W. Robert Dixon
William C. Morse

The Prediction of Teaching Performance:

Empathic Potential

REPORTS and summaries of the many attempts to predict teaching performance appear regularly in the literature. Yet the teaching profession is still faced with severe limitations in both the assessment and prediction of teacher performance.[1] It is small comfort to recognize that a similar condition exists in social work and clinical psychology, somewhat allied professions which require a broad complex of interpersonal relations.[2]

Two conditions account for this discouraging situation. First, more often than not, value judgments—evaluations in terms of "good" teaching or "poor" teaching—have been substituted for any actual assessment of teaching performance. The fundamental fact is that opinions differ on what is good and what is poor. Sophisticated judges differ in educational philosophy, for one thing, and then too the same teacher act might objectively be "good" in one context and "poor" in another. We lack any constant criterion in appraising teaching. The second major difficulty is the tendency to consider teaching as a global phenomenon. A single global assessment

runs counter to current psychological appreciation of a complex task. For example, we do not think of two well-adjusted persons having the same test profile, or two maladjusted individuals having equated pathologies. Likewise, teaching is not a unitary performance. It more likely consists of several inter-related but discernable functions, each of which has to be abstracted and differentiated from the others.[3]

But there are no universally recognized dimensions of teaching. By a time-study approach, one could identify the myriad of things teachers do. However, this approach would fail to distinguish between the significant and the trivial. It should be obvious that efforts to predict teaching performance must deal with significant aspects of teaching and at the same time these elements of teaching must be cast in the form of operative dimensions rather than value judgments. We need to identify and quantify the significant teaching syndromes.

After some preliminary exploration of possible patterns, predicting the teacher's capacity to empathize with pupils was selected for study.

Empathy

The general topic of empathy is currently receiving considerable research attention. Several criterion measures have been developed. Work has been done on factors relating to the presence or absence of this capacity which in turn means that we had suggestions available to use for predictive purposes. There are two basic theoretical definitions of empathy. In the first definition, empathy is perceived as the capacity to put oneself in the other fellow's shoes, as shown by the ability to predict the way the other person will respond in given situations. It can be seen that this is an intellectual type of experience. In contrast, the second definition of empathy implies an emotional condition, a feeling state which may range from close and warm to distant and formal.

It is not too difficult to arrange a situation to test the capacity to empathize when defining empathy as some kind of an intellectual performance. A given subject can respond as he thinks someone else would, his response then being checked against the actual response. The less the difference, the greater the empathy. While this appears simple and straightforward, difficulties appear as soon as one attempts to utilize these difference measures. Let us suppose that someone predicts accurately but the target responses are identical with the self responses of the subject. Is this equal to hitting the target when the responses are at variance with the subject's responses? Also, are differences equal regardless of where they fall on the scale? It may be that a difference near the mean is quite at variance with an equal distance further toward the extremes of the distribution. Also, what should be done about the direction of the difference? Is an underestimation equal to the same amount of overestimation? Should they cancel each other out? Painstaking studies have been conducted to find ways

to handle these difficulties.[4, 5, 6, 7, 8] It is interesting that the empirical findings of the present study parallel much of this work.

But the more one studies empathy in teaching the more it becomes evident that the capacity to intellectually perceive how another person will respond is not an index of the feeling tone which high empathy should produce. The important quality of empathy, as we recognize it in teaching *is a highly interpersonal phenomenon with the subject and object bound up in a mutual response.* The teacher and the pupil develop a positive relationship. There is ample theoretical basis for this use of the process in the literature. As pointed out above, this constitutes the second major use of the term. Christine Olden[9] has clarified the non-intellectual aspects in a thoughtful article. The term encompasses such expressions as "gifted with children," "understands children," and "feeling for children." The essence of her theory is found in the following quotation:

⁴ Margaret B. Luszki, "Empathetic Ability and Social Perception." (Unpublished doctor's thesis, University of Michigan, 1951) xi + 262 p.

⁵ Irving E. Bender and Albert H. Hastorf, "On Measuring Generalized Empathic Ability," *Journal of Abnormal and Social Psychology* 48:503-06; October 1953.

⁶ George C. Stone and others, "Generality of Accuracy in Perceiving Standard Persons." (Champaign-Urbana: Bureau of Educational Research, University of Illinois, 1954) 13 p. Mimeographed.

⁷ Lee J. Cronbach, "Processes Affecting Scores on Understanding of Others and Assumed Similarity," *Psychological Bulletin* 52:258-66; May 1956.

⁸ George S. Leavitt and others, "The Intermediate Key in the Analysis of Interpersonal Perception," *Psychological Bulletin* 53:258-66; May 1956.

⁹ "On Adult Empathy With Children," *Psychoanalytic Study of the Child* (New York: International Universities Press, 1953) Volume 8, p. 111-26.

Empathy is the capacity of the subject instinctively and intuitively to feel as the subject does. . . . Empathy may be described as a feeling that emerges spontaneously in social contact, that enables the subject instantaneously to sense the object's apparent emotions . . . to trespass the object's screen of defenses behind which the real feelings may hide.[10]

In Freudian terms, this must be a nonnarcissistic but libidinized feeling for the other person. Olden further illustrates the nature of mature and immature empathic responses. Briefly, an immature state would produce narcissistic self projection, working out one's own needs. The more mature person would respond in terms of the object's needs. Thus, Olden brings in the person on the receiving end of the relationship, though there would be expected individual differences in the pupil's potentials for response as well as differences in the ability of a teacher to relate equally with all types of pupils. To be mature requires some degree of self-realization, understanding, and objectivity. What is meant by "instinctive" and "intuitive" is not too explicit, but illustrations suggest that it need not be conscious. Coming spontaneously from the deeper levels of personality, mature empathy is a consequence of adequate socialization and concern for others—the opposite of the psychopathic personality response. All of this implies the generation of a positive interpersonal feeling between the teacher and pupil.

This meaning, derived from Olden's statement, is the teaching factor postulated as empathy. It is central to any mental hygiene approach. Akin to rapport, it lies at the core of the classroom tone. It is clear on both theoretical and statistical grounds that "intellectual" measures are not a promising source of

such an index of empathy. The appraisal of empathic response requires some assessment of the feeling of the recipients (here, the pupils). Examples of feeling might be obtained in different ways: (1) clinical observation of the quality of interaction between pupils and teachers; (2) clinical interviews with pupils to obtain depth material related to the feeling for the teacher; (3) projective test devices to collect feelings about teachers; (4) direct questioning via interview or questionnaire on a conscious level. Attempts to use direct observation on such an elusive quality as empathy proved upon actual test to be very difficult and expensive. With large populations, intensive interviews were prohibitive in time. Interest in projective devices has run the complete circle from skepticism to almost universal uncritical use until finally the point has been reached where it is considered one type of tool with both specific advantages and limitations. In this case, the specific point of reference being a given teacher, it would be very difficult to disguise the target and still relate projected feelings to that target with any assurance. Admittedly, the sense of being accepted by and feeling positive toward a teacher is certainly not a simple phenomenon to assess and probably appears at both the conscious and unconscious levels. However, it is an area in which pupils have definite ideas which they express freely.[11, 12, 13]

[11] Sister Mary Amatora, "Can Elementary School Children Discriminate Certain Traits in Their Teachers?" *Child Development* 23:75-80; March 1952.

[12] Sister Mary Amatora, "A Diagnostic Teacher-Rating Scale," *Journal of Psychology* 30:395-99; December 1950.

[13] Philip Diskin, "A Study of Predictive Empathy and the Ability of Student Teachers to Maintain Harmonious Interpersonal Relations in Selected Elementary Classrooms." (Unpublished doctor's thesis, the University of Michigan, 1955) x + 175 p.

[10] *Ibid.*, p. 115.

The Assessment Inventory

Our analysis of the methods of assessing the empathic status in the classroom led to the conclusion that an assessment inventory had to be worked out which could be used with pupils. There were two main problems to be solved in constructing such an inventory: first, deciding what items were to be included, and second, finding some way of comparing the attitudes of the various classes. We started developing the questionnaire with the idea of having items which struck at the important feelings pupils have about teachers and proceeding to items which would let us know how these feelings might influence other behavior. Hence the first questions have to do with such feelings as "like," "trust," "understand." Then, some items were added which required the pupil to mark on the opposite side of the scale if he was to be consistent, for example, questions on the frequency of the teacher's efforts to embarrass the pupil, or how often he felt uncomfortable in a teacher's class. Finally, items were added which sought to discover whether the basic feelings transferred to such areas as "talk over personal problems" or "converse about non-school matters."

The problem of identifying the overall acceptance attitudes of the various classes was important because we needed some way of knowing whether or not we were starting with comparable groups. Hence twelve "control" questions were placed in the questionnaire. These items asked about the same attitudes as the other questions, but used as the reference point adults, peers, or teachers in general. Using the pupils' responses to these items, it was possible to examine classes for differences in basic attitudes about teachers. An analysis of variance revealed no significant differences in this interclass comparison. Hence it was reasonable to treat student teachers as starting with classes in which the attitudes

with which we were concerned were similar. Another analysis of the control questions showed no significant differences in the responses of boys and girls. In addition, we checked the reliability of the questionnaire through responses to the control questions. With a sample of one hundred secondary pupils and using the test-retest technique, a reliability coefficient of .92 was obtained.

At this point, let us turn to a consideration of the second part of the problem: how to predict this empathic capacity. Ideally, we were looking for some kind of group-administered measure or measures which could be given to teacher candidates early in their training. Since we are involved in the preparation of teachers, it would be possible to administer the predictive measures somewhere during the professional training and then to ascertain the "empathic" reactions of pupils to these candidates during their student teaching. Thus we could check the validity of our predictions.

The literature offers many suggestions, though the criterion of "empathy" differed from that used in this study. Several dimensions of personality seem to offer the greatest potential for the predictive problem we faced. The first of these we might call the "basic adjustment level." There is a considerable emphasis in the literature on the need for the teacher to be well-adjusted to do a satisfactory job in the profession. It is a moot question whether "good" adjustment in the usual sense of the term is a prerequisite to the empathic aspect of teaching. One can argue that the more mature the teacher—the more objective he is in social relationships—the more accepting he can be of pupils. But it can also be postulated that the individual with heightened sensitivities—who may also be somewhat neurotic—is really in the best position to be accepting, unless of course the neurotic aspect happens to be of such a nature as to produce coun-

ter-effects. For example, a teacher who has concern for being rejected because of his own life history might be able to identify more closely with such a child, or, on the other hand, might be unable to accept such a child, depending on how the self-feeling of being rejected has been incorporated. Or, a person may, by all indications, be well adjusted but not be particularly focused on social relationships. For example, a scientifically or artistically oriented individual might be well-adjusted or "normal," but not involved enough to empathize with pupils. The fact is that adjustment ceases to have the once-implied unitary nature. For these reasons, there is a question as to what adjustment, as an over-all condition, may have to do with the capacity of a teacher to empathize.

It is difficult to develop confidence in personality scales, especially group-administered paper-and-pencil tests. Yet this is the direction of our effort, since the search was for some device useful for large groups which could be easily administered and scored. The Maslow Security-Insecurity Inventory was selected. This scale is designed to detect and measure security, which is considered almost synonymous with mental health.[14] The security aspect includes feelings of being liked and "belonging," friendliness, affection, emotional stability, and the like. In fact, the list of sub-syndromes reads almost like the supposed correlates of empathy. For this reason, it was well-suited for the present study.

The second factor relating to personality which we used was derived from the early life experience of the subjects. There is considerable feeling that empathic ability is related to the adequacy of family experience. Rosalind F.

Dymond[15] found that the high-empathy group rated their family relationship as close, with considerable affection being displayed.

While it is possible to obtain estimates of this from biographical data, the life histories at our disposal were not complete in this regard. It was thought that a more systematic approach to the problem might prove useful. To do this, Gabriel Elias' Family Adjustment Test was selected.[16] This test, according to Elias, measures the intra-family "homeyness" in a projected manner by asking about attitudes toward the general community life. The description of the "homey" feelings is a replica of Dymond's concept previously mentioned. More than three thousand individuals have been tested to secure the standardization. It is said to have "unusual predictive ability."

The third dimension of personality which we considered can be called "flexibility." It was reasoned that the rigid, stereotyped personality would be unable to take the many diverse roles necessary in empathic responses to a class of public-school pupils. The rigid person would project his own feelings rather than introject others. Thus, lower empathy would result. Cottrell and Dymond hold that empathy is related to rigidity.[17] The F scale[18] was used as a test for this reason. The latest available form was employed.

The fourth aspect of personality which we investigated is in the area of "self-acceptance." Persons with high capacity

[14] A. H. Maslow, Manual for the Security-Insecurity Inventory (Stanford, California: Stanford University, 1952).

[15] "A Preliminary Investigation of the Relation of Insight and Empathy," *Journal of Consulting Psychology* 12:228-34; July-August, 1948.

[16] *The Family Adjustment Test Manual.* (New York: Psychometric Affiliates, 1952).

[17] Leonard S. Cottrell, Jr. and Rosalind F. Dymond, "The Empathetic Responses," *Psychiatry* 12:355-59; November 1949.

[18] T. W. Adorno and others, *The Authoritarian Personality* (New York: Harper and Brothers, 1950) xxiii + 990 p.

to empathize are not only thought to be well-adjusted, but they are expected to be high in self-insight and self-acceptance. This idea has been discussed by Ronald Taft,[19] as well as by Dymond and Cottrell. The two instruments we selected derive their scores through assessing the stability of one's self concept. The hypothesis is that those whose self concept is relatively stable (those whose negative and positive self-ratings are minimally discrepant) are individuals with the highest self-insight and self-acceptance. The first instrument was developed by John J. Brownfain.[20] This inventory consists of twenty general personality traits such as intelligence, cheerfulness, popularity. Each trait was defined in terms of extremes. Directions are given so that the subject first makes the most accurate estimate of how he sees himself. A scale ranging from 1 to 8 is used. In the next situation, the subject on the same set of traits rates himself in such a manner as to take a very favorable view of his situation by giving himself the highest rating that he realistically thinks he should. The next direction asks the individual to refrain from giving himself the benefit of any doubt and to rate himself the lowest he realistically thinks he should. The discrepancy between the highest and lowest ratings represents the score derived from this instrument. In the same area we used a second instrument which was developed by Glen Rasmussen.[21] In this second inventory (a revision of one used in his thesis), Rasmussen has selected

items specifically relating to the performance of teachers. The same kinds of directions are included for Rasmussen's inventory as for the inventory developed by Brownfain.

The final dimension of personality which we investigated we can label "the capacity for affectionate responses." A recent test has been developed which assesses in a projected fashion a teacher's relationship to pupils. This has been identified as the TAT for Teachers after the original Thematic Aperception Test. Actually, the test is called the Adult-Child Interaction Test and was devised by Theron Alexander.[22] Although the stories secured from this instrument can be analyzed in several ways, the measure sought here is the capacity to show affectionate behavior. This test can be administered on a group basis, each of the eight pictures being shown in turn while the subject is given time to write out responses.

Results

Our analysis was based on the predictive data collected from seniors at the University of Michigan who were engaged in their student teaching, and on the assessment data secured from responses to the questionnaire by pupils in their classes. Ratings were also obtained from the supervising teachers. Although we originally started with both elementary and secondary groups, the detailed analysis was confined to the secondary group. Included were ninety-seven student teachers and more than two thousand pupils.

The responses of the pupils to the assessment inventory were made on a five-point scale.[23] For example, in an-

[19] "The Ability to Judge People," *Psychological Bulletin* 52:1-24; January 1955.

[20] John J. Brownfain, "Stability of the Self Concept as a Dimension of Personality," *Journal of Abnormal and Social Psychology* 47:597-606; July 1952.

[21] "Teacher's Membership in Informal Groups and his Potential for Feelings of Professional Failure." (Unpublished doctor's thesis, University of Michigan, 1952) v + 78 p.

[22] "The Prediction of Teacher-Pupil Interaction with a Projected Test," *Journal of Clinical Psychology* 6:273-76; July 1950.

[23] Copies of the assessment inventory are available upon request.

swer to the question, "How much do you trust this teacher?" possible answers, rated one to five, were "Very much," "quite a bit," "some," "only a little," or "not at all."

The relationships of the responses were determined by developing a correlation matrix. Low positive correlations were found for all but one combination. Hence, a standard "E" score was derived for each student teacher. This "E" score was computed by summing an individual's standard scores on the various questions and securing the average.

The first important question to be asked was: Are student teachers who received "good" empathy scores from pupil's ratings perceived as better teachers by these respondents? One of the questions in the pupil assessment inventory asked: "Everything considered, how would you rate this person as a teacher?" Responses were made on the following scale: " (1) exceptionally good, (2) quite good, (3) average, (4) somewhat below average, (5) no good at all." Students in the top 20 per cent E scores were compared with those in the bottom 20 per cent.

Three ratings were studied. First, the pupils gave a 1.59 mean over-all average rating to the "good" group and a 2.70 score to the "poor" group. A *t* test applied to these scores showed that they were significantly different (1 per cent level). In other words, as far as the pupils were concerned, the student teachers who had "good" empathy were also better teachers. Second, an analysis was made of the evaluation of the supervising teachers. They saw the "good" empathy groups as being significantly (1 per cent level) better teachers than the "poor" groups. How the student teachers rated themselves as teachers gave us a third comparison. Are you surprised to learn that there was *no* significant difference between the mean over-all self-ratings of the "good" and "poor" groups?

In a way, such a result is somewhat discouraging, since "self-evaluation" has been built up as an important and valid procedure. However, in this study, it did not give results which were consistent with data obtained from pupils and supervising teachers.

Nevertheless, we did find evidence that one of our original hypotheses was substantiated. Pupils and supervising teachers considered student teachers with "good" empathy to be better teachers than those with "poor" empathy. The next task was to ascertain whether or not the predictive instruments had performed as expected.

To start this phase of the analysis, the "good" and "poor" E groups were compared on the basis of their scores on each of the predictive measures. On the F scale, the Security-Insecurity Inventory, the Elias, the Teacher's TAT, and the Self-Concept Inventory, no significant differences were found. Only the Teacher's Self-Concept Inventory scores bore out our idea. A *t* of 9.33 was obtained and this result is significant at the 1 per cent level of confidence. Looking at the Teacher Self-Concept scores of the two groups, we saw that the "good" empathy group had much lower scores than the "poor" group. In other words, the student teachers who developed very positive feelings toward their pupils were significantly more stable in their appraisal of themselves as teachers.

Correlation coefficients were obtained for the various predictive instruments and the standard E score. Only the *r* of .42 between the Teacher Self-Concept Inventory and E attained significance. From the results described in the previous paragraph this was to be expected.

The extremes (top and bottom 20 per cent) on each of the predictive instruments were compared using the results from the question on over-all teacher effectiveness. In no instance did we find any significant difference in the ratings

of the pupils. Using the supervising teachers' ratings, one significant difference was found. The "secure" group was rated as being better (5 per cent level of confidence) than the "insecure" group.

Conclusions

We had not expected to find that an analysis of personality traits would be productive as far as predicting global performance in teaching was concerned. This expectation was verified. However, we had anticipated that personality measures would permit us to predict performance in an aspect of teaching which theoretically is especially dependent upon personality. We were not able to substantiate this expectation. It may be that the instruments used were not refined enough for this purpose. However, another possibility is that the "teaching situation" is a condition which generates responses that are not tied specifically to particular dimensions of personality. The personality factors considered in this study and with this population, within broad limits, do not seem to be significant for developing good personal relationships with pupils. It appears that the teaching role tends to mediate differences in personality.

Our results indicate that pupils do have different perceptions of the individual's ability to develop good interpersonal relationships. An important question to ask is, What things prevent individuals from establishing this relationship? One possibility is the perception one has of his role as a teacher. The results of the Teacher's Self-Concept Inventory indicate that stability in the perception of one's role as a teacher may enhance one's chances of developing good personal relationships. For individuals working in teacher education, this research would indicate that we should *not* try to set up the ideal image or model of a teacher. Instead the crucial thing seems to be to help the student develop confidence in *his* teacher role. The problem is not whether he has anxiety or lack of anxiety in general, but the extent to which any anxiety he may have effects his teaching role.

We found verification of our hypothesis that individuals who have high empathic capacity as measured by pupil responses are also seen as better teachers. The problem of predicting this attribute is still with us. The dimension which may have the most potential for predicting this aspect of teaching appears to be tied up with an individual's concept of his role as a teacher. We need to have more detailed information about the development of the teacher self-concept. The important factor is somewhere in the synthesis of self-concept and the teacher-role-concept which develops in each individual.

The Teaching Profession, 1961

1,400,000	teachers in public elementary and secondary schools
220,000	teachers in private elementary and secondary schools
130,000	school administrators, supervisors, consultants, researchers and other specialists in elementary and secondary schools
350,000	professional personnal in higher education institutions
25,000	professional staff members in professional organizations, in government offices of education, and in private agencies with educational programs
2,125,000	This is the size of the teaching profession, 1961

—*National Commission on Teacher Education and Professional Standards, National Education Association*, NEW HORIZONS FOR THE TEACHING PROFESSION *(Washington, D.C.: the Association, 1961) p. 5*