



The Relationship between Child's SES and Teacher Expectations: A Test of the Middle-Class Bias Hypothesis

Dale G. Harvey; Gerald T. Slatin

Social Forces, Vol. 54, No. 1. (Sep., 1975), pp. 140-159.

Stable URL:

<http://links.jstor.org/sici?sici=0037-7732%28197509%2954%3A1%3C140%3ATRBCSA%3E2.0.CO%3B2-B>

Social Forces is currently published by University of North Carolina Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/uncpress.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

The Relationship Between Child's SES and Teacher Expectations: A Test of the Middle-Class Bias Hypothesis

DALE G. HARVEY, *Comprehensive Care Center,
Lexington, Kentucky*

GERALD T. SLATIN, *University of Kentucky*

ABSTRACT

The purpose of this study was to assess the degree to which teachers' expectations are related to children's social class characteristics. A sample of 96 elementary school teachers drawn from four schools serving lower- and middle-upper class neighborhoods was asked to judge the performance potential and related characteristics, including the SES background, of a set of photographs of both black and white grade school children. While expectations were found to be positively and significantly related to perceived SES of the children in the photos, variations by school SES and number of years teaching experience in the mean number of judgments made suggest that teachers' class bias operates in a more complex manner than is usually thought. In addition, teachers were found to respond more often when choosing children for "success" than for "failure" categories. Regardless of perceived SES, white children were more often expected to succeed in school than black children.

Recently, Rosenthal (1973) and Rosenthal and Jacobson (1968) presented data supporting the hypothesis that a teacher's expectations of a child greatly influence the child's actual behavior. (Students of whom teachers were led to expect exceptional performance were found generally to live up to the expectations.) They speculated that the artificially created positive expectations the teachers had for certain children in their classes caused them to be more warmly encouraging; to give more feedback; to teach more effectively; and to ask more of these chosen few. Thus the teachers were transmitting high expectations to children who behaved accordingly. Though not directly tested in their study, there is the implied corollary that the teacher who expects poor performance of a child may also communicate this expectancy through her attitudes and interactions with the child, and the child will perform poorly.

Since the initial publication, Rosenthal and Jacobson's study has been attacked on points of methodology, procedure, and analysis, and has been criticized as being unreproducible (cf. Barber and Silver, 1968; Elashoff and Snow, 1971; Fleming and Anttonen, 1971). Nonetheless, the hypothesized relationship between expectations and behavior remains convincing, so much so that in some circles it seems to have become a basic tenet in the social psychology of the classroom, no doubt owing in large part to the considerable popularization of the Pygmalion study. (See the book review by Snow, 1969, on the uncritical acceptance and popularization of the study; and Johnson, 1970, especially pp. 141-52.) Furthermore, it is widely assumed that teachers develop expectations about their students on the basis of perceived SES of

students (Arnez, 1966; Davidson and Lang, 1960; Miller et al., 1968; Rist, 1970; Yee, 1968), physical appearance (Bersheid and Walster, 1972), classroom demeanor, occupation of the father (Barclay et al., 1972), and numerous other traits.

It is surprising, then, that there has been relatively little empirical testing of this assumption, especially regarding perceived SES: research on the assumed relationship between teacher perceptions of SES and performance expectations is indeed rare. With the few exceptions mentioned above, the writing on teachers' presumed middle-class bias has been mainly impressionistic with little or no supporting data. The important study by Rosenthal and Jacobson did not actually measure the relationship between the two variables of perception and expectation, but was more concerned with examining the effect of experimentally induced positive expectations on performance. Nor have attempts to replicate Rosenthal and Jacobson dealt with the question of a link between perceptions and expectations. As in the original investigation, the replicative studies have concentrated on the consequences of experimentally induced expectations.

Hence the need for research to appraise the extent to which middle-class bias operates in the classroom. Do teachers, in fact, have higher expectations for middle- (and upper-) class children as is so widely assumed in the literature on expectancy effects? The study reported here attempts to answer this question by asking a sample of elementary school teachers to respond to a set of anonymous photographs of grade school children, indicating performance expectations and perceived social class of the children in the photos. Anonymous photographs were used in order to avoid obvious ethical problems inherent in designs that use real children.

The hypothesis tested was this: a teacher's expectations of a child's school performance vary directly with the teacher's perception of the child's social class. Thus, it should be emphasized, this is a purely descriptive study that seeks to determine if a relationship exists and, if so, to measure the direction and strength of that relationship. Should the major hypothesis prove tenable, we would have some support for a related proposition: teacher attitudes may be an important intervening variable between the child's socioeconomic background and his academic performance. To provide an answer to the more basic question, *why* a direct relationship exists between perceived SES and expectations and how these expectations are communicated, would go beyond the scope of the present study, although we will argue that our findings have import independent of this more basic question.

DATA AND METHOD

THE INSTRUMENTS

With the help of a photographer assistant, we obtained 75 final copy prints of black and white children between the ages of eight and twelve years presently attending elementary schools in the metropolitan area from which a sample of teachers was later drawn. Most of the photographs were taken in the vicinity of local schools on

an average school day when children were arriving at, or leaving, school. The rest of the photographs were taken near shopping centers and community areas served by these schools. The backgrounds were removed so the final copy prints resembled studio photographs.

To insure a representative range of SES, the initial pool of 75 photographs was judged by a panel of 12 elementary school teachers (not participants in the field study). Each panel member was asked to place each child's photograph into one of three classes for each racial category—lower, middle, and upper—on the basis of the child's general appearance in the full-length photo. The children most consistently chosen by the panel as representative of a given social class within the two racial categories were used in the study. The final set of photographs consisted of 9 white children (5 boys and 4 girls) and 9 black children (7 boys and 2 girls), with each race equally distributed among the three class categories, i.e., 3 lower-class white children, 3 lower-class black children, etc. It should be emphasized that the panel's SES ratings were used only as a screening device for reducing the larger set of 75 photographs to a more manageable and representative subset: in the study, itself, individual teacher's perceptions of SES were correlated with expectations of performance.

Perception of academic performance was operationally defined by the teacher's placement of a child into eight categories related to academic performance, social adjustment, and related behavioral areas as follows:

1. Books read independently.
2. Emotional instability.
3. Aspirations for a professional career.
4. Participation in extra-curricular activities.
5. School adjustment as it relates to age leaving school.
6. Juvenile crime and/or school offenses.
7. Completion of homework.
8. School performance as it relates to parental involvement.

These eight areas of performance are suggested in work by Coleman et al. (1966), Miller et al. (1968), Rodman (1963), Whiteman et al. (1967), and Ziegler (1970). Four of the eight (1, 3, 4, and 7) are generally regarded as characteristics currently associated with academic success and the remaining four are felt to be associated with academic failure. It should be emphasized that the performance criteria implied in these designations do not represent our personal views on success and failure in the educational experience. But they do seem to be entirely consistent with the dominant educational ideology found among teachers and administrators in the U.S. today. Furthermore, the fact that the kinds of behavior involved in these eight items are generally thought to be class-linked begs the question of why teachers are wont to employ them as a basis for deciding success and failure in the classroom. Indeed, we take the position that none of these eight characteristics is *intrinsically* linked to any particular social class. Rather, the social class of the child as perceived by the teacher may itself dictate what behavior is reinforced, which behavior is later interpreted as academic success or failure.

The eight performance items constituted Section I of the questionnaire used by the teachers to record their perceptions and expectations about the 18 students in the photographs. Teachers were asked to indicate which children in the alphabetically lettered photographs "read a good many books independently," "have been diagnosed as emotionally unstable," etc. Section II of the questionnaire contained three items for making judgments as to each child's overall level of motivation, IQ and SES. Children could be ranked as "high," "medium," and "low" on each of the three dimensions. On every item, teachers were given the option of indicating that they could not determine which children could be expected to perform in a certain way or which children possessed certain characteristics. They were also instructed not to respond to children whom they recognized.

Section III of the questionnaire elicited background information on the teachers: sex, age, race, years of teaching experience (both in their present school and total years experience), marital status, and occupation of spouse. It was felt that such information would prove useful when introduced as control variables for elaborating the primary relationship between perceived SES and expectations.

SAMPLE

Using census data and the researchers' first-hand knowledge of the city (population about 100,000) all elementary schools were classified into two categories: those serving predominantly lower-class areas of the city and those serving predominantly middle- and upper-class areas. Four schools, two from each category, were chosen at random. Each school was given a pseudonym indicating the SES of the community served by the school. Of the 108 faculty in these four schools, 8 could not be reached due to illness and other commitments and 4 refused to participate in the study. This left a total of 96 distributed by schools as follows: *Elite*, $n=26$; *Average*, $n=22$; *Project*, $n=28$; and *Poverty*, $n=20$. Of the total, 88 were white women, 4 were black women and 4 were white men. Consequently, race and sex could not be used as variables due to the small numbers involved.

PROCEDURE

Working through the local school board, cooperation of the principals in the four schools was obtained and dates were set for one of the investigators to meet with the assembled faculty in each school for the purpose of administering the test. All data from the four schools were collected within a ten-day period. The procedure followed at each school involved:

1. Introduction of the investigator by the principal.
2. A brief presentation of a cover story which described the research as a study of "primary perceptions," or "the ability of people to make accurate estimates of the characteristics and performance capabilities of others on the basis of first impressions."

3. Distribution of an instruction sheet, supplemented by verbal instructions, which explained the nature of the task and how to use the questionnaire, which was then distributed.

The 18 photographs had been mounted on a large folding board with one-half of the board containing the photos of the 9 white children and the other half the 9 black children. Racial segregation of the photos was employed as both a diversionary tactic to increase the validity of responses touching social class and as a means of determining differences in perception by race.¹ Because the photographs were small (5" × 7") and because many respondents needed to examine the photos closely, it was impractical to test all teachers within a school at one time; therefore, teachers were asked to choose among themselves the order in which they would take the test. Test groups varied in size from 4 to 10 teachers with mean group size of 6. Test subjects sat in chairs near the photo board but were free to stand and/or move closer to examine the photographs if they desired. After all subjects at a given school had completed the task, they were again assembled and debriefed. All teachers at a given school were tested in a single afternoon.

To determine if the order in which the two sets of photographs were presented influenced perceptions, Set I (white children) was presented first to a third of the total sample of 96 teachers and Set II (black children) was presented first to another third. In both conditions, the second set remained hidden from view until all subjects in a test group were finished with the first set. The other third of the sample was shown Sets I and II simultaneously.

METHOD OF ANALYSIS

For purposes of data analysis, each of the eight performance categories of Section I of the questionnaire was taken as indicating either success or failure depending on whether the behavior would ordinarily lead to success or failure in school as judged by current standards of performance. Thus, for example, reading books independently (Item 1) and completing homework regularly (Item 7) were designated "success" categories, while having been "in trouble" at school or having been involved in juvenile crime (Item 6) was designated a "failure" category. The eight items were equally divided between the two categories with each item's designation shown in Table 3.

The Kolmogorov-Smirnov one sample test was used to test the null hypothesis that teachers' assignments of children to both "success" and "failure" categories were independent of teachers' perceptions of children's SES. A series of 3×1 frequency distributions (one for each of the eight items in Section I) showing aggregated teacher SES ratings for all children selected on that item were constructed for comparison by the K-S test with the theoretical distribution based on random assignment on SES. Table 1 is an example of one such distribution in which teachers' ratings from all four schools are combined. The total of 312 shown in Table 1 represents the number of selections from the nine photos of white children made by the

total sample of teachers on the success item "books read independently." The mean number of selections per teacher for this particular performance category was 3.25. This teacher average response pattern (TARP) serves as an indicator of the extent to

TABLE 1. *Illustrative 3 × 1 Distribution Showing K-S Analysis for SES and Children Selected in Response to Item 1, Reading Books Independently (all 4 schools combined): White Children*

Perceived SES of Child	Number of Children Perceived as Fitting this Category
Lower	18
Middle	120
Upper	174
Total	312

N = 96 *Critical D* = .09
Observed D = .27 (for significance at *p* = .01)
TARP = 3.25 (responses per teacher)

which teachers were willing and/or able to make selections for the various performance categories.² Thus, a TARP score of 3.25 indicates that slightly more than one-third of the maximum of 9 children were selected as positive on this trait.

Goodman and Kruskal's (1954) measure of association, *Gamma*, was used in the analysis of the ordinal variables of perceived motivation, IQ and SES, all scored as "high," "medium," and "low." An illustrative cross-tabulation showing the relationship between perceived SES and perceived motivation is given in Table 2. Here

TABLE 2. *Illustrative 3 × 3 Bivariate Distribution for Measuring Association Between SES and Overall Motivation (all 4 schools combined): Black Children*

Perceived SES of child	Motivation Level of Child			Totals
	Low	Average	High	
Lower	239	65	2	306
Middle	43	252	62	357
Upper	6	11	145	162
Totals	288	328	209	825

N = 96
Gamma = .93
TARP = 8.59 (responses per teacher)

individual teacher rankings on both variables are aggregated and combined in a single 3 × 3 table which shows the overall degree to which teachers' rankings on SES are related to their rankings on motivation. The TARP score for bivariate distributions of this type indicates the extent to which teachers were willing and/or able

to rank the pictured children on both variables and should be compared with the maximum of nine which would result if every teacher ranked all nine photos on both variables.

The data from sections I and II were analyzed separately for each school by race. Scores were then combined for the two lower-class schools (Project and Poverty) and the two middle- and upper-class schools (Elite and Average). Finally, data from all four schools were combined to give an overall view of teacher perceptions. However, due to space limitations and the large number of relationships generated in the analysis, results for the individual schools are not reported here. For the same reason, only summary statistics, including *D* values of the K-S tests and *Gamma* coefficients, are presented for the various tabulations.

RESULTS

Table 3 reports the results of the analysis of relationships between SES and the eight performance categories. Shown here are the differences between observed *D* values of the K-S test and the critical *D* values associated with the .01 level of significance. In every case the observed *D* exceeds the critical *D* and is thus significant at the .01 level or beyond. Given both our initial concern that teachers would be reluctant to make either positive or negative judgments about children on the basis of their photographs alone and the potentially stigmatizing and stereotypical nature of the kinds of judgments they were asked to make, such overwhelming rejection of the null hypothesis was not anticipated. In fact, out of the total of 192 K-S tests that were run there were only seven comparisons for which the probability of the observed *D* value exceeded .01: these occurred in the case of judgments about black children (Set II), all made by teachers in the two lower-class schools.

While these results are quite consistent in their support of the major hypothesis under investigation, two interesting patterns of variation in them may be noted. First, the strength of the relationships (between assignment to success-failure categories and perceived SES) varies as roughly gauged by the size of the difference between the observed *D* and its corresponding critical *D*. Examination of these *D*-value differences in Table 3 reveals that, in general, they are larger for failure items than for success items. This trend is summarized in the average *D*-value differences reported at the bottom of Table 3, Sets I and II.³

Note also that the spread of average *D*-value differences between success and failure items is greater for the black children. This reflects the tendency of teachers to respond differentially to the two races: perceived SES was more strongly associated with failure among the black children while being more strongly associated with success among the white children. Virtually all of the black children selected as potential failures were perceived as being lower SES and almost all of the white children expected to succeed were concentrated in the middle- and upper-SES categories.

A second pattern which was found in the data on which Table 3 is based concerns

TABLE 3. Relationships Between SES and Eight Performance Characteristics*

Item	Type of Item	D-value Differences for Middle/Upper Class Schools (Elite & Average)	D-value Differences for Lower Class Schools (Project & Poverty)	D-value Differences for all 4 Schools
<i>Set I (White Children)</i>				
1. Books read	Success	.16†	.12	.18
2. Emotional instability	Failure	.26	.16	.27
3. Professional career	S	.18	.15	.21
4. Activities	S	.19	.12	.20
5. Poor adjustment	F	.36	.37	.39
6. Trouble-crime	F	.29	.25	.32
7. Homework	S	.18	.18	.22
8. Parental non-involvement	F	.34	.33	.36
Average D-value difference:				
Success items:		.18	.14	.20
Failure items:		.31	.28	.33
<i>Set II (Black Children)</i>				
1	Success	.15	.05	.15
2	Failure	.33	.36	.38
3	S	.18	.07	.17
4	S	.15	.07	.15
5	F	.35	.31	.37
6	F	.30	.38	.38
7	S	.16	.11	.16
8	F	.34	.35	.38
Average D-value difference:				
Success items:		.16	.07	.16
Failure items:		.33	.35	.38

*Difference between observed and critical Ds all significant at .01 level or beyond by Kolmogorov-Smirnov test.

†D-value difference is the difference between observed D-value on this item (.28) and the Critical D-value based on N, the number choices on this item, (.12) or .28 - .12 = .16.

TABLE 4. *Teacher Average Response Pattern (TARP) to Eight Performance Characteristics*

Item	Type of Item	Elite & Average	Poverty & Project	All four Schools Combined
<i>Set I (White Children)</i>				
1	Success	3.69	2.80	3.25
2	Failure	2.85	1.79	2.32
3	S	3.69	2.60	3.14
4	S	4.56	3.48	3.83
5	F	2.85	2.10	2.48
6	F	2.71	2.00	2.35
7	S	4.81	3.79	4.24
8	F	3.10	2.52	2.81
Average by school		3.53	2.64	3.06
Item average				
Success:		4.19	3.16	3.62
Failure:		2.88	2.10	2.49
<i>Set II (Black Children)</i>				
1	Success	3.50	2.83	3.16
2	Failure	3.17	1.75	2.46
3	S	3.46	2.56	3.01
4	S	4.52	3.70	4.11
5	F	3.15	2.58	2.86
6	F	3.00	2.19	2.59
7	S	4.54	3.54	4.04
8	F	3.56	2.70	3.13
Average by school		3.55	2.96	3.22
Item average				
Success:		4.00	3.16	3.58
Failure:		3.22	2.30	2.76

the average number of responses per teacher (TARP). The TARP values presented in Table 4 reveal two general trends: (1) the average number of responses or choices made is larger among teachers in middle-upper-class schools than in lower-class schools; and (2) the averages are larger for success items than failure items. Although not as pronounced, there is also evidence of *higher* TARP values in respond-

ing to the photos of black children compared to white children when selecting those expected to fail (Items 2, 5, 6, and 8) and *lower* TARP values for blacks than whites when selecting children who are expected to succeed (Items 1, 3, 4, and 7). That is, regardless of perceived SES, white children were more often expected to succeed and black children more often expected to fail as indicated by the number of photos selected from each racial group for the two kinds of item.

TABLE 5. *Associations Between SES, Overall Motivation and IQ: Gammas*

Variables	Middle-Upper Class Schools (Elite & Average)	Lower Class Schools (Poverty & Project)	All four Schools Combined
<i>Set I (White Children)</i>			
Perceived SES × overall motivation	.94	.88	.94
Perceived SES × IQ	.96	.93	.95
<i>Set II (Black Children)</i>			
Perceived SES × overall motivation	.93	.93	.93
Perceived SES × IQ	.96	.94	.95

Equally significant findings emerged from the analysis of responses to items dealing with motivation and IQ. *Gamma* coefficients measuring the strength of the association between motivation and IQ and perceived SES are given in Table 5. With one exception, these coefficients are all larger than .90, indicating very strong positive relationships. The TARP values, not reported to conserve space, display the same general pattern noted above: teachers in the middle-upper-class schools made slightly more choices, i.e., judged more photos on motivation and IQ than their counterparts in the two lower SES schools. Also, a slight racial difference appears here in the form of generally fewer rankings of black children compared with white, with the most pronounced difference occurring in Poverty and Project schools where teachers judged an average of 8.60 white children versus 8.20 black children on overall motivation, and 7.85 white versus 7.50 black on IQ.

CONTROL VARIABLES

Because many respondents failed to answer questionnaire items on background characteristics, we were unable to introduce as many control variables related to teacher characteristics as had been planned. Missing data were most common for items on age, number of years teaching in present assignment, where taught before, and occupation of spouse (typically the husband). We have no explanation for our subjects' failure to respond to these items, except that they may have felt that such questions were too personal or could have been used for purposes of identification. However, for some unknown reason, response to the question on total years of teaching experience was fairly complete and this variable was used as a control.

The entire sample was divided at the median for the distribution on this item, creating the dichotomy: "6 years or less" and "over 6 years" teaching experience. Data were treated as before. As may be seen in Table 6, all relationships were significant at the .01 level or beyond with the exception of the four success items for black children involving teachers in lower-class schools who had taught for 6 or less years. In this special situation, perceived SES had no bearing on whether a child was expected to succeed.

Also, as in the zero-order K-S tests, differences between observed and critical *D* values tended to be greater for failure items than for success items (indicating higher significance). This pattern of *D*-value differences is most pronounced for teachers with over 6 years of teaching experience, especially when responding to the photos of black children. It would appear from these trends in *D*-value differences that the more experienced teachers more often regarded black children whom they expect to fit failure categories as also being from the lower SES category than do teachers with less experience. When selecting for failure among blacks, social class is perhaps a more salient characteristic for teachers who have taught longer.

The average number of responses per teacher (TARP) controlled by length of teaching experience reveals several interesting trends as shown in Table 7. First, as mentioned previously, teachers were less willing to respond to failure items than to success items, with the difference between response rate to the two types of items being generally greater for the more experienced teachers except when responding to black children in the middle- and upper-class schools. (Compare item averages given below the two main bodies of Table 7.) Second, independent of length of teaching experience, the overall rate of response to both types of item is seen to vary directly with the SES level of the school. Closer examination of the response patterns within the two SES levels of school, controlled for years of experience, indicates that teachers in *lower-class schools* with 6 years or less experience consistently made fewer choices than the over 6 group in response to both black and white children. This pattern persists for the less experienced teachers in *middle-upper-class schools* when judging black children for failure items; the pattern is reversed, however, in all other instances involving middle- and upper-class schools, i.e., less experienced teachers in these schools exhibited higher TARP values compared to the over 6 years group when judging whites for both success and failure and blacks for success.

TABLE 6. Relationships Between SES and Eight Performance Characteristics, Controlled for Years of Teaching Experience

Item	Type of Item	6 Years & Under			Over 6 Years		
		D-value Differences for		All 4	D-value Differences for		All 4
		Middle-Upper Class Schools	Lower-Class Schools		Middle-Upper Class Schools	Lower-Class Schools	
<i>Set I (White Children)</i>							
1	Success	.11	.09	.16	.09	.04	.12
2	Failure	.24	.03	.19	.16	.24	.27
3	S	.13	.10	.20	.10	.09	.15
4	S	.13	.09	.10	.15	.08	.16
5	F	.34	.28	.38	.29	.31	.37
6	F	.24	.14	.18	.26	.23	.31
7	S	.13	.14	.19	.13	.13	.17
8	F	.30	.18	.31	.28	.33	.38
Average D-value difference:							
	Success items:	.12	.10	.16	.12	.08	.15
	Failure items:	.32	.16	.26	.25	.28	.33
Average D-value difference between success and failure items		.20	.06	.10	.13	.20	.18
<i>Set II (Black Children)</i>							
1	Success	.10	.01*	.09	.10	.03	.12
2	Failure	.23	.24	.29	.33	.28	.36
3	S	.14	.00*	.13	.10	.08	.15
4	S	.13	.00*	.12	.05	.05	.10
5	F	.28	.24	.32	.30	.32	.37
6	F	.20	.32	.29	.27	.35	.37
7	S	.11	.02*	.12	.13	.06	.13
8	F	.25	.23	.30	.30	.37	.39
Average D-value difference							
	Success items:	.12	.00	.12	.10	.06	.12
	Failure items:	.24	.26	.30	.30	.33	.37
Average D-value difference between success and failure items		.12	.26	.18	.20	.27	.25

*Denotes significance at $P \geq .05$

TABLE 7. *Teacher Average Response Pattern (TARP) to Eight Performance Characteristics, Controlled for Years of Teaching Experience*

Item	Type of Item	6 Years and Under			Over 6 Years		
		Poverty and Project	Elite and Average	All Four Schools	Poverty and Project	Elite and Average	All Four Schools
<i>Set I (White Children)</i>							
1	Success	2.41	3.69	3.10	3.19	3.75	3.43
2	Failure	1.91	3.08	2.54	1.69	2.50	2.04
3	S	2.23	4.31	3.35	2.85	3.10	2.96
4	S	3.30	4.60	4.02	3.85	4.55	4.15
5	F	1.90	2.90	2.46	2.23	2.60	2.39
6	F	1.82	2.88	2.39	2.12	2.30	2.20
7	S	3.68	4.88	4.33	4.04	4.70	4.33
8	F	2.27	3.40	2.89	2.76	2.55	2.67
School average		2.44	3.72	3.13	2.84	3.26	3.02
(Item average)							
Success:		2.90	4.37	3.70	3.48	4.02	3.72
Failure:		1.98	3.06	2.57	2.20	2.49	2.32
Difference		.92	1.31	1.13	1.28	1.53	1.40
<i>Set II (Black Children)</i>							
1	Success	2.36	3.65	3.06	3.15	3.60	3.35
2	Failure	1.50	2.92	2.27	2.04	3.25	2.57
3	S	2.09	3.77	3.00	2.96	3.05	3.00
4	S	3.27	4.65	4.02	3.96	4.45	4.17
5	F	2.32	3.08	2.73	2.88	3.10	2.98
6	F	2.14	2.77	2.48	2.20	3.30	2.70
7	S	2.90	4.73	3.90	3.88	4.35	4.09
8	F	2.45	2.27	2.73	2.57	3.25	2.87
School average		2.38	3.48	3.02	2.96	3.55	3.22
(Item average)							
Success:		2.66	4.20	3.50	3.49	3.86	3.65
Failure:		2.10	2.76	2.55	2.42	3.22	2.78
Difference		.56	1.44	.95	1.07	.64	.87

With regard to the bivariate relationships involving perceived SES x motivation and SES x IQ, the more experienced teachers were found to exhibit generally higher associations but falling in the range of *Gammas* reported in Table 5. The difference in strength of association appeared to be greatest when judging photos of black children. All in all, though, the introduction of years of teaching experience as a control did little to elaborate the relationships revealed in the original *Gammas*. This control did produce some interesting differences in the related TARP values, however. Essentially the pattern of response rates here was the same as the previously mentioned pattern for selections on the eight success–failure items: the average number of photos of both black and white children ranked on motivation and IQ varied directly with length of teaching experience within the two lower SES schools, with the most marked difference occurring in selections on IQ where the less experienced group selected an average of 6.96 white children (out of a possible 9 photos) compared to an average of 8.76 for the more experienced group; this difference was even greater for selections on black children, 6.32 compared to 8.46. For teachers in the two middle-upper SES schools, years of experience made little or no difference in TARP values.

DISCUSSION

Before discussing the interpretation of our empirical findings, we must comment on the study design and alternative explanations of results obtained in a study of this type. First, one could argue that in an actual classroom a teacher has many cues from each child from which to form an empirically based judgment. In this design, where the only information available was that in the photographs, some stereotyping would be both inevitable and more obvious than in a natural setting. Some stereotyping, then, stems from the gap between research design and reality. However, the degree, direction, and consistency of the stereotyping and the willingness to do so went far beyond the limits expected by chance. We assume that teachers are more cautious in making judgments in their own classroom although the accuracy of these judgments is still an empirical question. But if teachers were willing to make judgments from a photograph alone, might they not continue to make judgments with additional cues, so promoting a self-fulfilling prophecy?

Second, there is a strong possibility that findings on the relationships between SES and teachers' expectations were distorted by a strong response set or repetitive response tendency. To determine if such a response set was operating, approximately one-fourth of the sample was randomly selected for questioning about the method by which they responded. Several of these subjects stated that they were not certain how they reached their decisions, suggesting in a few cases that their choices may have been random at times. But most of these 25 subjects indicated that they based their choices on such criteria as similarity to former students or other acquaintances and facial expression. None of these teachers felt that his/her answers were repetitive and most were adamant that the children they selected for *each* item truly

fit that characteristic. Although we made no statistical checks on response set, this substudy forces us to conclude that response set does not explain the highly consistent findings of this study.

A third explanation involves the strong possibility that subjects' responses reflected associations between physical attractiveness and performance characteristics rather than between perceived SES and performance. Physical attractiveness had not been considered as a potential confounding variable until the recent research of Berscheid and Walster (1972) demonstrated the significant impact of attractiveness on perceptions. Unfortunately the test instrument had already been developed and introduced to the sample when the Berscheid and Walster report appeared: thus no data were collected on teachers' perceptions of attractiveness. However, in a crude effort to assess the possible underlying influence of this variable on our results, a small availability sample ($n=20$) of the original 96 subjects was retested approximately six weeks later using the same instrument with the addition of an item that required the respondent to scale each photograph from 1-6 on a physical attractiveness continuum. This scale was then collapsed into a trichotomy of "very attractive," "averagely attractive," and "unattractive" and intercorrelated with perceived IQ, motivation, and SES. The six resulting *Gammas* (three for each racial category) were considerably lower in magnitude than those reported earlier, falling below an absolute value of .15 with the exception of SES x attractiveness for black children which was .30. In addition, K-S tests were run for each of the eight success-failure items with physical attractiveness and none was significant beyond the .10 level.

Obviously the negative results of this restudy do not wholly eliminate the possible importance of physical attractiveness in this investigation. The sample was small and the members had already been debriefed. It should be noted, however, that earlier research on this important variable (e.g., Berscheid and Walster, 1972) did not consider the possibly confounding effect of SES. It is certainly likely that SES, physical attractiveness, and perhaps a host of other variables, as yet unspecified, affect perceptions and the expectations people have of one another. (For a recent systematic attempt to assess the independent and joint effects of such variables on teacher expectations see Kehle, 1972.) In an undeveloped area of research with a limited literature it can be expected that crucial variables will often be ignored. We hope that future teacher expectancy research will not overlook the crucial variable of physical attractiveness.

Though the foregoing rival hypothesis, and others, cannot be ruled out entirely, given the somewhat limited data of this study, we take the position that, in the schools we studied, there is a rather strong relationship between teachers' perceptions of social class and their expectations on a number of performance and ability characteristics in children. To what extent the overwhelming evidence against the null hypothesis resulted from such unknowns as experimenter effects (Rosenthal, 1966), reactive effects of the test itself (Campbell and Stanley, 1963), and the demand characteristics of the experiment (Orne, 1962), we cannot know. It does seem unlikely, however, that the findings simply reflect contaminants of this kind. While

their removal would undoubtedly have an attenuating effect on the results, it seems improbable that they would be washed out completely.

Turning our attention to the details of the findings, we summarize and interpret several secondary results that may enhance our understanding of the major finding, the relationship between SES and expectations. We are referring to the patterning of TARP values by type of item, SES of school, teacher experience, and race of the children in the photos.

TYPE OF ITEM

As we commented earlier, the willingness and ability with which the subjects categorized and rated the photographs was unexpected. Yet there was a noticeable tendency for TARP values to be lower for failure items than success items. We interpret this as a reflection of the teachers' reluctance to make negative judgments as compared to positive judgments. It is significant that this reluctance was more common when judging white children than black children: differences between teacher average response rates for success and failure items were greater for white children.

SES OF SCHOOL

Perhaps more indicative of differences in teachers' willingness to respond and the underlying biases this reflects is the pattern of TARP values by school SES. Consistently, teachers in the two lower-class schools made fewer selections than teachers in the two middle-upper-class schools—to all items. We have no ready explanation for this finding. In fact, the pattern is contrary to the generally presumed negative effects of teaching in lower-class schools, i.e., such settings tend to strengthen class bias. One possible explanation which may be unduly optimistic is that the pattern was a manifestation of the "liberalizing" or "humanizing" effect of teaching in lower-class schools: as a result of their exposure to the hard realities of class differences, these teachers, who are middle class in terms of their achieved, if not their ascribed, status were less willing to make stereotypical judgments. Note, however, that the SES difference in TARP values was greater for white children than black children, thus somewhat qualifying our interpretation, i.e., the humanizing effect does not completely eliminate the presumed racial bias of middle-class teachers.

TEACHER EXPERIENCE

As might be expected, our findings show generally that the less teaching experience a teacher had, the less likely she was to make the kinds of judgments called for in this study. In part this tendency could result from the relative recency of her teacher training in college or, more generally, her total college experience, assuming that this training and/or total experience has a generally humanizing effect. And, either because this effect eventually wears off with time in service, or because older

cohorts of teachers did not receive the kinds of training that sensitized them to the relevant humane values, we observed that teachers with over six years experience had higher average response rates.

This interpretation is somewhat uncertain since the relationship between TARP and teaching experience appears to be contingent on: school SES; whether judgments are for success or failure items; and whether the children being judged are black or white. Recall that under certain conditions the less experienced group exhibited *higher* TARP values—among teachers in the middle-upper-class schools when judging white children for both success and failure and black children for success. This reversal of the general effect of experience could suggest that the hypothesized “humanizing” effects only manifest themselves when new teachers find themselves in an alien teaching situation in a lower-class school—a setting in which they encounter numerous conflicts with the deep-seated middle-class stereotypes and expectations which teacher training has failed to erase, or may, actually, have nurtured. A new teacher, finding herself in a comfortable middle-class school where in fact most black children and poor children have difficulty, may feel more confident in her role since little happens to invalidate her middle-class view. In short, the environment is not conducive to the germination and growth of the seeds of humane and liberal openmindedness that conceivably were sown during her training.

By contrast, the new teacher who enters a school serving a predominantly lower-class community may be more aware that his/her stereotyped expectations are in conflict with reality in the sense that not all of his/her students “fail.” With this awareness they may be more unsure of themselves and hesitant to prejudge children. Their generally lower associations between SES and IQ, SES and motivation, and lower TARP values would support this hypothesis. Even more persuasive in this regard is the fact that the only statistically nonsignificant K–S tests of the social class distribution on the eight success–failure items were for less experienced teachers in lower-class schools when selecting black children for success categories.

Although the more experienced group of teachers in lower-class schools exhibited to a smaller degree the same reluctance to respond and to relate expectations to SES, it is worth emphasizing that either through an increase in self-confidence in their role, or as a result of a waning of the liberal idealism of their earlier years in service, they did appear more like their counterparts in middle-upper class schools. We cannot determine from the data of this investigation whether this trend for experience results from increased confidence in the ability to prejudge children, the possible deliberalizing effect or increasing close-mindedness that comes with time in service, or more simply a qualitative difference in the training received by our more experienced group compared to the younger, more recent graduates of our less experienced group. This question would best be answered by a longitudinal study comparing teachers trained in various types of programs and working in different types of school.

CONCLUSIONS

Inherent in the foregoing discussion is the assumption that virtually all teachers enter their profession with many similar ideas, background experiences, and preconceptions about schooling. Their orientation is predominantly middle class in terms of their own career aspirations, the type of student they are likely to enjoy teaching, and, in most cases, the type of professional training they have received (Anderson, 1969; Charters, 1963; McGuire and White, 1957). What we find disturbing in the results of this study is not the presumed class identity of teachers (for it to be otherwise would require a major adjustment in the traditional occupational, professional, and class structures of American society), but rather that this identity appears to exert such a strong influence on the expectations teachers have about children. It would appear from the empirical results reported here that this influence is indeed a strong one in the elementary schools which we studied. Whatever the professional training of the teachers in these schools, it does not seem that much was accomplished in the way of sensitizing them to the possible effects of their basic middle-classness.

It must be emphasized that this has been a descriptive study not intended to establish causal relations. The interpretation of the data offered here can only suggest possible causal factors underlying the middle-class bias as it appeared to operate in the four schools we studied. To demonstrate, empirically, the existence of this bias does not, however, allow one to conclude that it negatively affects the performance or success chances of certain children. Only if the bias is communicated to under-classed and minority students could the self-fulfilling prophecy become a tragic reality.

Thus, one can best conceptualize teachers' expectations as related to their students' SES as one of many intervening variables that contribute to the complex relationship between class background and the educational experience. Even granting that the causal influence of teachers' expectations on the child's academic performance is at best minimal, we should still be concerned about the child's psychological well-being. Teachers' expectations surely must have some bearing on the child's feelings of worth. As called for by Jencks et al. (1972) in their emphasis on "short term" versus "long term" effects of school, teacher expectations would seem to be a variable whose control or modification would undoubtedly result in a marked improvement in the "internal quality of life" in our schools.⁴

FOOTNOTES

⁴Current social values regard discrimination by race or SES class as wrong. If teachers felt that they were being directly questioned on class discrimination, many might be unwilling to make candid judgments that discriminate by social class, especially if they believed that their discrimination might be viewed by others as undesirable. Further, since it seemed probable that many teachers might look for a hidden purpose beyond the cover story about "primary perceptions," an obvious racial segregation of the pictures was employed to sate their curiosity. It was intended that this seemingly overt racial orientation would cause the teachers to view the questions related to SES class as having racial overtones. Thus we

hoped they would overlook or fail to compensate for any social class bias in their attempt to respond to the photographs in a socially (racially) acceptable way. Segregation by race also made it possible to observe any class distinctions and/or differences in overall attitudes and expected performance levels between racial groups.

²It should be noted that this method of computing response averages confounds the number of teachers giving a rating with the number of photos each teacher rates. This is not felt to be a serious problem, however, because inspection of the raw data indicated that of these two sources of variation in TARP scores, the latter—differences in the number of photos rated by each teacher—is the main source. Thus TARP scores more accurately reflect differences between subsamples in the average number of responses made per teacher than differences in the number of teachers responding.

³As a general rule, the critical *D* value in the K-S test decreases as *N* increases, *N* being the total number of responses made by the subjects to a particular item. Thus, since teachers tended to make more responses to success items than to failure items, there would be a corresponding tendency for the difference between the critical *D* value and observed *D* value to be larger for success items, all other things being equal. However, it is also the general rule that observed *D* values tend to increase as responses become clustered in a single category with the result that the difference between critical *D* and observed *D* is inflated. Such clustering of responses appears to be present in the case of failure items due to the fact that teachers distributed their success choices into both the middle- and upper-class categories but their failure choices were almost always assigned to the lower class only. These two parallel tendencies, the one leading to inflation of the difference between observed and critical *D*s for success items and the other producing the same inflation in the case of failure items, appear to offset one another, making it appropriate to compare *D*-value differences for the two types of items.

⁴For a recent report on an apparently successful attempt to improve the educational experience of children by raising *their own* performance expectations see Entwisle and Webster (1973).

REFERENCES

- Anderson, James G. 1969. *Teachers of Minority Groups: The Origins of Their Attitudes and Instructional Practices*. Las Cruces: New Mexico State University Press.
- Arnez, N. L. 1966. "The Effect of Teacher Attitudes upon the Culturally Different." *School and Society* 94:149-52.
- Barber, T., and M. Silver. 1968. "Pitfalls in Data Analysis and Interpretation: A Reply to Rosenthal." *Psychological Bulletin Monograph* 6:48-62.
- Barclay, J. R., W. E. Stilwell, and L. Barclay. 1972. "The Influence of Paternal Occupation on Social Interaction Measures in Elementary School Children." *Journal of Vocational Behavior* 2:433-46.
- Berscheid, E., and E. Walster. 1972. "Beauty and the Beast." *Psychology Today* 5 (March): 42-6.
- Campbell, D. T., and J. C. Stanley. 1963. *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.
- Charters, W. W., Jr., 1963. "The Social Background of Teaching." In N. L. Gage (ed.), *Handbook of Research on Teaching*. Chicago: Rand McNally.
- Coleman, James S. 1966. *Equality of Educational Opportunity*. U. S. Department of Health, Education and Welfare, Office of Education. Washington: Government Printing Office.
- Davidson, H., and G. Lang. 1960. "Children's Perceptions of Their Teachers' Feelings Toward Them Related to Self-Perception, School Achievement and Behavior." *Journal of Experimental Education* 12:107-18.
- Elashoff, J. D., and R. E. Snow. 1971. *Pygmalion Reconsidered*. Worthington, Ohio: Jones.
- Entwisle, D. R., and M. Webster, Jr. 1973. "Status Factors in Expectation Raising." *Sociology of Education* 46 (Winter):115-26.
- Fleming, E. S., and R. G. Anttonen. 1971. "Teacher Expectancy or My Fair Lady." *American Educational Research Journal* 8:241-52.
- Goodman, L. A., and W. H. Kruskal. 1954. "Measures of Association for Cross-Classifications." *Journal of the American Statistical Association* 49 (December):732-64.

- Jencks, Christopher *et al.* 1972. *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books
- Johnson, David W. 1970. *The Social Psychology of Education*. New York: Holt, Rinehart & Winston.
- Kehle, T. J. 1972. "Effect of the Student's Physical Attractiveness, Sex, Race, Intelligence, and Socioeconomic Status on Teachers' Expectations for the Student's Personality and Academic Performance." Unpublished Ph.D. dissertation, University of Kentucky.
- McGuire, C. M., and G. D. White. 1957. "Social Origins of Teachers in Texas." In L. J. Stiles (ed.), *The Teacher's Role in American Society*. New York: Harper & Bros.
- Miller, C. K. *et al.* 1968. "Socioeconomic Class and Teacher Bias." *Psychological Reports* 23:806.
- Orne, M. T. 1962. "On the Social Psychology of the Psychological Experiment: With Particular Reference to the Demand Characteristics and Their Implications." *American Psychologist* 17:776-83.
- Rist, R. C. 1970. "Student Social Class and Teacher Expectations: The Self-Fulfilling Prophecy in Ghetto Education." *Harvard Educational Review* 40 (August):411-51.
- Rodman, H. 1963. "The Lower-Class Value Stretch." *Social Forces* 42(December): 205-15.
- Rosenthal, Robert. 1966. *Experimenter Effects in Behavioral Research*. New York: Appleton-Century-Crofts.
- _____. 1973. "The Pygmalion Effect Lives." *Psychology Today* 7 (September):56-63.
- Rosenthal, Robert, and Lenore Jacobson. 1968. *Pygmalion in the Classroom*. New York: Holt, Rinehart & Winston.
- Snow, R. E. 1969. "Unfinished Pygmalion." *Contemporary Psychology* 14 (April):197-9.
- Whiteman, M., B. R. Brown, and M. Deutsch. 1967. "Some Effects of Social Class and Race on Children's Language and Intellectual Abilities." In M. Deutsch *et al.*, *The Disadvantaged Child: Selected Papers of Martin Deutsch and Associates*. New York: Basic Books.
- Yee, A. 1968. "Interpersonal Attitudes of Teachers and Advantaged and Disadvantaged Pupils." *Journal of Human Resources* 3:327-45.
- Ziegler, E. 1970. "Social Class and the Socialization Process." *Review of Educational Research* 40 (February):87-110.

LINKED CITATIONS

- Page 1 of 2 -



You have printed the following article:

The Relationship between Child's SES and Teacher Expectations: A Test of the Middle-Class Bias Hypothesis

Dale G. Harvey; Gerald T. Slatin

Social Forces, Vol. 54, No. 1. (Sep., 1975), pp. 140-159.

Stable URL:

<http://links.jstor.org/sici?sici=0037-7732%28197509%2954%3A1%3C140%3ATRBCSA%3E2.0.CO%3B2-B>

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

Footnotes

⁴ **Status Factors in Expectation Raising**

Doris R. Entwisle; Murray Webster, Jr.

Sociology of Education, Vol. 46, No. 1. (Winter, 1973), pp. 115-126.

Stable URL:

<http://links.jstor.org/sici?sici=0038-0407%28197324%2946%3A1%3C115%3ASFIER%3E2.0.CO%3B2-7>

References

Status Factors in Expectation Raising

Doris R. Entwisle; Murray Webster, Jr.

Sociology of Education, Vol. 46, No. 1. (Winter, 1973), pp. 115-126.

Stable URL:

<http://links.jstor.org/sici?sici=0038-0407%28197324%2946%3A1%3C115%3ASFIER%3E2.0.CO%3B2-7>

Teacher Expectancy or My Fair Lady

Elyse S. Fleming; Ralph G. Anttonen

American Educational Research Journal, Vol. 8, No. 2. (Mar., 1971), pp. 241-252.

Stable URL:

<http://links.jstor.org/sici?sici=0002-8312%28197103%298%3A2%3C241%3ATEOMFL%3E2.0.CO%3B2-0>

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

- Page 2 of 2 -



Measures of Association for Cross Classifications

Leo A. Goodman; William H. Kruskal

Journal of the American Statistical Association, Vol. 49, No. 268. (Dec., 1954), pp. 732-764.

Stable URL:

<http://links.jstor.org/sici?sici=0162-1459%28195412%2949%3A268%3C732%3AMOAFC%3E2.0.CO%3B2-0>

The Lower-Class Value Stretch

Hyman Rodman

Social Forces, Vol. 42, No. 2. (Dec., 1963), pp. 205-215.

Stable URL:

<http://links.jstor.org/sici?sici=0037-7732%28196312%2942%3A2%3C205%3ATLVS%3E2.0.CO%3B2-J>

Interpersonal Attitudes of Teachers and Advantaged and Disadvantaged Pupils

Albert H. Yee

The Journal of Human Resources, Vol. 3, No. 3. (Summer, 1968), pp. 327-345.

Stable URL:

<http://links.jstor.org/sici?sici=0022-166X%28196822%293%3A3%3C327%3AIAOTAA%3E2.0.CO%3B2-2>