

Joan S. Stark

Malcolm A. Lowther

Predictors of Teachers' Preferences Concerning Their Evaluation

Personal backgrounds, skills, and attitudes toward professional success reported by experienced classroom teachers modestly predicted preferences for ways of evaluating their work. Teachers view their work as an amalgam: a labor, a profession, a craft, and an art. Achieving consensus regarding teacher evaluation may require multifaceted approaches that recognize varied teacher needs and backgrounds.

Joan S. Stark is Professor of Higher Education, and Malcolm A. Lowther is Professor of Education in the School of Education, The University of Michigan.

Based in part on recent national reports, the public seeks more effective teacher evaluation systems aimed at: (1) eliminating incompetent or ineffective teachers; (2) providing incentive systems for improved teacher performance; (3) improving skills of less effective teachers; and (4) recruiting qualified teachers. Recent initiatives by the states frequently leave the details of evaluation procedures for local development. Consequently, school districts may resort to variations of three possible implementation approaches: (1) direct imposition of evaluation procedures with little or no teacher consultation; (2) negotiation of evaluation procedures within a collective bargaining contract; and (3) broad involvement of teachers and consultants in the development of evaluation procedures.

Considerable professional literature has focused both on criteria for evaluating teachers and on the qualifications of evaluators.¹ In addition, investigators have analyzed the utilization difficulties of contemporary evaluation schemes so as to improve their effectiveness.² Even so, national publicity has highlighted the resistance of teacher professional

organizations, and presumably of teachers, to diverse evaluation systems. During the current period of public debate, it is difficult to obtain an accurate picture of how individual teachers actually view this evaluation. Data collected through neutral surveys prior to the publicity surrounding *A Nation at Risk*³ would seem more likely to contain meaningful teacher opinions on this subject than would be the case with more current data. As a result, this report is based on teacher opinions concerning preferred methods of evaluation collected in 1980 from teachers with three to thirty years of experience. The data base, available from a large university alumni follow-up study of individuals who received teaching certificates from 1946 through 1976, is not necessarily representative of all teachers. Nevertheless, it has the following advantages: the inclusion of teachers from a wide variety of schools and school districts, freedom of teachers to express their views concerning a politically sensitive issue, restriction to teachers who have persisted in their careers past the heavy attrition period of the first three years, and the availability of sufficient demographic and attitudinal information to permit multivariate analysis of the data.

This report represents an exploratory analysis of potential predictors of teacher attitudes toward teacher evaluation. Specifically, the analysis pursues the following two questions:

- (1) Do personal and environmental factors, teacher perceptions of their own skills, and teacher attitudes toward their professional success predict the preferences of teachers for various types of evaluation?
- (2) Can patterns of teacher background, skill perceptions, and attitudinal variables be identified that are related to preferences for various types of teacher evaluation?

Identification of factors that influence teacher views of evaluation may be helpful in (1) understanding forces that influence teacher support of or opposition to various methods of teacher evaluation; (2) identifying segments of the teacher population that might be involved when consensus about evaluation is desired; and (3) suggesting ways to develop shared views.

CONCEPTUAL FRAMEWORK

Although the data for this analysis were collected prior to publication of the recent review of the evaluation literature by Darling-Hammond, Wise, and Pease,⁴ their work proved useful in planning the analysis of

the data. They described four major conceptualizations of teaching work: teaching as labor, teaching as a craft, teaching as a profession, and teaching as an art. These conceptualizations of work fall on a continuum of school organization from the more rationalistic (labor and craft) to the less rationalistic (profession and art). Additionally, for each conceptualization of teaching work, the above authors described a corresponding evaluation scheme. "Teaching as labor" is associated with a rationalistic, hierarchical model of evaluation; in this sense, the work of teaching is well-defined and supervisors are responsible for direct inspection of teachers at work. Next along the continuum is "teaching as a craft," in which less direct supervision of teachers' work is required and evaluation is intended to determine if teachers possess the requisite skills. In the conceptualization of "teaching as a profession," teaching is viewed as problem-solving work. In this sense, teachers take total responsibility for their work according to standards developed by peers. Thus, problem-solving teachers bring to the task a body of theoretical knowledge and a range of instructional strategies that they adjust to meet the observed needs of students. Evaluative judgments by those of similar professional competence is most appropriate under this view of teaching. The conceptualization of "teaching as an art" implies an even more personalized, creative activity; in this view, teachers are primarily responsible for self-assessment while also subject to the critical assessment of others. Darling-Hammond, Wise, and Pease caution that these four idealized views of teacher work probably do not exist as pure entities in the real world.

These authors also provided examples of two additional conceptualizations of teaching work that apparently fell outside their school organization and continuum context and its associated evaluation methods. The first was the conceptualization of "teaching as the production of a product" that can be quantified. Although the emphasis here is on a product rather than a process, such a conceptualization, with an obvious emphasis on pupil achievement and developmental growth as the "educational product," seems to lie near the "teaching as labor" end of the continuum. The other conceptualization of teaching work was "teaching as a client service." This implies a role for both student and parent evaluation of teachers. Because such service is characteristic of professional activity, it is possible that Darling-Hammond, Wise, and Pease conceived of "teaching as a client service" within their view of teaching as a profession. Similarly, they might have subsumed it under teaching as an art because artists' work is subjected to critical assessment. Although the idea of sensitivity to clients may have

different meanings within other conceptualizations of teaching work, the present authors believe that "teaching as a client service" should be explored as a separate conceptualization of teaching.

Based on these considerations, the Darling-Hammond, Wise, and Pease scheme was expanded and indices were created from the available survey items concerning teacher evaluation that related to six views of teaching: teaching as labor, teaching as a craft, teaching as a profession, teaching as an art, teaching as the production of a product, and teaching as a client service.

The six conceptualizations of teaching, the kinds of assessments implied and the survey items that comprised the scales for this analysis are summarized in Table 1. The single- or multiple-item scales became the dependent variables in six separate hierarchical multiple regression analyses. Recognizing the difficulty of matching items to ideal types, the researchers' reasoning in this regard should be explained.

The survey items chosen to represent "teaching as a profession" and "teaching as an art" were very close to the evaluation strategies described for these conceptions by Darling-Hammond, Wise, and Pease. In these two categories, a single-item scale was judged to be a direct measure of the ideas embodied in each evaluation or assessment approach that was labeled Teacher Peer Assessment and Teacher Self-Assessment, respectively.

One survey item concerning classroom observation by administrators/supervisors directly represented "teaching as labor" and one item regarding judgments of professional growth seemed akin to the "teaching as a craft" conceptualization in which craftspersons attempt to improve requisite skills. A third survey item, concerning administrator assessment of objectives established or negotiated with teachers in advance, posed a dilemma in this regard. It incorporated aspects of both "teaching as labor" and "teaching as a craft." To determine the correct assignment of this item, a factor analysis was conducted of the ten survey items answered by the teacher sample. Clearly, teachers had placed emphasis on the judgment by administrators or supervisors rather than on the concept of professional growth. In fact, teacher responses on this item and the item concerning classroom observation by administrators/supervisors exhibited strong loadings (.81 and .73) on the same factor. Thus, the items concerning superordinates' judgments of professional growth and classroom observations by administrators/supervisors were grouped into a two-item measure called Administrator Judgments to represent "teaching as labor." Only a single-item scale (Objectives Accomplished) was used to measure "teaching as a craft."

TABLE 1
Operationalization of Conceptualizations of Teaching
as Dependent Variables in the Study

<i>Conceptualizations of Teaching (Dependent Variable)</i>	<i>A Teacher's Evaluation* Should Be Based on:</i>
Teaching as Labor (Administrator Judgments)	Classroom observation by administrators/supervisors Professional growth of teachers as perceived by administrators/ supervisors
Teaching as the Production of a Product (Test Results)	Students' achievement gains on standardized tests Students' achievement gains on locally developed tests
Teaching as a Craft (Objectives Accomplished)	Accomplishment of objectives stated or negotiated in advance (e.g., growth contracts)
Teaching as a Client Service (Student/Parent Judgments)	Students' ratings of the teacher Parents' evaluation of the teacher Numbers of students who desire to enroll in teacher's classes
Teaching as a Profession (Teacher Peer Assessment)	Evaluations by other teachers who are familiar with the teacher's work and students
Teaching as an Art (Teacher Self-Assessment)	The teacher's self-assessment

*This prefatory statement was used in the survey: "The issue of teacher evaluation for administrative decision making in elementary and secondary schools has received a great deal of attention in the last few years. Please indicate the strength of your agreement with the following items using these codes: (1) strongly agree, (2) agree, (3) disagree, (4) strongly disagree."

On the basis of both face validity and the factor analysis of the survey items, a three-item measure termed Student/Parent Judgments was assigned to the conceptualization of "teaching as a client service." (These items had related factor loadings of .72, .63, and .65.) A two-item measure called Test Results was used to represent "teaching as the production of a product" (factor loadings of .78 and .79). All multi-item

measures were formed by summing the non-missing scores of respondents on the scaled items and dividing by the number of items.⁵

METHODOLOGY

Hierarchical multiple regressions were used to assess the existence of patterns of teacher preferences for various types of evaluation. The entry order for the set-wise regression procedure was based on decreasing probability that the predictor variable sets were fixed characteristics of teachers or the teaching environment. The additional contribution of each set of variables (with the relatively fixed variables in the equation) to the prediction of views concerning the various evaluation methods was, thus, examined. A final step-wise regression provided additional information about the relationship of specific teacher characteristics to evaluation preferences.

Sample

A subsample of respondents for this study was selected from a more comprehensive study of individuals who received teaching certificates from the University of Michigan from 1946 through 1976.⁶ Of the original 2933 respondents to the spring 1980 survey (51% of the random sample of certificate recipients surveyed), 1054 individuals were still employed in teaching at the time of the survey. Elimination of college teachers, school administrators, and counselors as well as individuals who described themselves as special kinds of teachers (e.g., bible school, day care center, or private music teachers) reduced the subsample to 893 respondents that were teaching either full or part time in kindergarten through grade twelve school settings. The rather extensive set of independent variables used in this analysis, accompanied by list-wise deletion because of missing data, reduced the sample for analysis to 536 subjects. In Table 2, some key characteristics of the analysis sample are compared with the same characteristics for the larger sample of 893 employed teachers. As a result, it can be seen that the reduced sample slightly over-represents younger and male teachers in suburban public schools, but otherwise, it differs minimally from the larger sample that might have been used without the necessary deletion of respondents with missing data.

Although the sample was not intended to be representative of all teachers, its characteristics are similar to other presumably representative samples drawn at about the same time. For example, the National Education Association reported that the mean United States teaching salary in 1981 was \$17,209, the median number of years of teaching

TABLE 2
 Characteristics of Teachers in the Larger Respondent Sample
 and the Analysis Sample

	<i>Sample</i> (<i>N</i> = 893)	<i>Analysis</i> <i>Sample</i> (<i>N</i> = 536)
Sex (male)	23.8%	26.1%
Mean age	40.9	39.6
Year teaching began		
1945-1955	24%	22%
1956-1965	31%	29%
1966-1976	45%	49%
Teaching regularly since graduation	69%	72%
Level of teaching		
Elementary	44%	44%
Junior High/Middle	18%	17%
Secondary	31%	32%
Districtwide	7%	7%
1980 mean salary	\$17,730	\$18,000
Teach in public school	89%	91%
Area in which school is located:		
Urban	27%	25%
Suburban	57%	59%
Rural	16%	16%
Plan to continue teaching	82%	82%
Prefer current job	69%	71%

tenure was 12 years, the median age of teachers was 37 years, and the percentage of women teachers was 66.7%⁷

Independent Variables

Measures available from the survey data base that might predict teacher attitudes toward evaluation were classified into the following six groups: Personal Variables, Environmental Variables, Skill Perceptions, Job Status Perceptions, Important Criteria for Professional Success, and Perceived Achievement of Professional Success. Categorical variables that could be ordered along a meaningful continuum were treated as interval data for the analysis; variables that were clearly nominal were dummy coded. Several extensive sets of conceptually related survey items (assessment of skills possessed—16 items; profes-

sional criteria considered important—11 items; and professional criteria considered achieved—11 items) were condensed into more parsimonious and orthogonal sets through use of factor scores. Based on previous research,⁸ one scale measuring a locked-in perception was constructed from conceptually related items without factor analysis. Each variable set is described briefly below, and descriptive statistics for each of the variables not adequately described in Table 2 are reported in Table 3. The intercorrelations among the independent variables are reported in Table 4.

Personal variables. The personal variables analyzed were the following:

- (1) Age,⁹
- (2) Sex (dummy coded: male = 1; female = 0),
- (3) Recollection of initial commitment to teaching (four point Likert-type scale: 1 = high commitment to 4 = low commitment), and
- (4) Extent to which first position is recalled as positive experience (four point Likert-type scale: 1 = positive to 4 = not at all positive).

Environmental variables. The environmental variables analyzed were the following:

- (1) School location (dummy coded as two variables: urban, suburban; reference category; rural),
- (2) School control (dummy coded: public = 1; private = 0),
- (3) Teaching level (dummy coded as three variables: elementary, junior high, and high school; reference category district-wide), and
- (4) Gross annual income index had 14 intervals (1 = less than \$3000 to 14 = \$60,000 and over).

Skill perceptions. The measures of skill perceptions used were factor scores created by varimax rotated principal components analysis of sixteen skill assessment items answered on a four-point Likert-type scale (1 = possess to a large extent to 4 = possess not at all). Based on loadings greater than .45, factors were named by considering the item loadings shown below. (A low factor score indicates a perception of strong skills in the skill area named.¹⁰) The variables used were the following:

- (1) Skill in working with others (supervising and leading, .45; cooperating with a work team, .57; persuading others to accept ideas, .52; dealing with the public, .57; resolving conflicts in work settings, .56; and communicating with others, .63);

TABLE 3
Descriptive Statistics for Independent Variables (N = 536)

Variable Set	Mean	SD	Skewness	Kurtosis	Yes	SA	A	D	SD
						(Percentages)			
Personal Variables									
Age	39.61	9.75	.24	-1.08					
Sex (male)					26.1				
Initial commitment to teaching ^a	1.85	.90	.66	-.65		44.6	30.2	20.9	4.3
Extent first position was positive ^b	2.04	.86	.34	-.75		30.8	39.2	25.7	4.3
Environmental Variables (See Table 2)									
Skill Perceptions^c									
Working with others	2.08	.77	.53	-.18					
Working with data	3.04	.82	-.88	.39					
Planning and organizing	1.58	.74	.54	-.07					
Communicating	2.02	.78	.60	1.00					
Problem solving	1.47	.66	.30	-.28					
Job Status Perceptions									
Locked-in perception ^d	1.74	.30	-.99	.09					
Sufficient status					59.3				
Opportunity to advance					23.9				
Skills well-utilized					69.4				
Plan to continue teaching					82.0				
Prefer to remain in current job					71.3				
Important Criteria for Professional Success^c									
Leadership/responsibility	2.64	.85	.06	-.54					
Recognition by others	2.47	.89	.15	-.42					
Self-direction	1.87	.66	.81	.36					
Perceived Achievement of Professional Success^c									
Leadership/responsibility	2.32	.83	-.01	-.12					
Recognition by others	2.59	.84	.05	-.18					
Unnamed criteria factor	2.82	.62	-.60	.46					

a. SA to SD are stand-ins for responses: extensive commitment (SA) to no commitment (SD).

b. SA to SD are stand-ins for responses: extremely positive (SA) to not at all positive (SD).

c. Means are for unstandardized factor scores; low scores indicate perception of strong skill possession, high importance of professional criteria, and professional success extensively achieved.

d. Continuous scale range for locked-in perception is from 1.00 (low lock-in) to 2.00 (high lock-in).

TABLE 4
Zero-Order Correlations (with decimal points omitted) among Independent Variables (N = 536)

Independent Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
Personal Variables																											
1 Age	-																										
2 Sex (male)	06	-																									
3 Initial commitment	07	-04	-																								
4 First position positive	-03	-03	27*	-																							
Environmental Variables																											
5 Urban school	07	-02	09*	14*	-																						
6 Suburban school	01	-00	-08	-16*	-																						
7 Public school	12*	07	-05	-02	-00	03	-																				
8 Elementary teacher	03	-33*	-02	-08	03	-03	00	-																			
9 Jr. high teacher	-01	12	01	03	03	-02	06	-																			
10 High school teacher	-05	24*	-02	03	-08	09*	03	-																			
11 Annual income	36*	32*	-07	-03	-02	17*	31*	-09*	04	11*	-																
Skill Perceptions																											
12 Working with others	10*	13*	09*	12*	-02	-02	01	02	-08	08	04	-															
13 Working with data	05	-07	02	00	-07	07	-12*	09*	-01	-13*	-12*	01	-														
14 Planning/organizing	-04	14*	04	01	-01	-03	-02	00	00	-05	09*	01	-														
15 Communicating	-00	07	-01	01	-01	04	08*	10*	08*	10*	-02	23*	04	13*	-												
16 Problem solving	09*	-01	-01	01	04	-05	02	05	-03	-02	08	20*	06	10*	23*	-											
Job Status Perceptions																											
17 Locked-in (yes)	16*	-08	-07	-07	-04	11*	03	14*	-08	-09	18*	-00	06	-00	02	09*	-										
18 Sufficient status	08	05	03	-05	02	-01	01	05	-01	-04	07	-00	-10*	-01	10*	06	08	-									
19 Chance to advance	-08	05	06	00	09*	-08	05	03	07	-05	-01	-12*	-16*	-07	01	-03	-11*	17*	-								
20 Skills well-used	02	-09*	-03	-09*	00	01	-01	09*	-05	-06	-04	-04	01	-02	04	08	07	46*	13*	-							
21 Hope to remain	23*	-04	-04	-12*	-02	-01	02	10*	-02	-05	08	06	-01	01	03	08	17*	41*	-01	40*	-						
Import. of Prof. Criteria																											
22 Leadership/resp.	06	01	07	02	-05	-00	-06	09*	-07	-03	-14*	19*	00	12*	15*	17*	03	09*	-11*	13*	-20*	-					
23 Recognition-others	-04	06	05	09*	05	-07	02	-06	06	02	01	10*	-07	06	01	01	-06	-08	-01	-07	-10*	03	-				
24 Self-direction	16*	19*	05	07	-01	-02	04	-02	04	00	09*	16*	03	14*	08	14*	-03	01	-09*	-04	-00	22*	16*	-			
Profess. Criteria Achieved																											
25 Leadership/resp.	04	01	09*	19*	-02	-06	06	06	04	-08	04	-08	19*	-03	13*	00	14*	04	-22*	-21*	-23*	-12*	36*	07	14*	-	
26 Recognition-others	00	10*	13*	16*	01	-09*	-00	-05	05	-01	-04	30*	09*	14*	14*	18*	-03	-08*	-13*	-13*	-13*	-13*	07*	45*	32*	09*	-
27 Unnamed factor	-12*	-05	07	09*	-01	-07	-08	01	-06	05	-26*	13*	00	18*	11*	13*	02	-09*	-16*	-06	-06	28*	05	22*	26*	15*	

*p < .05; r_p < .05 = .0847.

- (2) Skill in working with data (interpreting numerical data, .67; and using computers and computer printouts, .66);
- (3) Skill in planning and organizing (organizing time effectively, .55; planning and organizing job activities, .66; and working on long-term projects, .54);
- (4) Skill in communicating (writing effectively, .61; and speaking effectively, .49); and
- (5) Skill in problem solving (developing new approaches to problems, .61; and analyzing and evaluating ideas, .58).

Job status perceptions. The job status perception variables employed in the analysis were the following:

- (1) "Locked-in" perception: A continuous scale developed from three items indicating low horizontal job mobility (1 = low perception of being locked in; 2 = high perception of being locked in);
- (2) Sufficient status: Current employment offers sufficient status (dummy coded: yes = 1; no = 0);
- (3) Opportunity to advance: Current employment offers sufficient possibilities for advancement (dummy coded: yes = 1; no = 0);
- (4) Skill use: Skills are well-utilized in current employment (dummy coded: yes = 1; no = 0);
- (5) Future plans: Plan to continue in teaching (dummy coded: yes = 1; no = 0); and
- (6) Job preference: Prefer to remain in current position (dummy coded: yes = 1; no = 0).

Important criteria for professional success and perceived achievement of professional success criteria. The three variables measuring importance of professional success were unstandardized factor scores created from a varimax rotation of factors derived from a set of eleven items asking respondents to indicate on a four-point Likert-type scale the importance they attached to the given criterion as an indicator of professional success (1 = very important to 4 = not at all important). The three factors were named according to the items that exhibited loadings greater than .40.¹¹ (A low score indicated that importance was attached to the factor.) These variables were as indicated below:

- (1) Importance of leadership/responsibility (chance to contribute to important decisions, .66; leadership activities in professional field, .78; and increased job responsibility, .68);

- (2) Importance of recognition by others (recognition by students, .60; recognition by peers, .80; recognition by superiors, .73; and approval from family/friends, .51); and
- (3) Importance of self-direction (an inner sense of knowing you are doing work well, .53; and opportunity to learn, .45).

The three measures of perceived achievement of professional success were factor scores parallel to those described above and were derived in the same manner. In this instance the respondents were asked to rate the eleven potential indicators of professional success on a scale from 1 = have achieved extensively to 4 = have not achieved at all. The first two factors, labeled Achievement of Leadership/Responsibility and Achievement of Recognition by Others, exhibited loadings of slightly less magnitude for the same items described above regarding important criteria of professional success. The third factor, accounting for only seven percent of the variance, was difficult to interpret. It included modest loadings (from .35 to .38) on such diverse professional criteria as leadership activities in a professional field, inner sense of knowing you are doing your work well, recognition by students, opportunity to learn, and publication in journals, while sharing variance with each of the other factors ($r = .28$ with Leadership/Responsibility and $.18$ with Recognition by Others for the sample on which the factor analysis was based). The third factor was retained in the analysis but was left unnamed.¹²

Dependent Variables

The six dependent variables were scores on single items or summated measures constructed from a series of ten items concerning how teachers should be evaluated. The rationale behind the choice of these single- or multiple-item scales, based on conceptualizations adapted from Darling-Hammond, Wise, and Pease¹³ has already been discussed. Descriptive statistics for the scales are reported in Table 5. Respondents were not asked to choose a single evaluation method as preferable but rated the potential use of each method on a four-point scale from "strongly agree" to "strongly disagree." Scale intercorrelations are reported in Table 6, and the correlations between the dependent and independent variables are reported in Table 7.

Statistical Analysis

In order to assess multicollinearity of the predictor variables prior to the multiple regression analysis, each predictor variable was regressed

TABLE 5
Descriptive Statistics for the Dependent Variables (N = 536)

Evaluation Scales	Mean ^a	SD	Skewness	Kurtosis	SA	A	D	SD
					(Percentages)			
Scale 1 Administrator Judgments (Items [A + B]/2)	2.08	.60	.50	1.28	N/A ^b			
Item A: Class observations by administrators	2.02	.63	.65	1.54	16.2	68.5	12.5	2.8
Item B: Professional growth as judged by administrators	2.14	.69	.63	.82	13.4	64.0	18.1	4.5
Scale 2 Test Results (Items [C + D]/2)	2.84	.75	.07	-.71	N/A			
Item C: Student achievement gains on standardized tests	3.02	.78	-.32	-.59	2.2	22.6	45.9	29.3
Item D: Student achievement gains on local tests	2.66	.85	.13	-.82	6.2	41.0	33.6	19.2
Scale 3 Objectives Accomplished	2.12	.71	.45	.32	16.0	59.3	20.9	3.7
Scale 4 Student/Parent Judgments (Items [E + F + G]/2)	2.95	.64	-.33	-.05	N/A			
Item E: Student ratings of teachers	2.78	.82	.01	-.77	3.9	34.9	40.5	20.7
Item F: Parents' evaluation of teachers	2.96	.75	-.22	-.51	2.1	24.3	49.6	24.1
Item G: Students' desire to enroll in classes	3.11	.81	-.57	-.32	3.2	17.9	43.5	35.4
Scale 5 Peer Assessment	2.11	.78	.46	-.02	20.1	53.5	21.0	5.2
Scale 6 Self-assessment	1.84	.65	.54	.85	28.5	60.6	9.1	1.7

a. A mean score signifies agreement that the evaluation method is viewed as appropriate.

b. Because multiple-item measure construction resulted in interval variables, percentage responses could not be computed.

on all other predictor variables. The largest R^2 obtained from these regressions was less than .40; thus, multicollinearity was not judged to be a problem.¹⁴ Additionally, although a few independent variables were skewed, no substantial deviations from the assumptions of homoskedasticity or linearity were noted in post-hoc scatter plots relating predicted values to full-model regression results.

Each of the dependent variables was regressed on the predictor variables in a set-wise procedure with all variables in each conceptually related set entered together. The variable sets were entered in the following order, based on the decreasing probability that the predictor variables were fixed characteristics of teachers or work settings: (1) self-reported personal characteristics of teachers; (2) self-reported character-

TABLE 6
Zero-Order Correlations (with decimal points omitted) among Dependent Variables (N = 536)

	<i>Administrator Judgments</i>	<i>Test Results</i>	<i>Objectives Accomplished</i>	<i>Student/Parent Judgments</i>	<i>Peer Assessment</i>	<i>Self-Assessment</i>
Administrator Judgments	—					
Test Results	10	—				
Objectives Accomplished	21	15	—			
Student/Parent Judgments	08*	36	27	—		
Peer Assessment	16	11	16	20	—	
Self-Assessment	18	-06*	25	08*	26	—

*Only correlations that are not significant at the .05 level.

TABLE 7
Zero-Order Correlations (with decimal points omitted) between Dependent and Independent Variables (N = 536)

<i>Independent Variables</i>	<i>Administrator Judgments</i>	<i>Test Results</i>	<i>Dependent Variables</i>			
			<i>Objectives Accomplished</i>	<i>Student/Parent Judgments</i>	<i>Peer Assessment</i>	<i>Self-Assessment</i>
<i>Personal Variables</i>						
Age	-01	-10*	04	-09*	04	09*
Sex (male)	14*	-07	08	02	-03	10*
Initial commitment	-00	-03	-05	00	01	10*
First position positive	12*	04	-03	05	06	06
<i>Environmental Variables</i>						
School location (urban)	02	-09*	-07	00	-02	06
School location (suburban)	-03	07	02	-06	-01	-04
School control (public)	07	06	04	06	04	00
Teaching level (elementary)	-04	04	00	01	13*	-03
Teaching level (middle)	-03	02	-03	02	-09*	-02
Teaching level (high school)	09*	-05	04	00	-06	06
Gross annual income	10*	08	05	03	06	-01
<i>Skill Perceptions</i>						
Working with others	12*	-00	11*	08	11*	07
Working with data	01	04	-06	-00	03	-01

Planning and organizing	11*	05	00	-01	-06	03
Communicating	-02	08*	10*	08	06	07
Problem solving	-00	-01	11*	09*	03	18*
Job Status Perceptions						
Locked-in perception	-00	02	12*	03	00	02
Sufficient status (yes)	-09*	-05	-01	-04	-04	00
Chance to advance (yes)	-13*	-10*	-15*	-05	-01	-16*
Skills well-utilized (yes)	-08	-03	-05	03	10*	-03
Prefer to remain in current job	-09*	-00	01	-01	01	-00
Importance of Professional Criteria						
Leadership/responsibility	02	-01	10*	-05	06	-00
Recognition by others	18*	07	04	11*	09*	08
Self-direction	02	-01	01	08	03	10*
Professional Criteria Achieved						
Leadership/responsibility	19*	01	10*	-03	02	15*
Recognition by others	16*	07	03	12*	-01	09*
Unnamed factor	-02	06	12*	15*	01	11*

NOTE: A low score on a dependent variable indicates a favorable response to the evaluation method.

* $p < .05$; $r < .05 = .0847$.

istics of the teachers' work settings; (3) teachers' perceptions of their own skills; (4) teachers' attitudes about their job status; (5) criteria teachers chose as important indicators of professional success; and (6) teachers' self-reports about their achievement of the same success indicators. The entry order of the variable sets that were entered after personal and environmental characteristics could differ under other assumptions. The model used assumes (1) that personal characteristics of teachers and environmental variables are either not subject to change or that change is most unlikely and (2) that skills teachers bring to the job are more fixed than are the variables related to job status and professional attitudes.

Summary results of the set-wise hierarchical regression are reported in Table 8. The percentage of variance explained is reported for each set alone, for the increment due to the addition of each set, and for the equation after the addition of each set to the preceding variable sets.

In a second regression procedure, personal and environmental variables were entered first as a fixed set, followed by all other variables entered in step-wise fashion with the entry criterion set at $p < .05$. Because fewer variables entered the regression equation using this criterion, less variance was explained, but the significant contributing variables were identifiable. The results of this combined set-wise procedure are presented in Table 9. (Those predictor variables that did not contribute significantly to any of the six regression equations are not included in Table 9.)

RESULTS

Results of the above analyses are reported in four sections. First, some characteristics of teachers in the sample based on the independent variables and their correlations are described. Next, teacher views of evaluation based on the dependent variables are reported. In the third section, based on the hierarchical multiple regression model, the extent to which it is possible to predict teacher views of evaluation from the characteristics and attitudes they reported in the survey is indicated. Lastly, the patterns detected in the step-wise regression analysis that indicate the types of covariates that may predict teacher endorsement of specific methods of evaluation are examined in more detail.

Information from the Independent Variables

The distributions of the independent variables (from Table 3) and the correlation patterns among the personal and environmental variables (from Table 4) present few surprises for a sample of teachers. For

TABLE 8
Percentage Variance and Increase in Percentage Variance Contributed by
Each Set of Predictor Variables for Each Dependent Variable

Independent Variable Set Added	Dependent Variables																	
	Administrator Judgments			Test Results			Objectives Accomplished			Student/Parent Judgments			Peer Assessment			Self-Assessment		
	(1) ^a	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Personal Variables (4,531; 4,531) ^b	3.57*	3.57*	3.57*	1.70	1.70	1.70	.91	.91	.91	1.13	1.13	1.13	.65	.65	.65	2.70*	2.70*	2.70*
Environmental Variables (7,528; 11,524)	2.05	1.22	4.79*	2.19	3.82*	5.52*	1.20	.90	1.81	1.42	1.59	2.72	2.73*	2.65*	3.30	.86	1.20	3.90*
Skill Perceptions (5,530; 16,519)	2.84*	2.08*	6.87*	1.11	1.36	6.88*	2.91*	2.59*	4.40	1.52	1.70	4.42	2.03	1.52	4.82	3.32*	2.67*	6.57*
Job Status Perceptions (5,530; 21,514)	2.64*	1.98*	8.85*	1.10	1.07	7.95*	3.41*	5.02*	7.61*	.70	.90	5.32	1.78	1.99	6.81*	2.72*	2.86*	9.43*
Importance of Professional Criteria (3,532; 24,511)	3.16*	2.46*	11.31*	.55	.42	8.37*	1.28	.85	8.46*	2.16*	2.04*	7.36*	1.11	.83	7.64*	2.35*	.75	10.18*
Professional Criteria Achieved (3,532; 27,508)	6.58*	2.12*	13.43*	.79	.57	8.94*	1.95*	.83	9.29*	3.80*	2.37*	9.73*	.04	.64	8.28*	3.22*	.46	10.64*

*p < .05.

a. (1) variance explained by set alone; (2) increment in variance due to addition of set; and (3) variance explained after addition of set to preceding variable sets.

b. First pair of degrees of freedom in parentheses is for set alone; second pair of degrees of freedom is for addition of set to preceding variable sets.

TABLE 9
 Statistically Significant Predictors of Evaluation Preferences with Initial Set-Wise Entrance
 of Personal and Environmental Variables Followed by Step-Wise Entrance of Other Variables

<i>Independent Variables</i>	<i>Beta Weights of Significant Variables</i>				
	<i>Administrator Judgments</i>	<i>Test Results</i>	<i>Objectives Accomplished</i>	<i>Student/Parent Judgments</i>	<i>Peer Assessment</i>
Personal and Environmental					
Age		-.17**		-.09*	
Sex	.11*	-.12*			.12**
Initial commitment to teaching					.09*
First position positive	.09*				
Gross annual income		.19**			
Skill Perceptions					
Working with others					.10*
Working with data			-.09*		
Communicating		.09*			
Problem solving			.09*		.16**
Job Status Perceptions					
Locked-in perception			.10*		
Sufficient status					-.10*
Chance to advance	-.10*	-.11*	-.14**		-.17**
Skills well-utilized					.10*

Importance of Professional Criteria									
Leadership/responsibility									
Recognition by others	.16**								-.09*
Professional Criteria Achieved									
Leadership/responsibility	.17**								.09*
Recognition by others									.16**
Unnamed factor		.12**							
Percentage of variance explained by personal and environmental variables	4.8	5.5	1.8	2.7	3.3	3.9			
Overall percentage of variance explained	10.8	7.3	7.4	6.7	6.1	8.9			
Overall F ratio	4.52**	3.16**	2.59**	2.66**	2.40**	3.94**			
Degrees of freedom	14,521	13,522	16,519	14,521	14,521	13,522			

*p < .05; **p < .01.

example, statistically significant correlations ($p < .05$) among the variables revealed that elementary teachers were more likely than junior high school or high school teachers to be women. Higher gross annual income was associated with being older, male, and with teaching in a suburban public school above the elementary level. Suburban teachers more often had positive recollections of their first position than did urban teachers.

Judging from the relatively high correlations between the professional criteria teachers consider important and the criteria they feel they have achieved (ranging from .22 to .45), teachers may well be considered to feel relatively positive about their accomplishments. As a group, the teachers tended to view themselves as strong in the skills of planning, problem solving, communicating, and working with others but quite weak in working with data. At least in teachers' self-assessments, skill in working with data appeared distinct from the other skills, although it may be characteristic of some high school teachers at high income levels. Teachers who believed they were skilled in working with data also tended to believe that teaching had sufficient status and saw opportunities for advancement in their current jobs.

However, only a small portion of the sample of teachers (23.9%) saw chances for advancement in their current positions. The teachers in this sample also expressed a strong feeling of being locked-in to their current jobs. Correlation patterns among the measures of perceived job status and the perceptions of professional success criteria achieved prompt the notion that there are at least two possible groups of teachers—the "complacent" and the "ambitious."

To illustrate, the correlational patterns indicated that teachers who felt most locked-in to their current jobs tended to be older, suburban, elementary school teachers at relatively high salary levels who saw little chance to advance in their positions. They were likely to believe that teaching had sufficient status, to think that their skills were well-utilized, and to prefer to remain in their current jobs. Based on the correlations, the feeling of being locked-in among these "complacent" teachers seemed unrelated to views regarding the importance of various professional success criteria and to feelings of having achieved these criteria. In contrast, teachers who believed that they had achieved various professional success criteria were also likely to believe that they had strong skills. They valued leadership and responsibility as important in professional success and tended to report that there were no opportunities to advance. Further, they saw their skills as not well-utilized. Teachers possessing this set of views might be termed "ambitious."

Clearly, the large number of independent variables used in this study and the complexity of the correlational patterns among them makes descriptions based on the bivariate correlation coefficients quite speculative. A multivariate analysis based on this set of teacher characteristics and attitudes should be more useful because, in such an analysis, the shared variance among the independent variables can be taken into account.

Information from the Dependent Variables

Self-Assessment clearly was viewed by the teacher subjects as the most appropriate method of evaluation (Table 5); about 89% of the teachers either agreed or strongly agreed that teachers should assess their own work. Surprisingly, however, Administrator Judgments were also viewed by teachers as appropriate. About 85% of the teachers agreed or strongly agreed that classroom observations by administrators should be used, and 77% were accepting of administrator judgments regarding personal growth.¹⁵ Close behind in order of acceptance were Teacher Peer Assessment and assessment of Objectives Accomplished; 73-75% of the teachers accepted these modes of evaluation. Examination of the zero-order correlations in Table 6 indicated that teachers view Administrator Judgments, the assessment of Objectives Accomplished, Teacher Peer Assessment, and Self-Assessment as related methods of evaluation.

Teachers viewed both the assessment of Test Results and Student/Parent Judgments in teacher evaluation negatively. From 52% to 79% of the teachers felt that the use of these modes of evaluation was not appropriate. Within the sets of items comprising these measures, locally constructed tests were viewed somewhat more favorably than standardized tests and student evaluations were viewed more favorably than parent evaluations.

Judging from the correlations in Table 6, there were hints of two distinct categories of evaluation methods—one that could be viewed as internal to the profession of teaching (peer, self, and administrator judgments) and the other that is based on external referents (students, parents, and test scores).

The Prediction Model

Based on the independent variables that were available in the data set, the prediction of teacher views concerning specific ways of evaluating their work was statistically significant ($p < .05$), but modest in percentage terms. In the multiple regression analysis, the percentage of var-

iance for each dependent variable that was explained by the predictor variables ranged from 8.28% to 13.43% (see Table 8). The best predictions were obtained with respect to teacher views on Administrator Judgments (13.43%) and Teacher Self-Assessment (10.64%). The strength of these predictions must be considered in light of (1) the restricted diversity of views regarding these two types of evaluation (in each case, over seventy percent of the teacher sample answered "agree") and (2) the substantial correlation among the evaluation methods that makes clear-cut prediction equations difficult to derive.

Patterns of Prediction

Although relatively weak predictions were possible, the findings indicated that teacher background and professional attitudes were related to views about evaluation. Also, it was possible to identify predictors of teacher views toward specific types of evaluation that are amenable to further study.

The bivariate correlations of the teacher characteristics and attitudes with reports of teacher views on the six evaluation methods were reported in Table 7. (Note: In interpreting the correlation coefficients in Tables 7 and 9, the reader should keep in mind that a low score on each evaluation method indicates a favorable view of that method.) Examination of the correlations allows one to gain a sense of those teacher characteristics and attitudes that are related to teachers' evaluation preferences. Because of slight collinearity of the predictor variables, not all of the predictors that were significant in the bivariate analysis emerged as predictors when other variables were held constant. The results of the set-wise hierarchical multiple regression (see Table 8) and the step-wise regression in which personal and environmental variables entered first were held constant are summarized below (see Table 9).

In predicting teacher views on the use of Administrator Judgments for evaluation, each set of independent variables, except those representing environmental conditions, contributed new statistically significant information when entered in the set-wise regression procedure. Information about important environmental variables appears to have been incorporated by the personal variables with which they shared variance. Teachers favoring administrator judgments were most likely to be female (and, thus, more likely to be teaching at lower school levels and have lower incomes) and to have had a positive experience in their first teaching position. These teachers were characterized by a positive view of their advancement opportunities, and they viewed recognition by others as important and felt that they had achieved leadership and responsibility.

Fixed personal characteristics were also important in predicting teacher views on Self-Assessment as an evaluation method. Teachers favoring self-assessment were likely to be female and to recall a strong initial commitment to teaching. With personal and environmental variables accounted for, perceived skills in problem solving and positive views of opportunities to advance were additional contributors to the equation. Self-Assessment was distinguished from Administrator Judgment (with which it was correlated in the bivariate analysis of dependent variables) by the lack of additional predictive power of professional attitudes regarding success once skill perceptions and job status variables were controlled.

Although the overall regression equation on the Teacher Peer Assessment measure achieved statistical significance, only the set of environmental variables contributed significantly to its prediction. Perceived skill in working with others, a belief that their current positions provided status, and a sense that their skills were under-utilized seemed to characterize teachers who preferred Peer Assessment. Although specific environmental variables apparently were masked by shared variance with other variables in the step-wise regression, it appeared from the bivariate correlations that such characteristics associated with teachers who favor peer assessment may be more frequently found among teachers beyond the elementary school level.

Teacher views about assessment on the basis of Objectives Accomplished were predicted neither by personal characteristics nor by characteristics of the teaching environment. As was the case for Teacher Self-Assessment, important predictors associated with views of Objectives Accomplished were teachers' perceptions of the skills they possessed (in this case problem-solving skills) and a positive view of their job status (not feeling locked-in and believing there were opportunities to advance). Unlike those most strongly favoring Self-Assessment, however, teachers favoring negotiation of objectives to be accomplished believed that they had already achieved criteria of professional success. Their expression of such achievement had enough unique variance to enter the prediction equation after all other prime variables had been included.

In predicting views of teachers regarding the appropriateness of using Test Results as evaluation measures, personal and environmental characteristics were most important. Being an older, male, high income teacher in a non-urban setting seemed to predict a favorable view of use of student test scores. Only modest additional contributions to the regression equation for Test Results were made by skill perceptions (communicating with others) and by job status (opportunity to advance).

Views on the use of Test Results were not predicted by attitudes toward professional success.

In contrast to the situation for Test Results, the prediction of views on use of Student/Parent Judgments as an evaluation method was clearly related to teachers' attitudes toward professional success and was minimally dependent on personal and environmental characteristics or the skill perceptions of teachers. Professional attitude variables predicting views on Student/Parent Judgments exhibited a different pattern than for other evaluation methods. Specifically, teachers favoring this assessment method appeared to attach less importance to leadership and responsibility as a criterion of professional success and more importance to recognition by others.

One of the interesting results of this analysis is the identification of teacher characteristics and attitudes that do not appear to be associated with preferences for different types of evaluation. The type of school (public or private), the type of community in which the school is located, and teaching level seem minimally related to evaluation preference once collinear variables such as age, sex, and income level have been considered. Similarly, teachers' views regarding evaluation methods seemed to have little direct relationship to whether they preferred to remain in their current jobs. It may be, however, that the differences between those teachers who are resigned to their current situations and those who optimistically seek advancement opportunities were captured in the analysis by the response regarding advancement opportunities. In fact, although optimism about advancement opportunities characterized less than one-quarter of the teachers in the analysis sample, it was a predictor of views regarding four of the six evaluation methods.

Some teacher characteristics seemed unrelated to views concerning specific evaluation methods. The analysis of data indicated that teacher views regarding the use of peer assessment, accomplishment of objectives negotiated in advance, or use of student/parent judgments were minimally related to teacher personal characteristics or to the demographics of the teaching environment. In contrast, teacher views regarding administrator judgments, self-assessment, and use of test results as evaluation methods seemed more strongly related to personal and environmental factors. In the case of test results, personal and environmental characteristics were, in fact, the primary factors that seemed to distinguish those few teachers who favored the use of tests from the vast majority of teachers who opposed them. In a similar fashion, professional attitudes toward success were predictive primarily for use of administrator judgments (which most teachers accepted) and for the idea of student/parent judgments (which most teachers rejected).

DISCUSSION

The above findings indicate that views about evaluation may, indeed, be based on different conceptualizations of teaching, but the interrelationships prompt the researchers to reinforce the caution of Darling-Hammond, Wise, and Pease¹⁶ that distinct, ideal conceptualizations may not exist in the real world. Judging from teachers' acceptance of evaluation methods, they seemed to endorse both the "teaching as labor" and the "teaching as an art" views of their work. In the abstract, at least, they accepted evaluation methods that were primarily executed by supervisors as well as critical assessment by themselves and colleagues. When personal, environmental, skill, and job status perceptions are taken into account, however, differences appeared between these two views that may have organizational implications. Those holding the "teaching as labor" viewpoint valued leadership roles and had gained such roles as well as recognition from others. Administrator judgments represent one way to obtain such recognition. Recognition from others was also important for those favoring judgments by students and parents, but these teachers seemed not to care about leadership roles. Possibly, the concept of recognition by others as a criterion for professional success may be based on different reference groups (supervisors, students, and peers) for those favoring the different methods. The composite factor used as an independent variable in the above analysis encompassed all these reference groups in one measure.

Nevertheless, for teachers most strongly endorsing self-assessment (the "teacher as an artist" conceptualization), peer assessment (the "teacher as a professional" conceptualization), and test results (the "teacher as the producer of a product" conceptualization), the extrinsic rewards characterized by either recognition from others or by leadership roles seemed less important.

When personal and environmental factors were held constant, teacher self-perceptions of their job status appeared to influence their views regarding the use of administrator judgments, self-assessment, and accomplishment of objectives stated in advance, but seemed to have little relationship to the idea of using student/parent judgments or test results. The patterns seemed to indicate that teachers (particularly women) who continue to see possibilities for advancement in their jobs may favor joint participation with administrators in multiple modes of self and administrator evaluation as a road to advancement and job satisfaction.

Emphasizing both intrinsic and extrinsic rewards may be an important factor in designing teacher evaluation systems. Although Darling-

Hammond, Wise, and Pease viewed the conceptualizations of teaching as labor and teaching as art separately, indicating that they may inhabit opposite ends of an implied organizational continuum vis-à-vis evaluation methods, it is possible that both models can comfortably coexist for teachers. Teachers most strongly supporting each view may be satisfied if an evaluation system presents opportunities for recognition by others, for job advancement, and for sufficient feedback about performance to bolster a feeling of worth.

Teachers were quite favorable toward the idea of being assessed on their success in achieving objectives negotiated in advance. Although it is still uncertain about the extent to which this evaluation method encompasses the idea of “teaching as a craft,” these results indicate that a positive sense about advancement opportunities and strong positive perceptions of their own problem-solving skills characterize teachers who favor such procedures. In this sense, they appear similar to teachers who view self-assessment positively.

Peer assessment clearly epitomizes the conceptualization of teaching work as professional; teachers in the analysis sample generally supported this evaluation method, and, perhaps as a consequence of their unanimity, few predictors in this regard were identified. It can only be concluded that such teachers believed they possessed strong skills in working with others and felt that their skills were not fully utilized in their current positions. The concept of teaching as a profession clearly exists, but it may not be fully capitalized on in the current school organization.

The general acceptance by the teacher sample of four different evaluation methods implies that evaluation systems that incorporate more than one of these conceptions may be successful. Thus, the ideal system should incorporate elements from each or provide options for teacher choice of the method(s) to be used. Based on its relative independence from personal characteristics, the findings in this respect indicate that peer evaluation may be one basis on which to develop mutually satisfying systems. Indeed, the notion of teaching work as a profession implies standards developed by practitioners, but it need not exclude aspects of other conceptualizations. If peer assessment is not used as a rallying point, small coalitions of teachers drawn together by common characteristics and job attitudes may support widely different evaluation methods.

The two conceptualizations of teaching work that were added to the analysis—“teaching as the production of a product” and “teaching as a client service”—seem unlikely bases on which to begin to achieve consensus on teacher evaluation. The view of teaching work as “production

of a product," at least as operationalized in this analysis (use of test results), was not accepted by many teachers. The minimal relationship of this evaluation mode to professional success criteria may well reveal that teachers reject teaching as product production. Such a conclusion has been supported by other research. For example, the National Education Association reported in 1979 that only 11% of the teachers agreed or strongly agreed to the use of standardized tests in their evaluation.¹⁷ In the analysis sample, those who held this view independent of other professional attitudes were relatively well paid, older, male teachers who saw themselves as good communicators. Perhaps they felt that student demonstration of subject matter competence was the most important educational goal.

Although client relationships are considered important in most professional occupations, teachers in the sample were unwilling to accept measures of client input as a basis for evaluation. The regression results provided one insight, namely, that teachers who did value student and parent opinion as a measure of their own success were older, had already received recognition from others, and were less interested in opportunities for leadership and responsibility than colleagues. Perhaps the teachers accepting client input are more content with, or sensitive to, the importance of their helping relationship with students. To illustrate, Harootunian and Yarger¹⁸ found that teachers often conceive of their own success in terms of student successes rather than in terms of the effects upon themselves or their careers. Unfortunately, the list of indicators of professional success used in this analysis contained few items about students. Similarly, the list of skills possessed, developed originally for a survey that included both teachers and individuals working in other occupations, did not include skills specifically related to children or instructional strategies. Consequently, it can be concluded that it would be premature to discard the idea of teaching as a client service without further information regarding teacher views based on survey instruments that include such skills and criteria. It is suggested that future research on teacher views should concentrate on finding new measures that are suitable and should focus separately on each of the various methods of evaluation in order to avoid the interdependence among the dependent variables that was unavoidable in this study.

The work reported in this analysis of teacher perceptions and attitudes needs to be distinguished from other types of work concerning teacher evaluation. For example, Soar, Medley, and Coker¹⁹ have analyzed three currently used methods of teacher evaluation and found each method deficient on rational, methodological, and psychometric grounds. They discovered little evidence that existing examinations

used to assess teacher skills (teaching as a craft) correlate very highly with effective teaching. They cited student variability and methodological difficulties in computing gain scores as problematic in using student test results for evaluation purposes (teaching as the production of a product). Finally, they described difficulties in obtaining valid and reliable administrator ratings of teachers as a result of personal subjectivity, the use of high inference measures, and varying conceptions of desirable teaching behavior. This research neither supports nor refutes their analysis. The validity or reliability of various evaluation methods have not been addressed; rather, the researchers have examined teacher characteristics and attitudes toward teaching and evaluation that might be influenced by personal needs, job status concerns, and individual views of professional success. Although the resulting patterns are not definitive, some support for the existence of such influences has been provided, and it is believed that they must be taken into account, along with rational concerns, in devising meaningful teacher evaluation systems. (Reliability may also be a problem in the present study because some of the very low correlations with the dependent variables could be at least partially attributable to measurement error.)

Soar, Medley, and Coker, in indicating that teachers should be evaluated as professionals who deal with complex problems rather than as technicians, viewed teacher resistance to evaluation as reasonable on the basis of currently used measures. They provided, however, little in the way of suggestions for alternative procedures. From the survey analysis reported here, it can be questioned whether teacher resistance is, in fact, as strong and pervasive as has been alleged. Based on the present findings, more provocative is the solution of Barber and Klein²⁰ called "peer-mediated self-appraisal" that involves an interactive system of evaluation based on clearly defined conceptions of teaching effectiveness, diverse assessment modes, and provision for professional growth and improvement, followed by reassessment. In this system, administrator judgments, peer review, self-appraisal, and negotiated objectives interact in a combination of evaluative steps that encompass several demonstrable conceptualizations of teaching. As outlined by Barber and Klein, the system appears so complex that schools might, rather, adopt simplistic unidimensional schemes for teacher evaluation and merit pay. This would be unfortunate because a complex problem may require a complex and systematically developed solution. It seems likely that attempted imposition of simple schemes without input from teachers could, in fact, meet with substantial resistance and eventual failure. Nonetheless, the analysis above indicates that there is sufficient unanimity concerning evaluation methods that teachers can and do perceive as

appropriate to being constructive discussion. Political objections to such discussions are already well known. Further research to refine and verify individual and professional concerns of teachers would be fruitful, especially if such inquiry helped school districts encourage coalitions that would engage in professional level discussions regarding workable evaluation systems. Such discussion should revolve first around conceptualizations of teaching and not around specific measures, thus allowing the development of multidimensional evaluation procedures that include input from a variety of legitimate participants.

NOTES

1. J. Millman, *Handbook of Teacher Evaluation* (Beverly Hills, Calif.: Sage, 1981).
2. R. S. Soar, D. M. Medley, and H. Coker, "Teacher Evaluation: A Critique of Currently Used Methods," *Phi Delta Kappan* (December 1983).
3. National Commission on Excellence in Education, *A Nation at Risk: The Imperative for Educational Reform* (Washington, DC: U.S. Government Printing Office, April 1983).
4. L. Darling-Hammond, A. E. Wise, and S. R. Pease, "Teacher Evaluation in the Organizational Context: A Review of the Literature," *Review of Educational Research* 53 (Fall 1983).
5. The principal components analysis with varimax rotation of the ten survey items (N = 624) resulted in three interpretable factors based on loadings greater than .60, accounting for 44.6% of the variance. The Kaiser statistic for unrotated factors was .67. These factors were named: Student/Parent Judgments, Administrator Judgments, and Test Results. The remaining items did not share substantial variance and, therefore, did not emerge as factors. Hence, they were used as single-item measures.
6. The authors express thanks to Ann E. Austin for initial coding and analysis of survey data; see J. S. Stark, A. E. Austin, M. A. Lowther, D. W. Chapman, and S. M. Hutcherson, "Teacher Certificate Recipients at the University of Michigan, 1946-76: A 1980 Follow-up Study," *Resources in Education* ED209 (January 1982).
7. National Education Association, *The Status of the American Public School Teacher, 1980-81* (Washington: NEA, March 1982).
8. R. Quinn and G. Staines, *The 1977 Quality of Employment Survey* (Ann Arbor, Mich.: Institute for Social Research, The University of Michigan, 1979).
9. The decade in which teaching began was not used as a variable because of its substantial correlation with age.
10. The factor analysis was based on 779 teachers with non-missing data on skill assessment. The Kaiser statistic for unrotated factors was .55, and 43.4% of the variance was accounted for by the five skill factors. One item, "using library and research facilities," did not load on any factor. The complete factor matrix is available from the authors.
11. The factor analysis was based on 763 teachers with non-missing data on assessment of importance of professional criteria. The Kaiser statistic for unrotated factors was .67, and the three factors accounted for 42.2% of the variance in the item set. Two items, "importance of salary" and "importance of publication in professional journals," did not load on any retained factor. Although the first two factors were judged orthogonal, the third factor, "importance of self-direction" was correlated with "importance of leadership/responsibility" ($r = .22$).

12. The factor analysis was based on 699 individuals with non-missing data. The Kaiser statistic for the unrotated factors was .62. The achievement of salary as an indicator of professional success did not load on any retained factor.

13. Darling-Hammond, Wise and Pease, "Teacher Evaluation in the Organizational Context."

14. M. S. Lewis-Beck, *Applied Regression: An Introduction* (Beverly Hills, Calif.: Sage, 1980).

15. The restricted variance of the measure of administrative judgments was noticeable in the strong kurtosis of the constructed two-item scale.

16. Darling-Hammond, Wise, and Pease, "Teacher Evaluation in the Organizational Context."

17. National Education Association, *1983 Nationwide Teacher Opinions Poll* (Washington, DC: NEA, 1983).

18. B. Harootunian and G. P. Yarger, "Teacher's Conceptions of Their Own Success," *Resources in Education* SP 017 372 (February 1981).

19. Soar, Medley, and Coker, "Teacher Evaluation."

20. L. W. Barber and K. Klein, "Merit Pay and Teacher Evaluation," *Phi Delta Kappan* (December 1983).

ERRATA

We sincerely apologize to our authors and readers for occasional lapses that result in errors in final copy in the *EAQ*. Unfortunately, the pressure of time and human oversight seem to be unavoidable in the production of a journal. Both result in inconvenience for all and, at times, embarrassment for authors, editors, and publishers. The following have come to our attention from the Spring 1984 (Vol. 20, No. 2) issue of the *EAQ*:

1. Page 33, line 9, should read: ($r = .121, p < .05$)
2. Page 34, table note b, should read: $p > .05$
3. Page 35, line 18, should read: $\beta = -.245$ (Influence)
4. Page 66, line 16, should read: (b), which states . . .
5. Page 68, line 8, should read: After the replay, each person was asked a question similar to that used by . . .
6. Page 74, line 4 after the heading, should read: (acri-)monious "showdowns" ¹⁵ appear(ing) at worst suspect and at best . . .
7. Page 78, line 4 after the heading: Elocutionary should be Illocutionary
8. Pages 86 and 87 are misnumbered: Page 86 should be page 87, and page 87 should be page 86 (the pages were reversed in production)
9. Page 87, line 15: reply should be replay
10. Page 90, line 7: Zeleznick should be Zaleznik

The publication of a journal requires the work of many people. Thus, the potential for human error at a number of points in the process is ever-present. All of us are aware of this and are attentive to presenting authors and the *EAQ* at their best. When lapses occur, we do regret them, and our apologies are extended to all for the above and other unacknowledged mistakes that have appeared in final copy of the *Quarterly*. As well, we continue our pursuit of accuracy and the highest level of quality in the publication of the *EAQ*.