DOCUMENT RESUME

ED 247 285 TM 840 464

AUTHOR Marshall, Hermine H.; Weinstein, Rhona S.
TITLE Classrooms Where Students Perceive High and Low Amounts of Differential Teacher Treatment. Draft.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.; National Inst. of Mental Health (DHHS), Rockville, MD.; Spencer Foundation, Chicago, Ill.
PUB DATE Apr 84
GRANT 1-ROH-MH34379-01; G-80-007
NOTE 47p.; Paper presented at the Annual Meeting of the American Educational Research Association (68th, New Orleans, LA, April 23-27, 1984).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Classroom Observation Techniques; Elementary Education; Expectation; *Feedback; *High Achievement; *Low Achievement; Questionnaires; *Student Attitudes; Teacher Attitudes; *Teacher Behavior

ABSTRACT

The purpose of this paper is to explore differences in the nature of classrooms where students perceive high and low amounts of differential teacher treatment between high and low achievers. The Teacher Treatment Inventory was used to measure students' perceptions of the frequency of 30 teacher behaviors towards a hypothetical male or female high or low achieving students in twelve classrooms, four each at grades 1, 3, and 5. In addition, teacher expectations were measured and the classrooms were observed. Hypotheses concerning the differential occurrence of variables affecting the formation of students' self-evaluation were supported to a greater extent at grade 5 than at grades 1 and 3. The operation of quantitative structural and interaction variables is enlightened by qualitative field notes which allow for an understanding of the dynamic interaction of variables within the larger context. (Author/BW)

DRAFT

Classrooms Where Students Perceive High and Low Amounts of Differential Teacher Treatment

Hermine H. Marshall and Rhona S. Weinstein

University of California, Berkeley

Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, April, 1984.

Work on this paper was supported by the National Institute of Education Grant G-80-007, by the National Institute of Mental Health Grant 1-ROH-MH34379-01, and by the Spencer Foundation. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education or the National Institute of Mental Health or the Spencer Foundation, and no official endorsement should be inferred.

Acknowledgement and thanks are due to Mary Lou Bedrosian Vernon, Marilyn Little, and Yu Wen Ying for help in collecting the data; to Leonard Marascuilo for advice on data analysis; and to Deborah Curtis for help in analyzing the data.

Classrooms Where Students Perceive High and Low Amounts of Differential Teacher Treatment

Research to ascertain differences between classrooms at grades 1, 3, and 5 where students perceive high and low amounts of differential teacher treatment between high and low achievers is reported. Hypotheses concerning the differential occurrence of variables affecting the formation of students' self-evaluation are supported to a greater extent at grade 5 than at grades 1 and 3. The operation of quantitative structural and interaction variables is enlightened by qualitative field notes which allow for an understanding of the dynamic interaction of variables within the larger context.

The research reported here is a part of a series of studies designed to investigate student mediation of teacher expectancy effects in the classroom. The purpose of this paper is to explore differences in the nature of classrooms where students perceive high and low amounts of differential treatment. Specifically, preliminary results are reported from an observational study of 12 classrooms at three grade levels selected from the extremes of a larger sample of classrooms where students perceive high and low amounts of differential treatment of high and low achievers. The exploration of the nature of classroom structure and interactional processes within which student perceptions of high and low amounts of differential treatment are embedded is based on a model of classroom factors which affect students' self-evaluations and expectations (Marshall & Weinstein, in press).

Communication of teacher expectations. Following publication of Rosenthal and Jacobson's study (1968) of self-fulfilling prophecies in classrooms, much research in this area has focused on the communication of teacher expectations in the classroom. (See Brophy, 1983, for a review.) Much of the early work was conducted using quantitative analyses of discrete teacher behaviors or sequences of teacher-student-teacher behaviors (Brophy & Good, 1970; 1974). A set of teaching behaviors emerged from this research which was often found to reflect differential treatment of high and low achievers, such as less frequent calling on students, demanding less work, more frequent criticism for an incorrect public response (Brophy, 1983; Good, 1980).

Other research has focused on the structural organization of the classroom (Bossert, 1979; Marshall, 1976) which has implications for the communication of teacher expectations. Aspects of the way the teacher structures the environment may serve as a means of providing clues to students about their performance relative to their classmates and thus as a means of affecting their self-evaluations. Bossert (1979) has identified four classroom organizational factors which contribute to the ease with which students can make comparisons about their relative ability: (a) grouping practices, (b) similarity or differentiatedness of tasks, (c) evaluation practices, and (d) student autonomy and choice.

Using a similar perspective, Rosenholtz and Wilson (1980) have shown that in classrooms with ability or whole class grouping, similar tasks, public evaluation, and little autonomy ("high resolution classrooms"), high consensus on academic ranking of students is found between teacher and students, and among students. Blumenfeld and her associates (1982) add that teacher practices which increase competition and discriminate between high and low achievers may also influence students' perceptions of ability by heightening the comparability of performance and its publicness and salience. Ames (1981) has shown that in competitive environments, success enhances evaluations of ability and failure decreases these evaluations to a greater extent than in cooperative environments.

Research has also addressed students' awareness of communications about expectations (Weinstein, in press). A student mediation model of the processes by which teacher expectations are communicated through patterns of differential teacher treatment of high and low achievers has

been postulated (Weinstein, in press). According to this student mediation model, patterns of differential teacher treatment are believed to contain cues about expected achievement which students can perceive, interpret, and act on, resulting in different levels of achievement for students about whom teachers hold high and low expectations. A series of studies has documented that (a) students do perceive differential teacher treatment [illegible] of behaviors derived in part from studies (e.g. Brophy & Good) of how teacher expectations are expressed in behavior, (b) classrooms differ in the amount of differential teacher treatment that students perceive (Weinstein, Marshall, Brattesani, & Middlestadt, 1982), and (c) teacher expectations are more closely associated with student expectations and with student achievement in classrooms in which students perceive high amounts of differential treatment than in classrooms where low amounts of differential treatment are perceived (Brattesani, Weinstein, and Marshall, 1984). Yet, at this point, we know little about the classroom context surrounding students' perceptions of teachers' differential treatment of high and low achievers on the selected behaviors measured in this research.

Limitations of behavioral categories. Investigations of the student perspective, especially an inspection of students' responses to interview questions about life in classrooms (Weinstein, 1980) suggest several limitations with much of the work so far. The first of these centers on the drawbacks of using behavioral categories that overlook the larger context that may change the meaning of the behavior to the participants. As Brophy (1979) has pointed out, most research on classroom process variables uses categories which include behaviors that, although similar on the surface, may have different meanings to students depending on the nonverbal accompanying behaviors, the situation and the context. For example, interviews with students reveal at least four student interpretations of the teacher behavior "call on" (Weinstein & Middlestadt, 1979). The teacher "calls on the smart kids for the right answer. ... She expects you to know more and won't tell the answers;" whereas the teacher sometimes calls on low achievers "to give them a chance" or "because they goof off." Or often she "doesn't call on them because she knows they don't know the answer." These variations in interpretation, which may not be apparent from classroom observational data, may affect students' interactions and student outcome variables.

Neglect of multiple interacting factors. The second limitation suggested by student perception data is the failure to recognize that other factors within the classroom environment may overcome a potentially detrimental effect of one factor or change the potentially beneficial effect of other factors. Studies using student perceptions of differential teacher treatment (Weinstein et al., 1982; Brattesani et al., 1984) as well as dyadic observational studies (Brophy & Good, 1974) have noted that only some teachers discriminate in their treatment of high and low achievers. Nevertheless, all of the teacher-subjects who participated in the research hold expectations for their students--since they can rank or rate their students on year-end achievement. Yet, the expectations of teachers who were perceived as showing little treatment differentiation on the behaviors investigated were not as related to student performance as the expectations of teachers who were perceived as showing high treatment differentiation (Brattesani et al., 1984).

Indeed, observations in some of the classrooms where students reported little differential treatment in behaviors traditionally found to distinguish those teachers whose expectations affect performance from those whose expectations do not, revealed that all students were not being treated the same. Students in one classroom were taking math tests at different levels of difficulty. Students in another of these low perceived differential treatment classrooms were assigned to different groups for special projects of different levels of complexity and some students were asked to help others.

These differences between classrooms raise several questions. First, do these "low perceived differential treatment teachers" differentiate in their treatment of high and low achievers in ways other than those previously found to discriminate between high and low achievers? For example, do some of these teachers hold types of expectations that can be translated into the provision of appropriate learning experiences which are likely to provide each student with optimal opportunities for learning? Second, do these teachers differentiate on some of the same behaviors, but in ways that do not communicate differential expectations. For example, do they vary the recipients of these behaviors so that stable differentiating patterns do not occur? Or third, do these teachers treat high and low achievers differently on some of the hypothesized behaviors but the existence of compensating features within the larger picture viewed by the students alters the interpretation that students make of these behaviors? That is, what is it about the context of these "perceived" low differential treatment classrooms that prevents the teachers' expectations from having a deleterious effect on the performance of students for whom they have low expectations?

In attempting to answer some of these questions, we have proposed a model of factors and their potential interactions which, within the climate created, are postulated to contribute to the development of students' self-evaluations (Marshall & Weinstein, in press). This model suggests how previously investigated factors interact with each other as well as with additional factors so as to compensate for or negate one another. Included in the model are elements of classroom organizational structure, teacher-student interactions, and the quality of relationships established by the teacher. The model focuses on structural strategies which make public the comparability of relative performance, as well as on other factors which serve to alter the potentially deleterious effect of social comparison.

According to this model, aspects of the classroom environment which may provide opportunities for students to observe differential teacher treatment reflecting differential expectations and to make comparative assessments affecting their own expectations and performance include (a) the task structure, (b) grouping practices, (c) feedback and evaluation procedures and information about ability, (d) motivational strategies, (e) locus of responsibility for learning (teacher vs. student), and (f) the quality of teacher-student relationships.

Several examples will serve to demonstrate how the various factors may influence students' ability to observe differences in treatment and

make comparative assessments. First, where the classroom is organized for whole class instruction or with stable, homogeneous (ability) groups, comparisons are more easily made than where flexible or heterogeneously composed small groups that are called together for particular short-term purposes are used. Second, where teachers encourage students to evaluate their own work, evaluation may be more private, made on varied criteria, and students may be less vulnerable to external evaluation pressures. And third, the assignment of different tasks may decrease opportunities for peer comparison. However, where the tasks require divergent processes or products, students may be less able to compare their work, even where the assigned tasks are the same. On the other hand, where tasks are different but from the same series in a sequence, students may be able to compare their own work with others who are in different places in the assigned series of tasks. The last example illustrates how one factor (e.g. task divergence) may overcome the effect of another factor (e.g. task similarity). (The model is further elaborated in Marshall & Weinstein, in press.)

In order investigate the nature of classrooms where students perceive high and low amount of differential teacher treatment as suggested by this model, we have also developed an observational system designed to overcome the limitations of behavioral categories that overlook the larger context in the classroom (Marshall & Weinstein, 1982). This system includes qualitative field notes as well as quantitative behavioral categories. This approach allows retrieval of explanatory information to enlighten the findings of the quantitative data, in addition to providing information about contextual factors.

Method

Subjects

Twelve teachers, 4 each at grades one, three and five, in nine urban ethnically mixed schools, in two school districts were recruited on a voluntary basis for the study. Only self-contained, single grade classrooms were used. These classrooms were selected on the basis of data collected in the fall of the school year from a larger sample of 30 classrooms as representing the extremes of high and low differential teacher treatment as perceived by students. The classrooms were selected according to the following criteria. In each classroom (described under Student Measures), for each of the three scales of the Teacher Treatment Inventory, a difference score was obtained between mean student responses on high and low achiever long forms (combined across male and female versions). The values of the three scale difference scores were added together (irrespective of sign) to yield a classroom level perceived differential teacher treatment score. Classrooms were then ranked within grade level on the amount of perceived differential teacher treatment and the two highest and lowest scoring classrooms within each grade were chosen for observation. Of the eleven female and one male teachers, three were Black, three were Asian, and six were Caucasian. Four classrooms were in one district and eight--including all four fifth grades--were in the other.

Student Measures

Students' perceptions of teacher treatment. The Teacher Treatment Inventory (TTI) was used to measure students' perceptions of the frequency of 30 teacher behaviors towards a hypothetical male or female high or low achieving student. Items on this instrument were derived from reviews of the literature on the relationships between teaching behaviors and student achievement, on the expression of teacher expectations in behavior, and student perceptions of classroom environments as well as from pilot interviews with students (Weinstein & Middlestadt, 1979). The instrument was further refined based on an assessment of the meaning and reliability of the items and reliability, stability, and validity of the scales (Weinstein & Marshall, 1984). Instrument refinement resulted in a reduction of the original four scales to three 10-item scales: (1) Negative Feedback and Teacher Direction; (2) Work and Rule Orientation; and (3) High Expectations, Opportunity and Choice, suitable for administration to first through fifth graders.

Internal consistency coefficients (Cronbach's alphas) for the three scales over both forms for students in grades one, three, and five were .70, .63, and .81 respectively. Two-week test-retest reliability coefficients, as indicated by Pearson's correlation coefficients, over both high and low forms and over all three grades were .73, .70, and .80 for the three separate scales.

The items on the TTI were administered according to whether the hypothetical student described was male or female and a high or low achiever. The descriptions of the target students follows:

> High achiever form. This boy/girl is someone who does really well in school. In fact, he/she always gets the best grades in the class. Everyone thinks he/she is very smart.
>
> Low achiever form. This boy/girl is someone who does not do very well in school. In fact, he/she always gets the lowest grades in the class. Everyone thinks he/she is not very smart.

Students were asked to pretend that this was a student in their own class and to rate how frequently their own teacher would work with him/her in the ways described. Students responded to each item by marking one of four different circles of decreasing size, labeled "Always," "Often," "Sometimes," and "Never." A sample item and practice trial were also provided.

Achievement measures. Grade placement scores on the Comprehensive Test of Basic Skills (CTBS) Reading and Math Achievement Tests were collected from the prior year-end and current year-end district-wide testing.

Teacher Measures

Teacher expectations for students. Teachers were presented with three decks of student name cards, one at a time, and asked to rank them in order of expected year-end performance in reading, in math and in

overall schoolwork.

Teacher Interview. An interview schedule consisting of both structured and open-ended questions was created to clarify the classroom observations. Interview questions focused on (a) grouping practices, (b) uniformity of curriculum sequence, (c) evaluation practices, (d) locus of responsibility (student or teacher), (e) conceptualization of students' abilities, and (f) effective teaching strategies for high and low achievers.

Observer Measures

Observers' perceptions of teacher treatment. A shortened form of the TTI was constructed consisting of eight items, four positive treatment items and four negative and structuring items. Items which significantly differentiated the treatment of high and low achievers at an item level in earlier studies were selected for this form.

Classroom Observation Methods

A two-part system for observing in classrooms, the Classroom Dimensions Observation System, was developed and refined based on previous work. (See Marshall & Weinstein, 1982.) This observation system includes both qualitative field notes and a quantitative observation scale (CDScale). This system centers on aspects of the classroom which are believed to have implications for the communication of achievement expectations: (a) Structure of the tasks, subject matter and materials, (b) Grouping practices, (c) Locus of responsibility in learning, (d) Feedback and evaluation, (e) Motivation, (f) Quality of teacher-student relationship and (g) Expectations.

Focused field notes. In using the Classroom Dimensions Observation System, the classroom observer first keeps a narrative record of events in the classroom, focusing on those aspects of the classroom and teacher-student interactions which are believed to have implications for the development of achievement expectations. Teacher statements other than subject matter content are recorded as closely to verbatim as possible. Individual students with whom interaction occurs are identified. Also recorded is whether the interaction occurs with the class as a whole, with a group or group with others around, with an individual alone, an individual within group (group setting), or an individual with others around. The observer also makes separate notes of impressions and interpretations of events. Field notes are typed immediately according to a format for ease of retrieval of teacher statements.

Observational scale. At the beginning of the observation period, the observer uses the Classroom Dimensions Scale (CDScale) to code an overview of the general structure of the learning environment. Following the observation period, the observer records the exact number of instances certain teacher behaviors were observed and rates aspects of the climate on the CDScale based on field notes. The CDScale is a low inference observational scale designed to provide both quantitative and qualitative information concerning the cognitive, affective, interpersonal, and structural aspects of the classroom. This scale is derived

from the Dimensional Occurrence Scale (Marshall, 1976).

The CDScale allows analysis of the recorded data according to subject matter, type of classroom structure (whole class, group or individual, and level of the group (high, middle, low). A new CDScale form was used each time the subject matter changed as well as each time that the teacher changed the group s/he worked with and each time there was a change in how the group functioned.

The scale is divided into three parts. Part I yields an overview of the general structure of the tasks, grouping, and evaluation which create the context for learning during the observation period. This section provides a general picture of (a) whether the students are working individually, in groups, or together as a class; (b) where the teacher is working; (c) the subject matter content and types of tasks; (d) the amount of choice that the students have; and (e) the predominant type of teacher evaluation. The items on Part I represent categories (nominal and ordinal), frequency counts, and amount of time.

Part II focuses more specifically on the nature of the teacher's interactions with the students or with groups of students. Items in this section provide additional information about (a) the type of task, (b) motivation, (c) responsibilities, (d) evaluation and feedback and (e) the quality of relationships. The items in Part II represent countable instances of behavior and are coded for the exact number of times (frequency) that the behavior occurs. This part also allows for the coding of whether the teacher's interaction is with individuals, groups of students, or the class as a whole.

In Part III, the frequency and intensity of the warmth and irritation conveyed to the class, groups or individuals are rated.

Observers undertook extensive training over a period of eleven weeks, including more than 30 2-hour training sessions, beginning with videotapes and moving into actual classrooms. As a check on inter-observer agreement, the transcripts of the field notes were inspected for correspondence of events between observers. Inter-observer agreement for the CDScale for each of the three observers with the trainer (who served as the fourth observer) was calculated using the percent exact agreement averaged over six observation periods, three observation periods for two teachers. Percent agreement ranged from .94 to .97 for the items on Part I, from .94 to .96 for the items on Part II, and from .88 to .92 for the overall level of variables on Part III. Calculation of agreement for the items on Part II is based on the exact number of times that the behaviors were observed to occur as well as the observation that the behaviors did not occur. Because many of these behaviors are infrequently occurring events and did not occur during the observation periods when the observers observed the same events, agreement concerning the actual occurrence of some of these infrequent behaviors could not be directly ascertained. To ensure agreement on these behaviors when they were encountered during the data collection, weekly meetings were held to discuss all occurrences of infrequent behaviors and to resolve other coding problems. In addition, the trainer read all transcripts and re-checked the coding of infrequent events for

consistency within and across observers.

Procedures

Fall teacher expectations. In the early fall, teachers were individually presented with three decks of cards, one at a time and were asked to rank their students from one to thirty on year-end expected achievement in reading, math, and schoolwork.

Student perceptions of differential treatment. The TTI was administered following collection of the teacher data. Based on prior year-end reading achievement scores, a randomized blocking procedure was used to assign forms to participating students within sex within each classroom. Students for whom achievement data were missing were blocked together and similarly assigned forms. Each student was assigned to a TTI long form of either a male or female high or low achiever version and to a TTI short form of the same sex but the opposite level achievement.

Trained testers administered a long form of the TTI (male or female, high or low achiever) followed by a non-related filler task and then an opposite achievement level same sex short form of the TTI. All instructions and items were read aloud to avoid problems of varying reading levels.

Classroom observations. Observers were assigned randomly to classrooms with the following restrictions: (a) Each observer was assigned one classroom at each of the three grade levels. (b) No more than two of the three classrooms assigned to an observer had the same level of perceived differential treatment (high or low). (c) During the fall data collection, one observer formed an hypothesis about the differential treatment level of one of the observed classes. Another observer formed hypotheses about four of the observed classes. These classes were assigned to observers with no prior hypotheses. All observers were blind to the actual amount of differential teacher treatment perceived by students in all of the classrooms.

The order in which classes at the three grade levels were observed was varied across observers.

Each observer observed in one classroom at a time for a period of two to four weeks. Preliminary observations were made to acclimate the observer to the classroom and the students to the observer as well as to learn the students' names. After these initial observations, an additional 12 hours of observations per classroom or more were made in an attempt to observe three periods during which high and low reading groups received instruction, three math lessons, and some whole class discussion or organizational time. The context of the observations during the remainder of the time varied according to the type of activity common to the particular classroom.

Observers used the Classroom Dimension Observation System to make a narrative record of the teacher's interactions with individuals and groups of students as well as with the class as a whole, and recorded

teacher comments in the areas hypothesized to be important to the development of achievement expectations. Observations were immediately typed and coded on the CDScale.

Observers' impressions of differential teacher treatment. After observing in a classroom, the observer completed a short form of the TTI for a high and low achieving male and female for that classroom. Observers also wrote a summary of their impressions based on what was observed and made a judgment as to whether the classroom was a high or low differential treatment classroom. Since the observers were familiar with an early version of the model of classroom factors which were postulated as contributing to students' self-evaluations on which the Classroom Dimension Observations System was based, their estimates may have been influenced by this model as well as by observed evidence of differential teacher treatment.

Teacher interviews. After the observations in each classroom, the observer interviewed the teacher. Teachers were asked to rank their students again on expected year-end achievement in reading and math and were then asked the questions on the interview schedule. Interviews were audiotaped and transcribed, and observers recorded the responses to the forced choice questions.

Supplementary Data

Supplementary data on the nature of two of the four first grade classrooms in this study comes from a dissertation study by Mary Lou Bedrosian Vernon (1983). Bedrosian Vernon conducted a study of reading groups in four first grade classrooms following the main observations. These classrooms were also chosen from the extremes of high and low perceived differential teacher treatment. Teachers in her study had to be willing to have their reading groups videotaped. Since not all of the first grade teachers in the current study were willing to participate in the videotaping, two classrooms from the current study and two additional classrooms were selected.

The videotaped and transcribed lessons for high and low reading groups were coded by two coders blind to the hypotheses of the study and to the identification of the classrooms. The transcripts were coded as to teacher behaviors of informality, trust, warmth, and support as well as student-initiated remarks, question-response-evaluation, chatting, and evaluation events.

Chi-square tests of binomial proportions were performed to test for differences between high and low reading groups within each of the four classrooms. The analyses of these variables were considered together in attempting to distinguish teachers that appeared to be high and low differentiating in their treatment of the high and low reading groups. Two of the four teachers were found to be high differentiating and two low differentiating. However, in only two cases did the level of coded differentiation correspond with the student-perceived level of teacher treatment differentiation (according to the TTI). The level of differentiation in the coded behaviors of the two teachers who participated as well in the current study was not consistent with students'

perceptions of differential treatment.

Results

The analyses reported here reflect preliminary work on the observational data. They focus on the quantitative variables drawn from Part I and Part II of the CDScale supplemented by a preliminary review of the qualitative field notes. Analyses of variables from Part III and additional work on the narrative records of classroom events are still underway. Further, the results reported here describe whole-class characteristics. We are in the midst of conducting between-reading group analyses of structuring and interactional strategies.

Quantitative Analyses

The CDScale yielded four kinds of data: nominal scores, ratings, and frequencies of teacher behavior as well as amount of time spent. Since we observed whole lessons within the classroom, the amount of time spent in each classroom and time spent observing reading groups and other subject matter lessons varied between teachers. Similarly, the number of CDS forms completed for each teacher and for different subject matters varied as well, since CDS forms were changed each time the subject matter was changed or within subject matter each time the grouping structure or group with whom the teacher worked changed.

In order to compensate for the varying lengths of time for each CDS form, for each subject matter, and for total amount of observation, raw scores were adjusted by the number of CDS forms used or by the number of minutes of observation, where appropriate. (Table 1 shows the mean number of CDS forms used and the mean number of minutes classrooms were observed.)

Creation of variables. Variables were created to describe three levels of classroom structure and process: (a) the class as a whole, which included class level ratings of structure as well as the summed frequency of teacher behaviors across individual, individual-in-group, group, individual-in-class, and whole-class contexts; (b) characteristics of groups, which included data from all times during which groups were in operation and (c) high and low reading groups, which included observations separated by subject matter (in this case, reading) and for all times in which the teacher worked with the high and low reading group. (In one classroom only, math groups were substituted since the teacher did not use group instruction for reading.)

Part I of the CDScale yielded structural information about the classroom. Four variables were created to describe the predominant type of classroom organization used: the proportion of individual structure over all observations, the proportion of group structure, the proportion of whole class structure and the proportion of mixed structure (some combination of individual, group, and whole class structure). Four variables were created to indicate aspects of task structure: the presence of student choice (occur vs. not occur), the use of divergent tasks (occur vs. not occur), the concurrent use of different subject matter

(occur vs. not occur) and the sameness of the tasks (on a 5 point score from same exact, same series, different tasks in a series, same broad topic, different activities).

Four variables described the nature of group instruction. These included the number of groups worked with during observation, the proportional use of short-term flexible grouping over the total number of group observations, the proportional use of heterogeneous grouping over the total number of group observations and the type of label given to the groups (no label or neutral label, consecutive label and imagery label). A fifth variable was created based on information obtained from the teacher prior to the observations: the number of groups identified for instruction in reading (as well as math and spelling).

Part II of the CDScale provided frequency data on 42 types of teacher interactions with students. These interactions concerned task strategies, motivational strategies, establishing responsibility, evaluative feedback, and interpersonal relationships. Items were combined on a conceptual basis and aggregated to create a set of proportional variables. These included the (1) proportion of encourage expressiveness (proportion of observations in which teachers encouraged expressiveness), (2) proportion of cooperative strategies (proportion of observations in which the teacher used cooperative strategies), (3) proportion of positive display (frequency of positive display for academic and behavioral purposes divided by the total display behaviors), (4) proportion of positive academic evaluation (all positive academic evaluative items divided by positive plus negative academic evaluative items), (5) proportion of positive behavioral evaluation (all positive behavioral evaluative items divided by positive plus negative behavioral evaluative items), (6) proportion of praise (frequency of praise divided by the frequency of praise plus criticism), (7) proportion of buffered criticism to total criticism, and (8) proportion of positive interpersonal behavior to total interactions.

Statistical procedures. Due to the unequal variances and non-normal distributions, nonparametric methods were used to analyze the data from the CDScale. Different methods of statistical analysis were used for the proportional variables and for the rating variables.

To test for equality of proportions, a series of a priori contrasts were performed. (These contrasts are commonly associated with the Chi Square Test of Homogeneity.) In order to retain the equal contribution of each teacher in the analyses, the proportions utilized by each teacher were given equal weight. Thus, for example, the proportion of "Praise" in Grade 1 actually represents the average proportion of "Praise" across the four teachers in that grade. The standard error of each contrast was computed under H_1(Goodman, 1963).

To test for equality of ratings, a series of a priori contrasts were performed based upon the Kruskal Wallis test. For these analyses, the model was "laid out" as a one-way design. To correct for tied values when ranking, average ranks were assigned. This correction for ties was also utilized in calculating the variance of each contrast. As in the proportional analysis, equal weighting was employed.

For both of the above analyses, the contrasts under consideration were (a) the difference between type of classroom (level of differential treatment), (b) differences among grades, and (c) type of classroom by grade interactions. Since each of these three groups of contrasts represents "families" of hypotheses, a family-wise error rate of .05 was used. For example, each of the three contrasts for grade were assigned an alpha of .0167, or .05/3 (Marascuilo & Levin, 1970).

Structural aspects of the classroom. Table 2 documents the proportion of observations during which each type of organization structure was in use. Only 4 of the 12 teachers in this sample used an individualized structure for teaching during our observation periods; hence, differences in usage of individualized structure could not be tested. To test the hypothesis that whole class structure would occur more frequently in high than low perceived differential treatment classrooms and to explore grade level differences in type of structure, a priori contrasts were conducted on the proportional use of group, whole class, and mixed structure. These contrasts revealed significant main effects for grade level on two of the three variables, with significant Grade x Type of Classroom interactions of two of the three variables as well. No main effect for type of classroom was documented. A greater proportion of whole class structure was observed in first grade compared to fifth grade ($z = 2.43$) with no other effects noted. Further, proportionally less mixed structure was observed at the first grade level compared to the third grade level ($z = 2.66$, $p < .01$) and the fifth grade (although not significantly). However, there was more use of mixed structure in perceived high compared to low differential treatment classrooms at the first grade; whereas at third and at fifth grades, the higher proportion of mixed structure was found in low differential treatment classrooms ($z = 2.48$). For the proportional use of group structure, these relationships were reversed. More group structure was documented in low differential treatment classrooms compared to high differential treatment classrooms at first grade; whereas the higher proportion of group structure was found in high differential treatment classrooms at third ($z = 3.10$) and at fifth grades ($z = 2.66$).

With regard to aspects of the task structure, Table 3 demonstrates that the proportion of student choice, divergent tasks and concurrent use of different tasks and subjects was not high. In fact, the observed instances of concurrent use of different subject matters was too low to allow statistical analysis. Of the remaining three task variables, significant Grade level effects were documented for two of these three variables and significant Type of Classroom effects for all three variables. The proportion of student choice was higher in third ($z = 4.91$) and fifth grades ($z = 3.03$) than in first grade, as was the use of concurrent different tasks (the latter tested by the Kruskal Wallis test). The proportional use of divergent tasks did not show overall grade level differences. Use of student choice, divergent tasks and concurrent different tasks were all found to be higher in perceived high compared to low differential treatment classrooms ($z = 2.84$; $z = 6.27$; $z = 3.62$ for the three variables respectively), contrary to our hypotheses. A significant Grade x Type of Classroom interaction for divergent tasks suggests that the type of classroom difference was greatest at the first grade level ($z = 2.79$, 2.93 for Grade 1 compared to Grade 3 and 5).

Characteristics of grouping. During meetings with teachers preceding the observations, teachers were asked for a list of their groups in reading, math and spelling, and in whatever other subjects they used grouping. The number of groups that the teachers identified in conference with the observer did not always correspond with what the observers actually recorded as they watched the teachers in the classroom. In some cases, the teachers identified groups to the observers, for example, by what book or level they were in, but in reality combined several of these groups for instruction. Another teacher identified three groups but never instructed groups in reading during our observations. Instead, this teacher circulated among the students in individualized instruction. Table 3 shows the mean number of identified and observed groups by type of classroom. Whereas we would have predicted a larger number of reading groups in low compared to high differential treatment classrooms, the means suggest only a slight difference in instructed groups in favor of low differential treatment classrooms. Perceived high differential treatment teachers identified more reading groups than did low differential treatment teachers. However, the difference between the number of groups identified and instructed is greater for perceived high than for low differentiating teachers.

The use of flexible short-term grouping and heterogeneous group composition (that is, not ability based) was not observed in all classrooms; thus differences in degree of usage could not be analyzed statistically. Table 4 shows the proportion of teachers in our sample who used these structural strategies by type of classroom. Inspection of these patterns suggest variability in use. For flexible groups, perceived high differential teachers appeared more likely to use this strategy in first and third but not fifth grade, contrary to our hypothesis. For heterogeneous grouping, perceived low differential teachers appeared more likely to utilize this strategy in first and fifth grade but not in third grade.

Concerning the degree of imagery in the labeling of groups, contrasts based on the Kruskal Wallis test were carried out to determine whether more neutral than imagery labels would occur in low as compared to high differential treatment classrooms. These contrasts suggest main effects for Grade level and for Type of Classroom. Labels conveying no or neutral messages were more likely to be found in first grade than in third or fifth grade ($z = 2.66$; $z = 2.83$). Only 2 of the 12 teachers used imagery labels for their groups. As well, contrary to our hypothesis, neutral labels were more likely to be used in perceived high differential treatment classrooms than in low differential treatment classrooms ($z = 2.17$). No significant interactions were documented.

Interactional strategies (Whole class findings). A priori contrasts were conducted on the eight teacher interaction variables to examine whether treatment is more positive in perceived low than high differential treatment classrooms and to explore grade level differences. Proportions are shown in Table 5 and significant effects are displayed in Table 6. These contrasts revealed a significant overall difference between perceived high and low differential treatment classrooms for seven of the eight variables, although not in the predicted direction. Teachers in perceived high differential treatment classrooms

were observed in general to be more encouraging of student expressiveness, use more cooperative strategies, more positive display, more positive academic evaluation, more positive behavioral evaluation, more praise, and more positive relationship behaviors. No overall classroom differences were documented for the amount of buffered criticism. However, these overall classroom type differences were qualified by significant Classroom x Grade interaction on six of the seven variables. A significant Classroom x Grade interaction was also noted for buffered criticism.

These analyses also demonstrated significant grade level effects in the frequencies of seven observed teacher interaction variables. Teacher's use of positive display, positive behavioral evaluation (but not academic evaluation) and praise was higher in first grade than in third grade (for display) and third and fifth grade (for evaluation and praise); whereas the observed frequency of buffered criticism and positive relationships was higher in the later grades. Third grade teachers encouraged student expressiveness more than did fifth grade teachers and fifth grade teachers used more cooperative strategies than did third grade teachers. Six of these grade level effects, too, are qualified by significant Classroom x Grade interactions.

The significant Classroom x Grade interactions suggest that the perceived high and low differential treatment classroom differences noted in the positivity of teacher interactions overall favored high differential treatment classrooms only for first grade, and sometimes for third grade, but that at fifth grade, perceived low differential treatment classrooms were observed to have more positive teacher interaction behavior than did high differential treatment classes. At fifth grade, teachers in perceived low differential treatment classrooms were observed to be more encouraging of student expressiveness, use more positive display, more positive academic as well as behavioral evaluation, more buffered criticism, and more positive relationship behaviors than did teachers in high differential treatment classrooms. Teacher use of cooperative strategies showed no such interactions and thus was higher in perceived high differential treatment classrooms at all grade levels. In addition, differences in levels of teacher praise were not documented between types of classrooms at fifth grade.

Summary. These results from the quantitative analyses conducted thus far at the whole class level do not provide strong support for our hypotheses about the manner in which the structural and interactional variables operate individually. Our hypotheses about the differential use of these structural and interaction variables generally seem to hold up better at Grade 5 than at Grade 1. Additional analyses concerning the teacher's treatment of high and low groups as well as of individual students who are high and low expectation students, and concerning differences in the public or private context of the interactions have yet to be completed (although an inspection of the means between high and low reading groups suggests that teachers do use different strategies with high and low groups).

Informal Profile Analysis

An inspection of the variables for the individual teachers (within each grade level and type of classroom) suggests that what is happening is much more complex than the results of a traditional statistical analysis would indicate. In an attempt to account for the results of the quantitative analyses conducted thus far and to provide a more complete picture of classrooms where students perceive high and low amounts of differential teacher treatment, an examination of the way the structural and interactional variables operate within individual classrooms was made in what might be considered an informal type of profile analysis. Such an analysis of the way variables operate within individual classrooms is consistent with the model of classroom factors which postulates that variables must be considered within the larger context of the classroom as a whole as well as in terms of how these variables are influenced by and have an influence on other variables. Certain variables may compensate for or negate the effects of other variables as noted earlier.

To carry out a preliminary profile analysis, several types of information were utilized. First, the CDS structural and interactional variable scores analyzed thus far were listed for each teacher. In addition, each teacher was ranked on each of these variables. (See Tables 7 and 8.) This listing of variable scores and ranking allows for an exploration of whether the scores and ranks for each teacher are consistent with the categorization of teachers as high or low differential treatment based on student perceptions.* For example, we can examine whether a student-perceived high differential treatment teacher was ranked high on each of the variables. A more complete picture can be portrayed when additional qualitative analyses have been conducted.

A second set of information was used in an attempt to classify the classrooms according to whether students' and observers' agreed in their perceptions of the amount of differential teacher treatment (as measured by the TTI) and according to whether the nature of the classroom context postulated in the model of classroom factors affecting student self-evaluations was consistent with that expected based on the level of student perceptions. Classification as to whether or not there was consistency between perceptions based on specific teacher behaviors and impressions based on the larger context expected to surround these perceptions provides a framework for looking more closely at the patterns of CDS variables within individual classrooms.

* Note should be made that this study was not designed as a validation study of the Teacher Treatment Inventory. Few of the behaviors on the observational scale were the same as those on the inventory. The inventory items were rated; most of the observation scale items were frequency counts. In addition, the two measures were taken in different time periods. Rather the purpose of the study was to examine the context of classroom structural and interactional strategies within which students perceived high or low amounts of differential teacher treatment.

The student and observer perceptions based on the TTI--although measured at different time periods--can provide an indication of agreement or disagreement about differentiation on specific teacher behaviors. As described earlier, the initial level of differentiation was determined by a median split within each grade level of the difference between student TTI scores for the high achiever 30-item-long form and the low achiever 30-item long form using the 30 classrooms in the larger fall sample. The level of differentiation for the observer TTI was based on a median split across grade levels in the 12 classrooms observed in the winter, using differences on the 12-item short form of the TTI.

In addition to the observer TTI, three sources of information were used to provide a picture of whether the classroom structural and interactional strategies were consistent with that expected from the level of student-perceived differential teacher treatment. The first source of information is the observers' impression of whether the classroom was a high or low differential treatment classroom. These observer impressions were based on information gleaned while observing in the classroom and from formal and informal interviews with the teacher, considered in conjunction with the observer's general knowledge of the model of classroom factors affecting the development of self-evaluations. The second source of information consists of the first author's impressions after reading the qualitative field notes and the teacher interview transcriptions for each teacher. Third, the level of differentiation in coded behaviors between high and low reading groups in the Bedrosian Vernon study was used as a supplementary source for two of the first grade classes. A median split based on the number of significant differences in proportions of favorable treatment between high and low reading groups was used to determine high and low levels of differentiation here.

Consistency between student-perceived level of differential treatment and information from all other available sources was found in five of the 12 classrooms. The influence of grade level on this correspondence of perceptions and impressions was evident in that consistency was found in three of the four Grade 5 classrooms, and only one of the classes at each of Grades 1 and 3. (See Table 9.)

The third set of information used in the profile analysis is comprised of the narrative records of each classroom. The transcripts were read for additional clues as to the nature of the classroom, themes that emerged in individual classrooms, and teachers' statements of expectations and of attributions for success and failure.

Finally, the difference in the residualized gain scores for reading achievement between students for whom the teacher held high expectations and those for whom the teacher held low expectations was explored for preliminary clues concerning the effects of the student mediation model of teacher expectations on student outcomes. (See Table 10.) These residualized gain scores could not be calculated at Grade 1 due to the nature of the prior year achievement scores for the first graders. Less difference between high and low expectation students in the amount of residualized achievement gain was expected in low than in high

differential treatment classrooms.

Classrooms with Overall Consistency Perceptions and Impressions

An examination of the CDS variable scores and rankings analyzed to this point for each of the five classrooms where student perceptions agreed with outside sources indicated that the direction of these CDS variables showed greatest consistency with both student and observer perceptions and impressions in two Grade 5 classrooms, one high and one low differential treatment classroom. That is, these two classrooms which were selected on the basis of student perceptions of differential teacher treatment in the fall fairly clearly fit the model of classroom factors which facilitate the development of students' self-evaluations as measured by the CDScale in the winter.

In Classroom K, the Grade 5 low differential treatment classroom, consistency with the low differential treatment level was found on all the interaction variables (Table 8). Discrepancies were noted on some of the CDS structural variables (Table 7). Relatively few instances of heterogeneous grouping, divergent tasks, and student choice and no instances of flexible grouping were recorded on the CDScale. A relatively high proportion of whole class structure was observed. However, the qualitative field notes reveal that this teacher used heterogeneously grouped "families" for seating and study periods, where students of differing abilities could help each other. Use of these "families" may have compensated for the lack of other types of heterogeneous or flexible groups. Further, the difference in residualized gain scores between high and low teacher expectation students was relatively low, as would be predicted.

In Classroom I, the agreed upon high differential treatment Grade 5 classroom, the only discrepancies between the CDS variables and level of student and observer perceptions were that student choice, task divergence, and concurrent different subject occurred more frequently than would be predicted for a perceived high differential treatment classroom. In this classroom, the difference in residualized gain scores between high and low teacher expectation students was relatively high, as would be predicted.

A brief summary from the qualitative field notes may highlight some of the factors which contribute to these two classrooms exemplifying a low and a high differential treatment setting. In Classroom K, major themes that are apparent are the learning orientation, student responsibility, and respect for individual differences. In this well-managed classroom, Teacher K makes learning a challenge. Competition in learning is against her, e.g. "I'm gonna catch you. This is tricky." Emphasis is on thinking rather than on just getting work done. Making mistakes is a part of the learning process rather than a sign of low ability. "I'm not interested in how many you got wrong; I'm interested in if I can help you." Expectation statements convey the sense that students can do it. Attributions for success or failure are often to external sources, such as the book going too fast, and sometimes to effort, rather than to ability. Responsibilities are assigned and students seem to know what to do and when. Teacher K responded to a

student request for a Valentine's Day celebration by indicating they could take responsibility for deciding about having "a little something." "You discuss it. It's up to you. You decide." Student responsibility for evaluation is exemplified by a comment to a low achiever about her writing: "I know you have beautiful writing when you want... Let me ask you: Are you proud of it?" Respect for each student and for individual differences is recurrent. "Finish [what you are saying even though the bell rang] because you're important." "Hands down. Give her a chance [to think]." In announcing those who received good scores on spelling, she said, "For some, spelling is the hardest subject. Your "families' are going to help you." Another discussion centered on different opinions about smartness and that it is not how much brains you have but how you use them.

Turning to the high differential treatment fifth grade classroom, Classroom I is managed reasonably effectively. However, the orientation is towards "work" rather than learning. Students need to finish their work so they can "go out to recess" or "get out of that book" rather than think. The lack of trust and responsibility is illustrated by Teacher I's refusal to allow a Valentine's party since she had to spend 45 minutes cleaning up from the Christmas party. She was also observed to respond for students and to cut off opportunities for them to respond, such as reading their responses for them, completing a problem, drawing a face. Errors and ability comparisons are often public. Teacher I asked how many people made mistakes; when no one responded, she announced, "I've got one here who got the whole row wrong." Seating was arranged according to high or low group (and consequently largely by race in this class). The high group is labeled "top group." Students are threatened about being removed from the "top group" if they do not "stop fussing." High expectations are expressed for high expectations students, e.g. to read "semi-adult books," "I don't expect you people in the top group to get low grades in spelling."--though some reservations were expressed about one student. Positive expectations were expressed for the middle level students. However, Teacher I commented aloud to the observer, "Now do you see why this is the lowest group?" Attributions were more frequently made to internal sources, e.g. effort, ability, immaturity, than to external sources such as task difficulty.

The other three classrooms where observer perceptions and impressions were consistent with the student perceived level of differential treatment exhibit elements of the model as evidenced on the CDS variables to varying degrees. In two of these classrooms, some of the CDS variables were consistent with the type of classroom and some were discrepant. One of these classrooms was a perceived low differential treatment classroom at Grade 5. The other was a high differential treatment classroom at Grade 3.

For Teacher G, a Grade 5 low differential treatment teacher, the interaction variables on the CDScale were less positive than would have been predicted by the model. The structural variables of amount of task divergence, concurrent different subject, student choice, and heterogeneous grouping were also lower than would have been expected. Only in the use of flexible grouping was there consistency between the relative frequency of CDS variables and type of classroom. Qualitative field

notes suggest an apathy and aloofness in Teacher G's interactions with high as well as low achievers in the classroom. This detachedness and poor management strategies may have influenced both student and observer perceptions and impressions of low differential treatment. The difference between high and low teacher expectation students in residualized gain scores was moderate, not lending much support to predictions based on the low differential treatment label.

Contrary to the classroom factors model, Classroom C, the Grade 3 high differential treatment classroom, showed high proportions of positive relationships and a relatively high use of the proportion of buffered criticism. Teacher C also used heterogeneous and flexible groups relatively frequently. The narrative records suggest that certain of Teacher C's strategies may have counteracted the potentially beneficial effect of the positive relationships and the grouping practices. For example, Teacher C was observed to give a clear public statement of negative expectations to a low achiever: "I bet you haven't [started on your project]. Bet a nickel." When the student replied that she had gotten a book, Teacher C continued: "She says she just got a book. [She] hasn't written anything yet." A moderate difference was found between high and low teacher expectation students in their residualized gain scores, although this difference was the highest of the four third grade classes. It may be that, despite the operation of certain variables in consonance with a low differential treatment classroom, the operation of other strategies and the teacher's verbal message about differences in expected performance may dilute the potentially beneficial effect of the former strategies.

Although students and observers agreed that Teacher F (Grade 1) was a high differential treatment teacher, the general pattern of CDS variables reflects the model of a classroom where students would be less able to perceive differences in treatment. However, closer inspection of individual items on the CDS scale--some of which were incorporated into the aggregated CDS variables--indicates that Teacher F used some strategies that may have counteracted the expected effect of the more positive variables and accounted for the perception as a high differential treatment teacher. For example, Teacher F encouraged expressiveness relatively often, but sometimes discouraged expressiveness. Both cooperation and competition were observed. Negative comparison was used more frequently than positive comparison. Although rewards were used relatively frequently, punishment also occurred occasionally. (Residual gain scores could not be calculated at Grade 1.)

Thus, even in these classrooms where students' and observers' perceptions agreed and where other impressions were consistent with the model of classroom contextual factors in which these perceptions would be expected to be embedded, the variation in the way the CDS variables operated within classrooms may account, in part, for some of the lack of clearcut findings in the statistical analyses conducted to date.

Classrooms Lacking Consistency between Student Perceived Level of Differentiation and Outside Impressions about Classroom Context

Of the six classrooms where some discrepancy existed between

students' perceptions of the amount of differential teacher treatment and impressions of the nature of the classroom expected to surround these perceptions, the observers' perceptions of the specific behaviors measured by the TTI agreed with those of the students in three of these classrooms. In making judgments about these classrooms as a whole, the observers seemed to be basing their impressions on a broader set of indicators than the specific teacher behaviors on the TTI. These indicators seemed to include elements from the model of classroom factors influencing the development of self-evaluations which are also reflected in variables on the CDScale. Two of these classrooms where student and observer TTIs agreed were at Grade 1 (one high and one low differential treatment) and one was at Grade 3 (low differential treatment). The observers' impressions of the two classrooms at Grade 1 (Classrooms L and J) are generally consonant with the classroom factors model as indicated by the level of CDS variables. This is not the case for the Grade 3 classroom.

In Classroom D, the Grade 3 low differential treatment classroom, the structural variables on the CDScale were discrepant from the low TTI level but in accord with the observer impression. In contrast, most of the CDS interaction variables were at a moderate level. As revealed in the qualitative field notes, the observer's impression was influenced by this teacher's common use of the strategy of having students read out their math scores, since this would make comparative assessments clearly visible, as well as the observation of a child crying after reporting her score. It may be that the students knew who did well in math but did not see those students as being treated differently. Similarly, the observer did not see the students being treated differently on the behaviors sampled on the TTI but was influenced by the structural aspects of the classroom factors model and by the observation of the display of math scores and student reaction to this display. It should be noted, however, that the difference between high and low teacher expectation students on the residualized gain scores in this classroom was very small, and in fact, favored the lows--more consistent with expectations for a low differential treatment classroom.

In three classrooms--one at each grade level--the observer TTI was discrepant from that of the student. In these classrooms, patterns seemed less clear-cut, leading to varying impressions depending, perhaps on which elements were the focus. In the Grade 1 classroom, Classroom H, the observer made a judgment that this was a low differential treatment classroom in accord with the students' perceptions. The observer noted that the atmosphere and teacher behaviors in this classroom were highly negative but that the teacher seemed to be highly critical of all students, not just low achievers. Nevertheless, in contrast to the student TTI and the observer estimate, the observer TTI indicated that this was a high differential treatment classroom. The results of the Bedrosian Vernon study and the impressions of the first author reading the narrative records were also that this was a high differential teacher treatment classroom. The level of the CDS variables indicates as well that this was a high differential treatment classroom. The original observer estimate seemed to be influenced by the heavy (critical) atmosphere in the classroom.

Teacher A, the third grade teacher, was perceived by students as showing low amounts of differential treatment; but according to the observer TTI, she displayed high amounts of differential treatment. However, with some ambivalence, the observer made a judgment that this teacher was a low differential teacher, stating that the teacher showed rather negative interactions towards both high and low achievers. The first author's impressions are that Teacher A is more likely a high differential treatment teacher. Despite multiple reading groups, all students are given the same spelling words which may therefore be too difficult for some students. Yet the scores of all students are announced and displayed. The interaction variables on the CDScale were relatively negative. Only in task divergence and student choice were the CDS structural variables consistent with the placement of this teacher. Examination of the residualized gain scores indicates a moderate difference between high and low teacher expectations students, although this difference is relatively high compared to the third grade classes. This difference does not strongly support the designation of this teacher as a low differential treatment teacher.

The last teacher where there was disagreement between student and observer perceptions (TTI) was a Grade 5 high differential treatment teacher, Teacher E. The observer TTI as well as the level of most of the CDS interaction variables and the structural variables of flexible grouping and concurrent different subject are consistent with the observer impression of this classroom as a low differential treatment classroom. The major unique feature of this classroom is its highly individualized nature and relatively frequent use of flexible grouping. Based on the model of classroom factors influencing the development of self-expectations, individualized classrooms would be expected to minimize the opportunities for comparative assessment. Yet somehow, these fifth grade students appeared to perceive differential treatment between high and low achievers. Further, the differences between the high and low teacher expectation students in residualized gain scores appear to support the designation of this classroom as a high rather than a low differential treatment classroom.

Classroom Where Observer Found Elements of Consistency and Discrepancy in the Classroom Context Expected from the Student-Perceived Level of Differentiation

Finally, for Teacher B, a Grade 3 student-perceived high differential treatment teacher, the observer (and first author) found elements of both high and low types of classrooms. The observer TTI and the variables on the CDS generally indicated a classroom on the low end of the scale. However, the observer notes--as did the teacher in an interview--that the teacher does treat high achievers differently from low achievers, "more as adults." The high group is given more independent work; and when the teacher meets with them, he discusses both reading and math assignments, rather than having separate sessions for each subject area. Further, as evidenced by the observer impressions and many of the CDS variables, many factors of the model which would minimize the opportunity to make comparative assessments were present, e.g. concurrent different subjects, divergent tasks, student choice.

In addition, the difference in residualized achievement gain between high and low teacher expectation students was more consistent with a low than a high differential treatment classroom. In this case, it seems as if the larger contextual factors reflected in the model and the CDS variables may have influenced the interpretations that students make of the differences in treatment such that their learning behaviors were maximized.

Interestingly, student perceptions of differential treatment on the High Expectation, Opportunities and Choice scale were at the low end of the continuum, ranking ninth out of the ten third grade classrooms in the original sample. This may imply that although some differences in treatment were perceived, crucial differences in the teacher's expectations were not perceptible.

Discussion

This study has attempted to begin answering questions raised from prior research regarding the nature of classrooms where students perceive high and low amounts of differential teacher treatment and whether teachers who are perceived as showing low amounts of differential treatment differentiate between high and low achievers in ways other than those found in previous research or in ways that minimize the communication of negative expectations.

The study and the observational system were designed so that both quantitative and qualitative data would be available to answer these questions. The results of the quantitative analyses of the Classroom Dimension Scale (CDScale)--the categorical section of the Classroom Dimension System--support our hypotheses concerning the nature of high and low differential teacher treatment classrooms to a greater extent at Grade 5 than at Grade 3 or Grade 1. At Grade 5, a lower proportion of whole class structure and a greater proportion of mixed structure (indicating flexibility) was observed in perceived low than high differential teacher treatment classrooms. In addition, teachers in these low differential teacher treatment classrooms showed a greater proportion of positive display (to positive plus negative display), positive academic as well as positive behavioral evaluation (to positive plus negative evaluation), buffered criticism (to total criticism), and positive interpersonal relationships (to total interactions).

A number of factors may have contributed to the failure to support our hypotheses tested in these preliminary analyses at other grade levels as well as those concerning other individual variables. The first set of factors concerns the time line for this study. The Teacher Treatment Inventory (TTI), upon which selection of classes was based, was administered in the fall. In our previous work, it had been administered in the spring. The age and maturity level of the first graders in the fall may have diminished the reliability of the measure at this grade level.

Furthermore, the observations were made in the winter. It is possible that some teachers change their behaviors over the course of the

school year. Different strategies may be used in the fall than in the winter or spring. Some teachers may emphasize rules and work orientation to a greater extent in the fall as they attempt to socialize their students into the student role as it exists in their classroom. Once this socialization process has taken place, these teachers may have less need for these strategies and may bring more student choice and responsibility into play. On the other hand, some teachers attempt to establish an "open" type of classroom early in the year yet lack the skills to implement this type of classroom (cf Marshall, 1981). When these teachers become frustrated in their attempts, they may revert to more teacher-directed strategies. Therefore, some teachers may utilize behaviors which can be perceived as differentiating between high and low achievers in the fall but not the winter, and some teachers may show more differentiation in the winter than the fall. Since low achievers have been described by students as showing less academic task conformance than high achievers (Marshall, Weinstein, Sharp & Brattesani, 1982), they may therefore be perceived as the recipients of more negative and teacher-directed behaviors.

In his review of research on the self-fulfilling prophecy, Brophy (1983) also describes the potential effect of time of year. At the beginning of the school year, teachers may make an extra effort with low achievers in an attempt to keep the class together; but as end of the year pressures to meet curriculum requirements build, teachers may shift their concentration to high achievers.

A second set of factors which may have influenced the results of the quantitative analyses refers to differences in the scores on the TTI from our earlier studies. The absolute difference between the sum of the scores over the three scales for the high achiever form and the low achiever form appears less for this study than in previous studies. These differences will be explored further. In addition, a great deal of variability was observed between the scores on each of the scales within each classroom. That is, a teacher may have been categorized as a high differential treatment teacher due to high difference scores on two of the three scales but that teacher may have had a very low difference score on the third scale. These factors together with the difference in the time of year of administration may have contributed to the selection of a somewhat less than ideal set of classrooms from the extremes of a high-low perceived differential teacher treatment continuum, particularly in the earlier grades.

A third set of factors concerns the variables on the observation system. Varying numbers of items were aggregated to create the proportions for the interaction variables. The creation of these variables was conceptually based. However, it may be that the individual items which comprise these variables may not all work in the same direction as had been anticipated. Further work on individual items is in progress to investigate this possibility.

In addition, some of the items were infrequently occurring behaviors, but behaviors that are considered important to the classroom factors model. In some of the classrooms, some of these infrequently occurring behaviors were not observed to occur at all--as would be

expected in certain types of classrooms (high or low differentiating, depending on the behavior). Despite their importance, low frequency items are difficult to analyze.

An alternative method of explaining the findings of the variables on the CDScale is a consideration of how these individual variables work together within individual classrooms in an informal profile analysis. This approach is consistent with the model of the operation of classroom variables which postulates that the operation of one variable may compensate for or negate the effect of another variable (Marshall & Weinstein, in press). The preliminary profile analysis reported herein was based on an inspection of the scores on the individual variables within each classroom as well as on clues from the qualitative field notes.

Consistent with the findings of the quantitative analysis, the clearest examples of high and low differential teacher treatment classrooms about which both students' perceptions and observers' perceptions agreed and where the impressions based on the model of classroom factors affecting self-evaluations matched these perceptions occurred at Grade 5. Yet even in these two relatively clear-cut examples, not all of the individual variables were found to occur in the predicted direction. For example, flexible and homogeneous grouping was infrequently recorded on the CDScale in the low differential treatment classroom. However, the qualitative field notes documented the existence of heterogeneously grouped "families" for seating and study period which may have compensated for the lack of flexible and heterogeneous groups during subject matter periods. Other classrooms where students' and observers' perceptions and impressions were consistent included fewer variables that worked in the predicted direction. These classrooms may have been farther from the prototype.

In those classrooms where outside judgments of the nature of the context in which the student perceptions were embedded did not correspond with that anticipated based on the level of student perceptions, several patterns were found. In some of these classrooms, the observer's TTI agreed with the students' TTI but differed from the observer's actual judgment of the level of the classroom. In these classrooms, it may be that students and observers discerned high and low achievers being treated differently on many of the specific teacher behaviors included on the TTI, but the observers (and perhaps the students) may be influenced in their interpretation of these behaviors by the larger context of the classroom as a whole and by other particular factors which may overcome the potentially negative impact of the differential teacher treatment or nullify the beneficial effect of low differential treatment.

In two other classrooms where outside impressions of the type of classroom did not correspond to what would be expected from the level of student perceptions of differential teacher treatment, it may be that the observer was unable to understand and capture adequately the influences on student interpretations and actions. In one of these two cases, the observer was influenced by a striking example of negative public comparison and its effect on children. In the other case, the observer was influenced by the individualized nature of the classroom

which would correspond to a type of classroom where comparative assessments would be less apparent. The observer did note that due to the private nature of the individualized interactions, she was unable to ascertain the content of the interactions in many instances. Nevertheless, other factors may have been operating which affected the students so that their residualized gain scores corresponded more with what would be predicted based on their own perceptions of differential treatment level than with the observer's judgment.

Finally, in one student-perceived high differential treatment classroom, the observer perceived elements of both high and low types of classrooms. Students, observer, and teacher recognized differences in treatment between high and low achievers. Yet, because the differences between high and low teacher expectation students in residualized gain scores corresponded more with a low level of treatment differentiation, it may be that in this classroom, the larger contextual factors which included student input and responsibility, support for student effort and errors, statements of positive expectations, and so on may have compensated for the observable treatment differences and allowed students to be unaffected by these differences. This speculation is supported by the fact that the difference perceived between high and low achievers on the High Expectations, Opportunities and Choice scale was low, rather than as high as the differences on the other two scales.

Summary and Conclusions

In sum, this study has shown that variations exist in classrooms where students perceive low amounts of differential teacher treatment as well as in those where students perceive high amounts of differential teacher treatment. Greater support for our hypotheses was found in the quantitative analyses at Grade 5. A trend towards greater correspondence between the level of student-perceived differential treatment and information concerning the nature of the classroom contextual factors in which these perceptions are embedded was also documented at higher grades. Further, relatively consistent examples of a high and a low differential treatment classroom which corresponds to the model of classroom factors that contribute to students' self-evaluations was reported at the fifth grade. Other findings suggest that older children may be able to consider more contextual factors in interpreting and reporting differential teacher treatment. In addition, some support has been found for the existence of factors which may compensate for the effect of other factors and for the uniqueness of each classroom, regardless of its classification.

Finally, because of the importance of considering the classroom context, this study demonstrates the importance of collecting qualitative data to supplement and enlight the interpretation of quantitative findings.

References

Ames, C. Competitive versus cooperative reward structures: The influence of individual and group performance factors on achievement attributions and affect. American Educational Research Journal, 1981, 18, 273-287.

Bedrosian Vernon, M. L. The social-emotional dimension of teacher-student ineractions during beginning reading instruction. Doctoral dissertation, University of California, Berkeley, 1983.

Blumenfeld, P. C., Pintrich, P. R., Meece, J., & Wessels, K. The formation and role of self-perceptions of ability in elementary classrooms. Elementary School Journal, 1982, 82, 401-420.

Bossert, S. T. Tasks and social relationships in classrooms. Cambridge: Cambridge University Press, 1979.

Brattesani, K. A., Weinstein, R. S., & Marshall, H. H. Student perceptions of differential teacher treatment as moderators of teacher expectation effects. Journal of Educational Psychology, 1984, 76, 236-247.

Brophy, J. E. Teacher behavior and its effects. Journal of Educational Psychology, 1979, 71, 733-750.

Brophy, J. E. Research on the self-fulfilling prophecy and teacher expectations. Journal of Educational Psychology, 1983, 75, 631-661.

Brophy, J. E., & Good, T. L. Teacher's communication of differential expectations for children's classroom performance: Some behavioral data. Journal of Educational Psychology, 1970, 61, 365-374.

Brophy, J. E., & Good, T. L. Teacher-student relationships. New York: Holt, Rinehart & Winston, 1974.

Good, T. Classroom expectations: Teacher-pupil interactions. In J. C. McMillan (Ed), Social psychology of school learning. New York: Academic Press, 1980.

Goodman, L. A. On methods for comparing contingency tables. Royal Statistical Society Series B, 1963, 126, 94-108.

Marascuilo, L. A., & Levin, J. R. Appropriate post hoc comparisons for interaction and nested hypotheses in analysis-of-variance designs: The elimination of Type IV errors. American Educational Research Journal, 1970, 7, 397-421.

Marshall, H. H. Dimensional occurrence scale: Manual. School of Education, Berkeley, CA: University of California, 1976.

Marshall, H. H., & Weinstein, R. S. Classroom dimensions observation system. Psychology Department, Berkeley, CA: University of California, 1982.

Marshall, H. H., & Weinstein, R. S. Classroom factors affecting the development of students' achievement expectations. Review of Educational Research, in press.

Marshall, H. H., Weinstein, R.S., Sharp, L., & Brattesani, K. A. Students' descriptions of the ecology of the school environment for high and low achievers. Paper presented at the meeting of the American Educational Research Association, New York, March, 1982.

Rosenholtz, S. R., & Wilson, B. The effect of classroom structure on shared perceptions of ability. American Educational Research Journal, 1980, 17, 75-82.

Rosenthal, R., & Jacobson, L. Pygmalion in the classroom: Teacher expectations and pupils' intellectual development. New York: Holt, Rinehart and Winston, 1968.

Weinstein, R. S. Student perceptions of differential teacher treatment: Final report. Berkeley, CA: Department of Psychology, University of California, 1980.

Weinstein, R. S. Student mediation of classroom expectancy effects. In J. B. Dusek (Ed.), Teacher expectancies. Hillsdale, N.J.: Lawrence Erlbaum Associates, in press.

Weinstein, R. S., & Marshall, H. H. Ecology of students' achievement expectations: Final report. Psychology Department, University of California, Berkley, 1984.

Weinstein, R. S., Marshall, H. H., Brattesani, K. A., & Middlestadt, S. E. Student perceptions of differential teacher treatment in open and traditional classrooms. Journal of Educational Psychology, 1982, 74, 678-692.

Weinstein, R. S., & Middlestadt, S. E. Student perceptions of teacher interactions with male high and low achievers. Journal of Educational Psychology, 1979, 71, 421-431.

Table 1

Mean Number of CDS Forms Used and Minutes of Observation

		CDS'S	Minutes Obs.
Grade 1	High Diff	38.50	745.00
	Low Diff	28.50	709.50
Grade 3	High Diff	27.50	618.00
	Low Diff	32.00	723.00
Grade 5	High Diff	26.00	746.50
	Low Diff	31.50	728.50
Sample Mean		30.67	711.75

Table 2

Proportions and Standard Errors for Structural Variables by Grade Level and Type of Classroom

Variables		Grade 1		Grade 3		Grade 5		High Diff Classes	Low Diff Classes	Grade		
		High D	Low D	High D	Low D	High D	Low D			1	3	5
Type of Organization												
Group Structure	P	.21	.36	.34	.15	.40	.24	.32	.25	.29	.24	.32
	Se	.05	.07	.06	.10	.07	.05	.03	.03	.04	.04	.04
Whole Class Structure	P	.58	.49	.39	.48	.37	.39	.45	.45	.54	.44	.38
	Se	.06	.07	.07	.06	.07	.06	.04	.04	.05	.04	.05
Mixed Structure	P	.19	.13	.25	.32	.17	.35	.20	.27	.16	.29	.26
	Se	.04	.04	.06	.05	.05	.06	.03	.03	.03	.04	.04
Task Structure												
Student Choice	P	.11	.05	.36	.27	.31	.12	.26	.15	.08	.31	.22
	Se	.03	.03	.07	.05	.07	.04	.03	.03	.02	.04	.04
Task Divergence	P	.56	.07	.42	.22	.44	.27	.47	.19	.31	.32	.36
	Se	.05	.03	.07	.05	.07	.05	.04	.03	.03	.04	.05
Task Difference:												
Same task		.54	.80	.38	.41	.33	.38	.42	.53	.42	.40	.36
Same series		.17	.12	.21	.21	.10	.44	.16	.26	.15	.21	.32
Different task in series		.025	.015	.24	.30	.36	.13	.21	.15	.02	.27	.24
Same broad topic		.125	.04	.11	.08	.06	.03	.10	.05	.08	.09	.05
Different task		.155	.04	.07	.02	.16	.02	.13	.02	.10	.04	.09

Table 3

Group Structural Variables by Grade Level and Type of Classroom

Variables	Grade 1		Grade 3		Grade 5		High Diff Classes	Low Diff Classes	Grade		
	High D	Low D	High D	Low D	High D	Low D			1	3	5
Mean # identified reading groups	5.0	3.0	6.5	5.5	3.0	3.0	4.8	3.8	4.0	6.0	3.0
Mean # instructed reading groups	2.5	2.0	6.0	4.5	1.5	3.0	2.7	3.2	2.3	4.3	2.3
Mean # math groups	0	1.5	1.5	0	3.0	2.5	1.5	1.3	0.8	0.8	2.8
Group labels											
No or neutral labels[a]	.50	.87	.69	.04	.50	.44	.56	.45	.69	.37	.47
Consecutive labels[a]	.51	.13	.32	.93	.25	.56	.36	.54	.32	.63	.41
Image labels[a]	0	0	0	.03	.25	0	.08	.01	0	.02	.13

[a]Proportion of observed groups where these labels were used.

Table 4

Proportion of Teachers within Grade and Type of Classroom Using Flexible and Heterogeneous Grouping

		High Differential	Low Differential
Flexible			
	1	1.00	0
	3	1.00	.50
	5	.50	.50
Heterogenous			
	1	.50	1.00
	3	1.00	.50
	5	.50	1.00

Table 5

Proportions and Standard Errors for Teacher Interaction Behavior by Grade Level and Type of Classroom

Variables		Grade 1 High Diff	Grade 1 Low Diff	Grade 3 High Diff	Grade 3 Low Diff	Grade 5 High Diff	Grade 5 Low Diff	High Diff Classes	Low Diff Classes	Grade 1	Grade 3	Grade 5
Encourage Expressiveness	P	.33	.08	.36	.25	.15	.19	.28	.17	.20	.31	.17
	Se	.05	.03	.07	.05	.05	.05	.03	.03	.03	.04	.03
Cooperative Strategies	P	.29	.20	.24	.08	.39	.16	.31	.15	.24	.16	.28
	Se	.05	.06	.06	.03	.07	.05	.03	.03	.04	.03	.04
Positive Display	P	.86	.57	.67	.44	.58	.74	.70	.58	.71	.55	.66
	Se	.03	.06	.08	.06	.07	.09	.04	.04	.03	.05	.06
Positive Academic Evaluation	P	.92	.68	.81	.66	.68	.83	.81	.72	.80	.75	.75
	Se	.02	.03	.03	.03	.03	.04	.02	.02	.02	.02	.03
Positive Behavioral Evaluation	P	.46	.16	.07	.12	.06	.13	.20	.14	.31	.09	.09
	Se	.04	.02	.03	.02	.02	.03	.02	.02	.02	.02	.02
Buffered criticism	P	.38	.18	.30	.15	.47	.69	.38	.34	.28	.22	.58
	Se	.05	.03	.05	.03	.03	.04	.03	.02	.03	.03	.03
Praise	P	.92	.43	.67	.49	.56	.52	.71	.48	.68	.58	.54
	Se	.02	.03	.04	.03	.02	.03	.02	.02	.02	.02	.02
Positive Relationships	P	.09	.02	.09	.06	.08	.13	.09	.07	.06	.08	.10
	Se	.01	.00	.06	.01	.01	.01	.01	.00	.00	.01	.01

Table 6

Significant Effects on Teacher Interaction Variables

Interaction Variables	Differential Treatment Classroom Effects		Grade Level Effects		Interaction Effects	
	Z	P level Direction	Z	P level Direction	Z	P level Direction
Encourage student expressiveness	2.61*	H>L	2.53*	3>5	3.10*	H>L (1>5)
Cooperative strategies	3.78*	H>L	2.25*	5>3		
Positive display	2.14*	H>L	2.76*	1>3	3.48* 2.60*	H>L (1>5) H>L (3>5)
Positive academic evaluation	3.40*	H>L			6.08* 4.71*	H>L (1>5) H>L (3>5)
Positive behavioral evaluation	2.64*	H>L	7.77* 7.67*	1>3 1>5	6.19* 6.59*	H>L (1>3) H>L (1>5)
Buffered criticism			9.21* 7.66*	5>3 5>1	5.41* 4.94*	H>L (1>5) H>L (3>5)
Praise	9.80*	H>L	3.49* 5.09*	1>3 1>5	5.57* 8.09*	H>L (1>3) H>L (1>5)
Positive relationships	2.99*	H>L	2.90*	5>1	3.56* 2.26*	H>L (1>5) H>L (3>5)

*Significant

Table 7

Proportions and Ranks for Structural Variables (CDS Part I) for Each Teacher

		Grade 1				Grade 3				Grade 5			
		High Diff		Low Diff		High Diff		Low Diff		High Diff		Low Diff	
		F	L	H	J	B	C	A	D	E	I	G	K
Grouping Structure/Total Operations													
Indiv	prop	.023	.015	.027	.013	.015	.019	.093	.014	.097	.024	.013	.016
	rank	5	8.5	3	11.5	8.5	6	2	10	1	4	11.5	7
Group	prop	.159	.269	.421	.304	.245	.436	.015	.278	.258	.545	.250	.222
	rank	11	6	3	4	9	2	12	5	7	1	8	10
Class	prop	.591	.567	.526	.455	.389	.400	.369	.585	.323	.409	.313	.476
	rank	1	3	4	6	9	8	10	2	11	7	12	5
Mixed	prop	.227	.149	.026	.299	.351	.145	.523	.123	.322	.022	.406	.286
	rank	7	8	11.5	6	3	9	1	10	4	11.5	2	5
Flexibility and Heterogeneity of Grouping/Total Observations													
Flex	prop	.583	.286	0	0	.400	.188	.059	0	.706	0	.143	0
	rank	2	4	10	10	3	5	7	10	1	10	6	10
Hetero	prop	.333	0	.125	.158	.133	.188	.059	0	.118	0	.048	.067
	rank	1	11	5	3	4	2	8	11	6	11	9	7
Task Divergence, Student Choice, Different Subject/Total Observations													
Task Divergence	prop	.386	.727	.055	.077	.571	.273	.406	.031	.355	.523	.250	.296
	rank	5	1	11	10	2	8	4	12	6	3	9	7
Student Choice	prop	.205	.015	.027	.077	.464	.261	.375	.156	.194	.429	.125	.111
	rank	5	12	11	10	1	4	3	7	6	2	8	9
Different Subject	prop	.227	.030	0	.026	.536	.074	.625	0	.419	.429	.063	.126
	rank	5	9	11.5	10	2	7	1	11.5	4	3	8	6

Table 8

Proportions and Ranks for Interaction Variables (CDS Part II)

for Each Teacher

	Grade 1				Grade 3				Grade 5			
	High Diff		Low Diff		High Diff		Low Diff		High Diff		Low Diff	
	F	L	H	J	B	C	A	D	E	I	G	K
Positive Relationships/Total Interactions												
proportion	.074	.109	.012	.030	.098	.090	.063	.049	.119	.047	.036	.214
rank	6	4	12	10	5	3	9	7	2	8	11	1
Positive Academic Evaluation/Positive and Negative Academic Evaluation												
proportion	.851	.978	.448	.920	.894	.788	.489	.838	.887	.464	.727	.923
rank	6	1	12	3	4	8	10	7	5	11	9	2
Positive Behavioral Evaluation/Positive and Negative Behavioral Evaluation												
proportion	.352	.576	.074	.250	.119	.015	.034	.200	.068	.046	.004	.259
rank	2	1	7	4	6	11	10	5	8	9	12	3
Buffered Criticism/Criticism												
proportion	.433	.320	.034	.322	.250	.353	.036	.259	.784	.146	.643	.739
rank	4	7	12	6	9	5	11	8	1	10	3	2
Positive Display/Positive and Negative Display												
proportion	.773	.938	.370	.764	.933	.400	.280	.600	.889	.263	.556	.929
rank	5	1	10	6	2	9	11	7	4	12	8	3
Praise/Praise and Criticism												
proportion	.960	.875	.243	.623	.729	.600	.279	.704	.890	.228	.223	.797
rank	1	3	10	7	5	8	9	6	2	11	12	4

Table 9

Categorization of Teachers into High and Low Differential Treatment

According to Various Criteria

	Grade 1				Grade 3				Grade 5			
	High Diff		Low Diff		High Diff		Low Diff		High Diff		Low Diff	
	F	L	H	J	B	C	A	D	E	I	G	K
Student TTI	High	High	Low	Low	High	High	Low	Low	High	High	Low	Low
Observer TTI	High	High	High	Low	Low	High	High	Low	Low	High	Low	Low
Observer Impressions	High	Low	Low	High	High & Low	High?	Low?	High	Low	High	Low	Low
Author Impressions	High	Low	High	High	High & Low	High?	High	High	Low	High	Low	Low
Bedrosian Vernon	--	Low	High	--								

Table 10

Difference in Mean Residualized Gain Scores

between High and Low Teacher Expectation Students

Grade 1				Grade 3				Grade 5			
High Diff		Low Diff		High Diff		Low Diff		High Diff		Low Diff	
F	L	H	J	B	C	A	D	E	I	G	K
------not available------				.38	.51	.44	-.18	1.03	.75	.50	.32