯ ▯ ▯

▯ ▯ ▯ ▯ ▯ ▯ ▯

2. Now (using the code below), please assign the child to graded classes in the following subjects.

SUBJECT	Graded Class
ART	▯ ▯ ▯ ▯ ▯
GRAMMAR	▯ ▯ ▯ ▯ ▯
PHYS. EDUCATION	▯ ▯ ▯ ▯ ▯
SOCIAL STUDIES	▯ ▯ ▯ ▯ ▯
MATHEMATICS	▯ ▯ ▯ ▯ ▯
SPELLING	▯ ▯ ▯ ▯ ▯
MUSIC	▯ ▯ ▯ ▯ ▯
COMPOSITION	▯ ▯ ▯ ▯ ▯
READING	▯ ▯ ▯ ▯ ▯

GRADED CLASS CODE
1 = remedial class
2 = below average class
3 = average class
4 = above average class
5 = far above average class

DO NOT USE

IDENTIFICATION

A ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯
B ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯
C ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯
D ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯
E ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯
F ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯
G ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯
H ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯
I ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯
J ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯
K ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯
L ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯ ▯

UTILITY FORM 7188 OPTICAL SCANNING CORPORATION

1970-71

jec

TABLE V-1

Overall factor analysis of speech evaluation scales and class assignments.

Variables	Factors I	II	III
Speech Ratings			
unsure	.38	.15	-.70
marked ethnic style	.77	-.18	-.19
non-Anglo-like	.76	-.22	-.18
home life unlike yours	.73	-.13	-.25
passive	.17	.14	-.76
reticent	.27	.16	-.82
low social status	.74	.04	-.36
hesitant	.32	.23	-.77
disadvantaged	.73	.05	-.42
dislike talking	.16	-.06	-.79
Graded Classes			
art	.17	.68	-.03
grammar	.78	.34	-.20
physical education	-.09	.64	-.23
social studies	.64	.49	-.18
mathematics	.63	.48	-.16
spelling	.73	.41	-.17
music	.02	.73	-.09
composition	.73	.40	-.21
reading	.75	.40	-.23
(Percentage total variance)	(33%)	(14%)	(19%)

They were then told to proceed with marking the three sheets which contained the stereotype label stimuli. When it was observed that the teachers completed these first three sheets, they were then told that they would see a series of videotapes. The videotapes were played on a Sony AV3600 Videotape unit through an 11-inch monitor. Special care was taken that teachers could see and hear the tapes without interference. Following evaluations of the six tapes, the test booklets were collected. The staff reintroduced themselves then continued with an actual in-service training session which took advantage of the tapes that the teachers had just seen.

Preliminaries to the analyses. Prior to the analyses directed at the main questions of the study, one overall factor analysis of the intercorrelations among all variables on the response form was conducted. The purpose here was twofold. First was the necessary assessment to determine whether the two-factor model again would be found as in previous studies. A second reason was to provide a preliminary indication of the degree and type of relation between the data on assignments to graded classes and the speech evaluations. The overall factor analysis was conducted across all subdivisions of the data including teacher race and experience, and child ethnicity and status. Results are summarized as a rotated factor matrix presented in Table V-1.

The first interpretation of the factor analysis results was that the usual factors of confidence-eagerness and ethnicity-nonstandardness could be identified as factors I and III in the rotated matrix. Thus it was assured that the two-factor model was again relevant in the speech evaluations. Factor scores were calculated by summing over the scales mentioned earlier (under "Evaluation Form").

A second interpretation of the factor analysis was that assignments to the graded classes could be reduced from nine to three variables. The first of these variables was defined as assignment to language arts subjects--namely, grammar, spelling, composition, and reading. The average graded class assignment constituted the score on this variable. Second and third variables were constructed from classes presumed either related to language arts (social studies, mathematics) or relatively unrelated (art, music, physical education). Again the scores were calculated as average assignments to the graded classes.

The technical report of the analyses and results is next presented. (The reader who is uninterested in these details may skip to the "Discussion" section where the conclusions are presented in summary form.)

Analyses and Results

Relation of Teachers' Experience and Ethnicity to the Differentiation of Videotaped Speech Samples

Anglo and Black teachers. Because of the relatively low number (18) of Mexican-American teachers included in the survey, the main analysis of teacher ethnicity and experience was focused upon Anglo and Black teachers. By dividing the teachers into 5 experience categories (0-4 yrs., 5-9, 10-19, 20-29, >30) of 35 Ss each, it was possible to analyze teacher ethnicity and experience in the same overall analysis of variance. Each of the experience groups consisted of 10 Black teachers and 25 Anglos. The equalization of ethnicity ratios within each of these groups was accomplished by random selection.

Two four-way analyses of variance were calculated, one for confidence-eagerness as a dependent variable, the other for ethnicity-nonstandardness. Independent variables in both analyses were teacher experience and ethnicity, as well as the videotape differences of child status and child ethnicity. The detailed design of the analysis as well as the summary table of the results are presented in Table V-2.

Confidence-eagerness ratings. As shown in Table V-2 there were no differences due to teacher experience or ethnicity in ratings of confidence-eagerness. The only differences could be interpreted in the anticipated interaction between the child status and child ethnicity variables. A summary of the mean confidence-eagerness ratings found in that interaction are presented in Table V-3. As was expected from the selection of the videotapes, the low status children in each of the ethnic categories were rated as less confident and eager than the middle status children. A main source of the interaction appears to be due to the much larger difference in rating the confidence-eagerness of the two status groups of Black children as compared with the other two groups. Also there were differences among ethnic groups within the status categories--namely that the ratings in the low status group ranged from a low for the Black children to a high for the Anglo children, but in the middle status group the Mexican-Americans were lowest, and the Blacks the highest.

Ethnicity-nonstandardness. As shown in Table V-2, the variables of teacher ethnicity and experience were involved in higher order interactions with either child status, ethnicity, or both. Interpretations of the results were based upon examination of the mean ratings in these higher order interactions.

TABLE V-2

Analyses of variance of confidence-eagerness and ethnicity-nonstandardness ratings.

Teachers: Anglo, Black
Stimuli: Videotape

Source		Confidence-Eagerness		Ethnicity-Nonstandardness	
	d.f.	MS	F	MS	F
Between-Ss					
A (experience)	4	57.0	1.1	57.9	1.1
B (teach. ethnicity)	1	141.1	2.6	322.1	6.4**
AB	4	48.5	.9	25.1	.5
error	165	54.2		50.6	
Within-Ss					
C (child status)	1	5751.3	83.5**	1745.4	35.6**
AC	4	161.0	2.3	63.1	1.3
BC	1	32.2	.5	2.7	.1
ABC	4	46.1	.7	139.4	2.8*
error	165	68.8		49.0	
D (child ethnicity)	2	554.7	11.1**	6571.8	86.2**
AD	8	41.4	.8	46.7	.6
BD	2	17.1	.3	422.1	5.5**
ABD	8	20.7	.4	37.4	.5
error	330	50.1		76.2	
CD	2	928.6	27.7**	186.7	7.3**
ACD	8	25.7	.8	15.9	.6
BCD	2	96.8	2.9	59.6	2.3
ABCD	8	32.9	1.0	27.3	1.1
error	330	33.6		25.7	

**$p<.01$

*$p<.05$

TABLE V-3

Means of confidence-eagerness ratings in the interaction of child status and ethnicity.

Teachers: Anglo, Black
Stimuli: Videotape

Ethnicity:		Anglo	Black	Mex.-Am.
Status:	Low	23.0_c*	18.3_a	20.2_b
	Middle	26.1_d	27.6_e	23.3_c

*Means with common subscripts are not significantly ($p<.05$) different from one another.

Table V-4 shows the means for the three-way interaction of teacher experience, teacher ethnicity, and child status. With only few exceptions, the low status children were rated as more ethnic and nonstandard, (shown by lower factor scores) than middle status one. The magnitude of this difference varied somewhat according to the particular combination of teacher experience and ethnicity. Also there was some tendency for Black teachers to rate all children slightly higher than did Anglo teachers.

Table V-5 shows the mean ratings of ethnicity-nonstandardness within the significant interaction of teacher ethnicity and child ethnicity. Some generalizations of this table are as follows: Both Anglo and Black teachers similarly rated Anglo children as the least ethnic and nonstandard. Anglo teachers, however, rated Black children as more ethnic and nonstandard than did Black teachers. Also, Anglo teachers rated Black children as more ethnic and nonstandard than Mexican-American children, whereas for the Black teachers there was no significant difference in rating Black children and Mexican-American children in terms of this variable.

Table V-6 reports the means of ethnicity-nonstandardness ratings in the significant interaction of child status and child ethnicity. As assumed for the design of the study, low status children were generally rated as more ethnic and nonstandard than middle status children, and this differentiation held for each of the child ethnic subgroups. However, similar to the confidence-eagerness ratings, there was a greater magnitude of differentiation between the low and middle status Black children than within the other two ethnic groups.

Mexican-American teachers. Separate analyses of variance were conducted for the group of Mexican-American teachers for their ratings of confidence-eagerness and ethnicity-nonstandardness of the videotape stimuli. Each of these was a two-way analysis involving child status and child ethnicity. As shown in Table V-7 both analyses resulted in an expected child status by child ethnicity interaction, which was the main focus for interpretation (Tables V-8, V-9).

Confidence-Eagerness. Unlike the other two teacher ethnic groups, Mexican-American teachers differentiated only the Black children in terms of low and middle status (Table V-8). Status differences within the Anglo and Mexican-American child groups were not statistically significant. (In fact there were slight reversals in the expected status differentiations.)

Ethnicity-Nonstandardness (Table V-9). Again, Mexican-American teachers differentiated only the Black children in

TABLE V-4

Means of ethnicity-nonstandardness ratings in the interaction of teacher experience, teacher ethnicity, and child status.

Teachers: Anglo, Black
Stimuli: Videotape

Child status:	Low		Middle	
Teacher ethnicity:	Anglo	Black	Anglo	Black
Teacher experience:				
0-4 yrs.	16.3_{a}*	18.5_{abc}	20.1_{cdefgh}	22.0_{fgh}
5-9	20.0_{cdefg}	18.4_{abc}	19.1_{bcde}	21.6_{efgh}
10-19	16.8_{ab}	19.7_{cdef}	21.1_{cdefgh}	19.4_{bcdef}
20-29	18.8_{abcd}	18.3_{abc}	21.1_{cdefgh}	23.7_{h}
30	17.0_{ab}	19.9_{cdefg}	21.3_{defgh}	22.5_{gh}

*Means with common subscripts are not significantly ($p<.05$) different from one another.

TABLE V-5

Means of ethnicity-nonstandardness ratings in the interaction of teacher ethnicity and child ethnicity.

Teachers: Anglo, Black
Stimuli: Videotape

Teacher ethnicity:		Anglo	Black
Child ethnicity:	Anglo	25.8_c*	24.5_c
	Black	14.2_a	17.7_b
	Mex.-Am.	17.5_b	19.0_b

*Means with common subscripts are not significantly ($p<.05$) different from one another.

TABLE V-6

Means of ethnicity-nonstandardness ratings in the interaction of child status and ethnicity.

Teachers: Anglo, Black
Stimuli: Videotape

Ethnicity:		Anglo	Black	Mex.-Am.
Status:	Low	23.7_{d}*	13.7_{a}	17.7_{b}
	Middle	26.6_{e}	18.2_{bc}	18.9_{c}

*Means with common subscripts are not significantly ($p<.05$) different from one another.

TABLE V-7

Analyses of variance of confidence-eagerness and ethnicity-nonstandardness.

Teachers: Mexican-American
Stimuli: Videotape

Source		Confidence-Eagerness		Ethnicity-Nonstandardness	
	d.f.	MS	F	MS	F
A (status)	1	106.0	<1	11.3	<1
error	17	110.3		74.9	
B (ethnicity)	2	123.6	2.8	415.6	11.5**
error	34	43.9		36.1	
AB	2	415.6	11.1**	173.8	6.4**
error	34	37.3		27.1	

**$p<.01$

TABLE V-8

Means of confidence-eagerness ratings in the child status by ethnicity interaction.

Teachers: Mexican-American
Stimuli: Videotape

		Ethnicity		
		Anglo	Black	Mex.-Am.
Status:	Low	22.3_a*	19.6_a	23.8_a
	Middle	19.3_a	29.3_b	23.1_a

*Means with common subscripts are not significantly ($p<.05$) different from one another.

TABLE V-9

Means of ethnicity-nonstandardness ratings in the child status by ethnicity interaction.

Teachers: Mexican-American
Stimuli: Videotape

		Ethnicity		
		Anglo	Black	Mex.-Am.
Status:	Low	24.0_c*	13.8_a	22.9_{bc}
	Middle	22.1_{bc}	19.6_b	21.0_{bc}

*Means with common subscripts are not significantly ($p<.05$) different from one another.

the low and middle status groups. Mean ratings of low and middle status Anglo and Mexican-American children did not differ significantly.

Relations of Teacher Experience and Ethnicity to the Differentiation of Stereotype Labels of Anglo, Black, and Mexican-American Children.

Anglo and Black teachers. The aforementioned subdivision of Anglo and Black teachers in the five experience groups was used in the main analyses of stereotype ratings. Dependent variables were again ratings of confidence-eagerness and ethnicity-nonstandardness. Independent variables were teacher experience, teacher ethnicity, and child ethnicity. The detailed analysis model as well as the summary of results are presented in Table V-10.

Confidence-Eagerness. The only significant variation in ratings of confidence-eagerness was in terms of child ethnicity. As shown in Table V-11, comparisons of the mean ratings of confidence-eagerness indicated that the stereotype of Anglo children was rated as more confident and eager than Black children, and Black children more than Mexican-American children.

Ethnicity-Nonstandardness. As shown in Table V-10 two significant interactions were found--i.e., teacher ethnicity by experience, and teacher ethnicity by child ethnicity. Table V-12 shows the individual mean ratings of ethnicity-nonstandardness for the interaction of teacher experience and ethnicity. One source of the interaction appears to be that although Black teachers typically rated stereotypes as less ethnic and nonstandard than did Anglo teachers, this difference varied somewhat within the different experience groups. No other pattern appears for teacher experience difference.

Table V-13 shows the mean ratings of ethnicity-nonstandardness in the significant interaction of teacher ethnicity and child ethnicity. Interpretations are as follows: Both Anglo and Black teachers rated the stereotype of Anglo children as less ethnic and nonstandard than the other two child ethnic groups. Within the Anglo and Black teacher groups there were no differences between the ratings of Black and Mexican-American stereotypes. However, the Black teachers rated Black and Mexican-American stereotypes as less ethnic and nonstandard than did the Anglo teachers.

TABLE V-10

Analyses of variance of confidence-eagerness and ethnicity-nonstandardness ratings.

Teachers: Anglo, Black
Stimuli: Stereotype-Labels

Source		Confidence-Eagerness		Ethnicity-Nonstandardness	
	d.f.	MS	F	MS	F
Between-Ss					
A (experience)	4	11.4	.3	14.7	.5
B (teach. ethnicity)	1	.9	.0	407.1	14.1**
AB	4	68.6	1.8	70.0	2.4*
error	165	38.4			
Within-Ss					
C (child ethnicity)	2	1978.5	54.0**	4827.0	106.0**
AC	8	24.6	.7	22.7	.5
BC	2	100.8	2.8	465.7	10.2**
ABC	8	36.3	1.0	11.6	.3
error	330	36.6		45.5	

**$p<.01$

*$p<.05$

TABLE V-11

Means of confidence-eagerness ratings in main effect of child ethnicity.

Teachers: Anglo, Black
Stimuli: Stereotype-Label

Ethnicity:	Anglo	Black	Mex.-Am.
	26.0_a*	22.0_b	18.6_c

*Means with common subscripts are not significantly ($p<.05$) different from one another.

TABLE V-12

Means of ethnicity-nonstandardness ratings in the interaction of teacher experience and ethnicity.

Teachers: Anglo, Black
Stimuli: Stereotype-Label

Ethnicity:		Anglo	Black
Experience:	0-4 yrs.	17.1_{ab}*	19.8_{cd}
	5-9	17.2_{ab}	18.2_{abc}
	10-19	16.3_{a}	21.1_{d}
	20-29	18.1_{abc}	19.1_{bcd}
	30	18.5_{abc}	18.8_{bc}

*Means with common subscripts are not significantly ($p<.05$) different from one another.

TABLE V-13

Means of ethnicity-nonstandardness ratings in the interaction of teacher ethnicity and child ethnicity.

Teachers: Anglo, Black
Stimuli: Stereotype-Label

Teacher Ethnicity:		Anglo	Black
Child Ethnicity:	Anglo	26.2_d*	24.0_c
	Black	13.7_a	17.7_b
	Mex.-Am.	12.5_a	16.5_b

*Means with common subscripts are not significantly ($p<.05$) different from one another.

Mexican-American teachers. As in the videotape analyses, separate analyses were run for Mexican-American teachers' ratings of the stereotypes. Again, dependent variables were confidence-eagerness and ethnicity-nonstandardness. The results of two, one-way analyses of variance across the child ethnicity categories are presented in Table V-14. In both of these analyses significant differences were found across the child ethnic categories. Comparisons of the individual means are presented in Table V-15.

Mexican-American teachers rated Anglo stereotypes as more confident and eager and less ethnic and nonstandard than the children in the other two groups. They did not differentiate between Black and Mexican-American stereotyping in ratings of confidence-eagerness or for ethnicity-nonstandardness.

The Predictability of Videotape Speech Ratings Based upon Stereotype Ratings

Data and analyses. The statistical model for this analysis was the regression of the videotape ratings upon stereotype ratings. Ratings by the Anglo and Black teachers as used in the ethnicity by experience analyses (Table V-2) were employed. Because the stereotype ratings were not differentiated by status as were the videotapes, analyses were done separately of the predictability of the low status videotape ratings from the stereotype ratings and again for the prediction of the middle status videotape ratings from the stereotypes. Four such analyses were conducted, two each for the confidence-eagerness and ethnicity-nonstandardness dimensions, and within these, one each for the low and the middle status videotape ratings. Results of the equation were interpreted upon the basis of the derived correlation (R) and determination coefficient (R^2), as given in Table V-16.

Confidence-Eagerness. Only within the low status group did stereotype ratings of confidence-eagerness show any relationship with videotape ratings. Although statistically significant, the degree of correlation was interpreted as a slight, almost negligible, relationship.

Ethnicity-Nonstandardness. In both the low and middle status analyses, a significant correlation was found between the stereotype and videotape ratings. Although the degree of correlation was only moderate, it was nevertheless large enough for both groups to consider that an interpretable relationship could be identified between the stereotype and videotape ratings.

TABLE V-14

Analyses of variance of confidence-eagerness and ethnicity-nonstandardness ratings.

Teachers: Mexican-American
Stimuli: Stereotype-Label

Source		Confidence-Eagerness		Ethnicity-Nonstandardness	
	d.f.	MS	F	MS	F
Child Ethnicity	2	301.2	10.1**	776.4	26.2**
error	34	29.7		29.6	

TABLE V-15

Means of confidence-eagerness and ethnicity-nonstandardness ratings.

Teachers: Mexican-American
Stimuli: Stereotype-Label

Child Ethnicity:	Anglo	Black	Mex.-Am.
Confidence-Eagerness:	26.1_a*	18.9_b	19.1_b
Ethnicity-Nonstandardness:	25.9_a	15.6_b	13.7_b

*Means with common subscripts across rows are not significantly ($p<.05$) different from one another.

TABLE V-16

Prediction of videotape ratings from stereotype ratings.

Rating/Subgroup	a	b	R	R^2
Confidence-Eagerness				
(low status)	15.8	.200	.19**	.037
(middle status)	22.9	.042	.04	.002
Ethnicity-Nonstandardness				
(low status)	9.2	.500	.48**	.232
(middle status)	12.1	.389	.40**	.157

**$p<.01$ with d.f.=503.

The Predictability of Teachers' Expectations

Data and analyses. The statistical model in these analyses was the regression of the class assignment variables each upon the predictor variables of confidence-eagerness and ethnicity-nonstandardness. Six two-variable equations were calculated, two each for the dependent variables of graded assignment to language arts classes, math-social studies, and music-art-P.E. Within each of these pairs one equation was based upon videotape ratings, the other upon the ratings of stereotypes. Results (Table V-17) of the equations were interpreted in terms of the multiple correlation coefficient (R), determination coefficient (R^2), and the relative proportions of variance predicted by the two predictor variables.

Videotapes. Graded class assignments in all three areas could be predicted to different degrees upon the basis of speech ratings. Assignment in language arts was the best predicted, and ratings of ethnicity-nonstandardness contributed almost twice as much to this prediction as did ratings of confidence-eagerness. Assignments to the math-social studies area were also well predicted by the speech ratings, but the degree of prediction was less than for language arts; further, ratings of confidence-eagerness and ethnicity-nonstandardness were more equal predictors than for the language arts subjects. Assignments to the music-art-P.E. combination could be predicted upon the basis of speech ratings, but substantially less so than the other two subject areas. In this prediction, however, a marked characteristic was that most of the predictability was due to confidence-eagerness ratings as compared with ethnicity-nonstandardness.

Stereotypes. Generally, the prediction of class assignments based upon stereotype speech ratings was similar to the findings for videotape ratings, although the degree of predictability was somewhat less. This was especially so in the prediction of assignments to the music-art-P.E. combination where the degree of prediction, although statistically significant, was so low as to be negligible. For both the language arts and the math-social studies predictions, ethnicity-nonstandardness was the more salient predictor.

Discussion

Results Summary

1. What are the relations of teacher experience and ethnicity to the differentiation of videotape speech samples of Anglo, Black, and Mexican-American children? Findings:

TABLE V-17

Predictions of graded class assignments from confidence-eagerness and ethnicity-nonstandardness ratings.

Stimulus Class	R	R^2	Relative Contributions	
			Confidence-Eagerness	Ethnicity-Nonstandardness
Videotapes				
Language Arts	.70**	.488	.17	.32
Math-Soc. St.	.61**	.377	.16	.21
Music-Art-P.E.	.36**	.129	.12	.01
Stereotypes				
Language Arts	.65**	.420	.08	.34
Math-Soc. St.	.54**	.291	.07	.22
Music-Art-P.E.	.19**	.038	.04	.00

**$p<.01$ for d.f.=1156 (videotapes) or 577 (stereotypes)

(A) Status and ethnic differentiations of children on ratings of confidence-eagerness appear independent of variations in teacher experience and ethnicity (Anglo, Black).

(B) Although Black teachers tend to rate children as slightly less ethnic-nonstandard than do Anglo teachers, this difference tends to vary unsystematically within particular experience levels of Black and Anglo teachers.

(C) Black and Anglo teachers both tend to rate the Anglo child as the least ethnic-nonstandard; however, Black teachers rate Black children as less ethnic-nonstandard than do Anglo teachers.

(D) Mexican-American teachers tend to make status differentiations on confidence-eagerness and ethnicity-nonstandardness ratings only in the case of Black children, and not for Anglos or Mexican-Americans. (There were no experience analyses for these teachers.)

2. What are the relations of teacher experience and ethnicity to the differentiation of stereotype labels of Anglo, Black, and Mexican-American children? Findings:

(A) Ratings of confidence-eagerness of Anglo, Black, and Mexican-American child stereotypes appear independent of teacher experience and ethnicity (Anglo, Black).

(B) Black teachers tend to rate children's stereotypes as less ethnic-nonstandard than do Anglo teachers; however, this difference tends to vary unsystematically within particular experience levels of Black and Anglo teachers.

(C) Both Anglo and Black teachers rate the Anglo child stereotype as less ethnic-nonstandard than the other two child ethnic groups; however, Black teachers tend to rate the Black and Mexican-American children's stereotypes as less ethnic-nonstandard than do the Anglo teachers.

(D) Mexican-American teachers tend to rate the Anglo stereotype as more confident-eager and less ethnic-nonstandard than they rate the Black and Mexican-American child; they do not differentiate the latter two from each other on either rating dimension.

3. To what degree can speech ratings based upon stereotypes be used to predict ratings of videotaped samples? Findings:

(A) There is low, if not negligible, predictability of confidence-eagerness ratings of videotapes based upon ratings of stereotype labels.

(B) There is a moderate degree of predictability of ethnicity-nonstandardness ratings of videotapes based upon ratings of stereotype labels.

4. To what degree can speech ratings based upon either videotaped samples or stereotype labels be used to predict teachers' expectations of pupils' academic performances?

ndings:

(A) Assignments to graded classes in the language arts area can be reliably predicted upon the bases of speech ratings of either videotapes or stereotype labels; ratings of ethnicity-nonstandardness contribute more than ratings of confidence-eagerness to this prediction.

(B) Assignments to graded classes in the math-social studies area can be reliably predicted from videotape or stereotype speech ratings; for videotapes, ratings of confidence-eagerness and ethnicity-nonstandardness contribute about equally to the prediction.

(C) Assignments to graded classes in music-art-P.E. can be somewhat predicted from videotape speech ratings but only negligibly from stereotype ratings; ratings of confidence-eagerness are the more salient predictor.

Additional Generalizations

Several additional observations transcend the main findings of the study. For one, it was assumed in selecting the videotape samples that differentiations of child status and ethnicity would reveal rather consistent patterns. This was borne out in Studies I through IV, and was again shown in the results of the field investigation. Although some ratings were found to interact with teachers' ethnic or experience characteristics, the overall pattern of child differentiations generally held (perhaps with the exception of the Mexican-American teachers).

A second observation was the striking parallel between overall patterns of differentiation involving the videotapes as compared with the stereotype labels. Such parallel extended even to the interactions (or lack of same) between patterns of child differentiation and the teacher variables of ethnicity and experience.

Finally, there was clear evidence that speech ratings could be associated with academic expectations in the various subject areas. That the teachers differentiated among the subject areas provided some evidence of their face validity.

However, the most supportive evidence of the association between speech attitudes and academic expectation came in the differential--and interpretable--associations between the two rating dimensions and the three subject matter areas. It seems particularly clear that attitudes about the ethnicity-nonstandardness of a child's speech are related to academic expectations in language arts subjects and even related somewhat to expectations in subjects such as math and social studies where language is more a means than a goal of instruction.

CONCLUSIONS AND IMPLICATIONS

Results of the experiments are interpreted here relative to the four main questions of the project.

1. Can the two-factor judgmental model of confidence-eagerness and ethnicity-nonstandardness be replicated with the use of videotapes and with the use of other teacher and pupil populations?

Results

As reported in Study I, a two-factor model very similar to the one found in the earlier audiotape studies (Williams, 1970) was found. The scales for the two factors were again obtained by a process of analyzing teachers' descriptions of children's videotaped speech, taking adjectives from these descriptions, then subjecting the adjectives to use in a prototype semantic differential instrument. Factors clearly identified as confidence-eagerness and ethnicity-nonstandardness were found in separate analyses involving their use with audio-visual presentations, visual-only presentations, and audio-only presentations. Although there was some evidence of modality-scale interaction, the two-factor model obtained in each of the three presentation modes. Additionally in Study I, the two factors were found in separate factor analyses of Black, Anglo, and Mexican-American child groups.

In Study III the two-factor model was found applicable when evaluations were made of the stereotyped written descriptions of Anglo, Black, and Mexican-American children. In Studies IV and V, the two-factor model also applied in the differentiation of children identified only by stereotype labels ("Anglo," "Black," "Mexican-American").

In all studies, the anticipated status and ethnic differentiations on the two factors could be taken as evidence of their validity. Such evidence was particularly persuasive in Study V where the two-factor model was used by Anglo, Black, and Mexican-American teachers representing experience groups ranging from one to over thirty years of teaching.

In sum, the conclusion was that the two-factor judgmental model of confidence-eagerness and ethnicity-nonstandardness could be replicated with the use of videotape language samples and with the teacher and pupil populations studied in the present project.

Discussion

There is reason to believe that the two-factor judgmental model is compatible with several of the performance continua described in social dialect research. In this latter area of research, the social dialect continuum itself is typically found to reflect differences in linguistic usages--as in the pronunciation /r/ in New York City (Labov, 1966) or in the use of multiple negation by children in Detroit (Shuy, 1970). These linguistic differences (i.e., "variables") are said to vary along a "linguistic continuum." This latter continuum may often represent the range of variation from the most prestigious dialect (or in Stewart's (1965) terms "acrolect") to the least prestigious one ("basilect"). It seems reasonable to assert that the present judgmental dimension of ethnicity-nonstandardness is simply a definition of this continuum in the attitudes of the teacher-evaluators.

Another continuum discussed (Labov, 1966; Shuy, et al., 1967; Shuy, 1970) in social dialect research reflects variations in the formality of the linguistic situation. Here the distinction has typically been made between the extremes of "careful" and "casual" styles of speech. It seems worth speculating that the judgmental dimension of confidence-eagerness reflects at least one of the key performance manifestations of the above stylistic continuum--i.e., fluency. Presumably, an exact attitudinal reflection of careful as against casual speech would incorporate consideration of both dialect features as well as fluency. These are the two characteristics found typically to vary according to the formality of speech situations. What the present results suggest is that the main attitudinal correlate of such variation is primarily a fluency judgment which may have an appreciable degree of independence from dialect judgment per se.

Further research could, no doubt, ferret out an answer to the question of whether variations in casual-careful speech (where dialect is possibly held constant) would result in judgmental variations only upon the confidence-eagerness dimension. In a broader sense, this same question suggests that studies be done to investigate the relationship between the judgmental dimensions and the characteristics of speech situations. In the present project, the speech situation was the somewhat formal context of an interview. Presumably, if the same children had been evaluated in peer speech situations, there probably would be an increase in fluency as well as a tendency for "less careful" speech. Both such variations should have their attitudinal correlates in ratings of confidence-eagerness and ethnicity-nonstandardness.

There is the obvious question of the generality of the two judgmental dimensions in populations other than teachers and pupils. The above reasoning suggests that the two judgmental dimensions might have as much generality as the dialect and stylistic (formality) continua just described. In short, ethnicity-nonstandardness and confidence-eagerness may be the attitudinal side of the coin of social dialects and language styles. If this is the case, then the two judgmental dimensions should be found in populations beyond teachers and pupils as well as in language communities other than English.

2. Does the judgmental process tend to reflect a social stereotype elicited quickly and by only a few cues, or does it reflect detailed perception of whatever is presented as representative of a child?

Results

Evidence from the studies indicates that the judgmental process is probably a combination of stereotype and evaluation behavior. In Study II, Ss, when given control over the amount of videotape playback, did provide themselves with an average of approximately one and one-half minutes exposure for purposes of making evaluations. In this same study, it was also found that there was a slight, but statistically nonsignificant, bias toward making ethnicity-nonstandardness ratings prior to those of confidence-eagerness. This order could have been a reflection of the greater frequency of dialect cues than fluency cues, thus suggesting that attention was given to the sequence of videotape cues in making judgments.

On the other hand, in most of the studies, and especially in the fieldwork of Study V, it was noted that teachers sometimes rapidly filled out the scales without waiting for the entire tape playback. This suggests that some teachers, at least, made judgments based upon preliminary cues, and that these cues may have been all that was necessary to elicit a stereotype basis for their judgments. That stereotypes do operate in the judgmental process was directly shown in Studies III, IV, and V where speech ratings were obtained without using speech samples as stimuli. In Study III, it was found that ratings could be obtained in response to a written description of a particular type of child, and in Studies IV and V, ratings were obtained in response to stereotype labels. It seems reasonable to assume that such descriptions and labels did elicit stereotypes and that the stereotypes were the bases for ratings on the two-factor judgmental model.

That the foregoing stereotype judgments were at least partially relevant to actual videotape judgments was shown particularly well in Study IV. Here, it may be recalled, the differences in rating stereotype labels of Anglo children as compared with Black and Mexican-American children could be used to predict biases found in ratings of standard English samples which were paired with video images of Black and Mexican-American children. The "misdirection" found in the latter ratings seems to be a direct reflection of stereotyping behavior.

Study V also provided evidence of the relevance of stereotyping in videotape judgments. It was found that ratings of ethnicity-nonstandardness for stereotype labels of Anglo, Black, and Mexican-American children were partial, yet statistically significant, predictors of the same ratings of the videotape stimuli.

In conclusion, stereotyping as well as stimulus evaluation appear to both enter into the judgmental process. Given minimal cues, it may be that the stereotype is relied upon more heavily in making evaluations than when there is the opportunity to experience a relatively detailed series of cues pertaining to a child's language performance.

Discussion

Although it is felt that the present research provided substantial evidence of the presence of stereotyping in language evaluations, the detailed role of the stereotype in such evaluations was not sufficiently revealed. The speculation that early forced judgments will result in more stereotyping than judgments based upon a continuous input of cues should be subjected to direct study.

The stereotyping process itself suggests the need for multiple avenues of investigation. Based upon findings of social dialect research as well as the earlier attitude study (Williams, 1970) where language characteristics were used to predict ratings, it seems clear that specific dialect and performance characteristics can be identified as cues in the judgmental process. One question is whether certain of these language cues particularly invite a stereotyped evaluation rather than careful attention to further details in rating a child. The research by Shuy, et al., (1969) suggests that it is the socially stigmatized cues that prompt the judgment of a speaker as being "low class," whereas it is the lack of such cues that may be the main basis for judging a speaker as "middle class." Perhaps, then, some language cues are particularly stigmatized so that they elicit social stereotypes and may direct the evaluator's attention away from

further cues. By manipulation of speech cues available for language judgments, it should be possible to determine which cues, if present, may lead to judgments that are more predictable upon the basis of stereotypes than upon what actually appears in the stimulus materials.

The above also suggests the need to study some of the visual correlates of social stereotyping. As found in Study IV, the video images did impose ethnic biases upon speech ratings. What are the detailed visual cues which stimulate stereotyping? Are they static features of ethnic identification such as skin color, hair texture, or facial structure, or are they dynamic behavioral characteristics such as eye contact or gestures? Perhaps they are some combination of the two. Just as research may be conducted where speech characteristics are varied, it is possible to vary visual cues (perhaps through sketches) and to determine which cues are the best predictors of stereotyped responses.

As suggested in the study by Seligman, et al. (unpublished), a child's written performances may serve as a basis for social stereotyping. Relative to the present research, one question is whether the judgmental dimension of ethnicity-nonstandardness might apply to written performance as it does to spoken performance, and whether children would be rated the same way if evaluated upon the basis of their writing as well as their speech. This, of course, opens up the practical series of questions about how teacher attitudes are elicited by, and in turn may affect, a child's written performances in the classroom.

There is the final, if not the most important, consideration of the personality, background, and other individual correlates of social stereotyping. A very practical issue for teacher education is the degree to which social stereotyping can be identified in a teacher or teacher-candidate. There is also the important question of how such behaviors may be modified in the process of teacher training, particularly when a teacher will be dealing with ethnic and status groups unfamiliar to her. It is felt that the present measurement procedures as well as research strategies such as the ethnic guise technique could be used to indicate stereotyping behavior in individuals, as well as to index whatever changes in such behavior might result from training experiences.

3. What are the relationships of teacher characteristics to judgments of children's speech?

Results

As found in Study V, teacher experience was unrelated to the major patterns of status and ethnic differentiation of the Anglo, Black, and Mexican-American children in the study. This was the case in judgments of videotape stimuli as well as stereotype labels. The only differences interacting with experience were in the slight variation between the Black and Anglo teachers' patterns in rating low and middle status children. These differences were negligible relative to the interpretations of the foregoing patterns.

Teacher ethnicity was related to speech judgments as reflected in the following patterns: Although both Anglo and Black teachers rated Anglo children as the least ethnic-nonstandard, Black teachers rated Black and Mexican-American children as less ethnic-nonstandard than did Anglo teachers. Subject to the hazard of Mexican-American teachers being under-represented in the sample, they tended to differentiate only the status of Black children on the confidence-eagerness and ethnicity-nonstandardness dimensions.

In conclusion, teacher ethnicity, but not teachers' amounts of experience, is related to judgments of children's speech.

Discussion

Admittedly, the researchers anticipated differences due to teacher experience, and some evidence of this was found in the early stages of the data analyses. However, the overall results of the project show no such relationship. The anticipation of the relationship was based on the idea that teachers who were inexperienced might have relatively negative stereotypes of children from ethnic and status groups other than their own. This reasoning was a reflection of Lambert's (1967) speculations on ethnocentrism and language attitudes. That is, the more a person lacked experience with communities other than his own, the more he would tend to see them in subjective-evaluative, rather than an objective-evaluative view. Also, the opposite was envisaged--that is, that experienced teachers would be highly sensitive to individual differences in the types of children which they had taught. In the present analyses, the foregoing lines of reasoning were not found to result in any experience relations with the overall patterns of ethnicity and status differentiations. It may be, however, that the experience variable correlated with patterns other than those investigated in the present analyses. For example, further analyses of the same data can be employed to answer the question whether the predictability

of videotape ratings from stereotype ratings varies as a function of teacher experience.

There is also the possibility that the details of experience--i.e., specific backgrounds--are more relevant to speech judgments than simply amounts of experience. In retrospect, the notion of "amount" seems like a relatively gross variable. Another further suggestion for subsequent analyses of the data is to undertake a factor analysis of the teachers to see whether they can be statistically separated into groups upon bases of commonality in their rating behaviors. It may be recalled that this was done in an earlier study by Naremore (1969; in press) where it was found that subgroups were often distinguished not only by ethnicity of teacher, but also by commonality in rating children of different ethnic groups. Such analyses of the present data might reveal further teacher characteristics related to judgments of children's speech.

Research should also be directed in a more experimental way into studying possible relations between manipulated experiences and judgment behaviors. Thus, to pursue the ethnocentrism thesis, as a teacher candidate is administered increasing amounts of experience with children of a particular type, one should find an increase in sensitivity to individual differences. This sensitivity can be measured by use of the present two-factor model and by such stimulus manipulations as involved in the ethnic guise technique. At the same time, it would probably be important to incorporate personality variables into the study. What would be the relationship, for example, between dogmatism of the teacher and changes in rating behavior? Such experiments not only promise answers to the above questions, but will serve as practical research into possible innovations in teacher training. The present series of studies suggests that speech evaluation could well be a measurable behavioral objective in teacher education.

4. What type of preliminary evidence might be obtained of the relation between judgments of speech and a teachers' expectations of academic performances?

Results

From Study V, it was clearly evident that speech judgments could be used to predict, to a substantial degree, assignments to graded classes in the language arts area, in areas partly related to language arts, and to some degree in areas presumably unrelated to language arts. It may be recalled that teachers had assigned children to graded

classes in nine subjects, and that the factor analysis of these assignments indicated some independence between language arts and non-language arts subjects. The category of "partly related to language arts" was comprised of assignments to math and social studies classes which were correlated with the language arts subjects, although not as highly as the relatively "pure" language arts subjects were intercorrelated among themselves. Ratings of ethnicity-nonstandardness, followed by confidence-eagerness ratings, provided the greatest degree of prediction in the case of assignments to language arts subjects. As might be expected, these language ratings were somewhat less salient predictors of assignments in the math and social studies class areas, and still less for the non-language arts classes (art, music, P.E.). In the case of these latter classes, the confidence-eagerness ratings rather than ethnicity-nonstandardness were the better predictor. It was found also that predictions involving speech judgments and graded classes obtained both in the ratings of videotape stimuli and of the stereotype labels.

In conclusion, judgments of speech could be used to predict teachers' assignments to graded classes, and particularly so for classes in the language arts area.

Discussion

As was reflected upon in the introduction to this research project, the concept of speech attitudes was a likely part of the overall process of self-fulfilling prophecies in the classroom as described by Rosenthal and Jacobson (1968). The connection between the teachers' evaluations of children's speech and the Rosenthal and Jacobson thesis was sought in the predictability of teachers' academic expectations of a child based upon the same teachers' attitudinal reactions to the child's speech. The results of Study V provide strong, yet preliminary evidence, of this predictability and, hence, of the relation between the present research into speech attitudes and self-fulfilling prophecies in the classroom. One question is, of course, whether speech attitudes themselves become a self-fulfilled prophecy in the classroom or whether speech attitudes may simply contribute to overall attitudes about a child's academic performance.

A teacher may not only expect a minority group child to remain fixed in his dialect, but, moreover, she might not be sensitive to change even if it did come about. As shown in Study IV, using the ethnic guise technique, the standard English performance, when paired with a minority group child, was judged differently than when paired with an Anglo child.

Like the other implications of this research, the relationship between speech attitudes and academic expectations strongly suggests the need to incorporate speech evaluation into the curricula of teacher training. Lest speech attitudes serve as false inputs into the overall process of self-fulfilling prophecies, or worse yet, become themselves self-fulfilled prophecies, teachers should be trained to be sensitive to the variations in social dialects and the variables of performance. Speech evaluation, which incorporates the attitudinal side of the social dialect coin, should be included as a part of the teacher training process. The present project not only points to the need for such training, but provides a number of ideas for implementation and evaluation.

REFERENCES

Anisfeld, E. and Lambert, W.E. Evaluational reactions of bilingual and monolingual children to spoken language. Journal of Abnormal and Social Psychology, 1964, 69, 89-97.

Anisfeld, M., Bogo, N., and Lambert, W.E. Evaluational reactions to accented English speech. Journal of Abnormal and Social Psychology, 1962, 65, 223-231.

Buck, J. The effects of Negro and White dialectical variations upon attitudes of college students. Speech Monographs, 1968, 35, 181-186.

Duncan, D.B. Multiple range and multiple F tests. Biometrics, 1955, 11, 1-42.

Harms, L.S. Listener judgments of status cues in speech. Quarterly Journal of Speech, 1961, 47, 164-186.

Harms, L.S. Status cues in speech: Extra-race and extra-region identification. Lingua, 1963, 12, 300-306.

Labov, W. The Social Stratification of English in New York City. Washington, D.C.: Center for Applied Linguistics, (1966).

Lambert, W.E. Psychological approaches to the study of language, Part II: On second-language learning and bilingualism. The Modern Language Journal, 1963, 47, 114-121.

Lambert, W.E. A social psychology of bilingualism. The Journal of Social Issues, 1967, 23, 91-109.

Lambert, W.E., Anisfeld, M., and Yeni-Komshian, G. Evaluational reactions of Jewish and Arab adolescents to dialect and language variations. Journal of Personality and Social Psychology, 1965, 2, 84-90.

Lambert, W.E., Frankel, H., and Tucker, G.R. Judging personality through speech: A French-Canadian example. Journal of Communication, 1966, 16, 305-321.

Lambert, W.E., Hodgson, R.C., Gardner, R.C., and Fillenbaum, S. Evaluational reactions to spoken languages. Journal of Abnormal and Social Psychology, 1960, 60, 44-51.

Naremore, R.C. Teachers' evaluational reactions to pupils' speech samples. Unpublished doctoral dissertation, University of Wisconsin, 1969.

Naremore, R.C. Teachers' judgments of children's speech: A factor analytic study. Journal of Speech and Hearing Research, in press.

Preston, M.S. Evaluational reactions to English, Canadian, French, and European French voices. Unpublished master's thesis, McGill University, 1963.

Putnam, G.N. and O'Hern, E. The status significance of an isolated urban dialect. Language Dissertation, No. 53. Language, 1955, 31, No. 4, Part 2.

Rosenthal, R. and Jacobson, L. Pygmalion in the classroom. New York: Holt, Rinehart, and Winston, 1968.

Seligman, C.R., Tucker, G.R., and Lambert, W.E. Teachers' attitudes and pupils' performance. Unpublished research report, McGill University, 1970.

Shuy, R.W. The sociolinguists and urban language problems. In F. Williams (Ed.), Language and poverty. Chicago: Markham, 1970.

Shuy, R.W., Wolfram, W.A., and Riley, W.K. Linguistic correlates of social stratification in Detroit speech. U.S. Office of Education Cooperative Research Project No. 6-1347, Michigan State University, 1967.

Shuy, R.W., Baratz, J.C., and Wolfram, W.A. Sociolinguistic factors in speech identification. National Institutes of Mental Health Research Project No. MH-15048-01, Center for Applied Linguistics, 1969.

Stewart, W.A. Urban Negro speech: Sociolinguistic factors affecting English teaching. In R.W. Shuy (Ed.), Social dialects and language learning. Champaign, Ill.: National Council of Teachers of English, 1965.

Tucker, G.R. and Lambert, W.E. White and Negro listeners' reactions to various American-English dialects. Social Forces, 1969, 47, 463-468.

Veldman, D.J. Fortran programming for the behavioral sciences. New York: Holt, Rinehart, and Winston, 1967.

Williams, F. Psychological correlates of speech characteristics: On sounding "disadvantaged." Journal of Speech and Hearing Research, 1970, 13, 472-488.

Williams, F. and Naremore, R.C. On the functional analysis of social class differences in modes of speech. Speech Monographs, 1969, 36, 77-102. (a)

Williams, F. and Naremore, R.C. Social class differences in children's syntactic performance: A quantitative analysis of field study data. Journal of Speech and Hearing Research, 1969, 12, 778-793. (b)