

RESEARCH IN STUDENT TEACHING: TRANSFER  
FROM  
THE IVORY TOWER TO THE CLASSROOM

TERRY LOVELACE

University of Southwestern Louisiana

CHARLES MARTIN, JR.

Southeastern Louisiana University

MARTA COLLINS

Louisiana State University

BONNIE CRAMMOND

Southeastern Louisiana University

JOANN DAUZAT

Grambling State University

SAM DAUZAT

Louisiana Tech University

ROGER DESANTI

University of New Orleans

SUTTON FLYNT

University of Kansas-Pittsburg

Successful performance on the *National Teacher Examination* (NTE) is required by law in 17 states in order for graduates of approved teacher training programs to apply for certification by the state departments of education. Little research has been done on the correlation between student teachers' scores on the revised NTE and their actual classroom performance. This study, funded by the Louisiana Board of Regents, was designed to answer two research questions:

(1) How well do the scores from the revised NTE, the *American College Test* (ACT), cumulative grade point average, sex, race, age, and certification level predict student teachers' classroom performance, as measured by three instruments from the *Teacher Performance Assessment Instruments* (Capie, Johnson, Anderson, Ellett, & Okey, 1979)?

(2) How well do student teachers chosen at random from seven state universities perform in the classroom, as measured by the TPAI?

#### Review of the Literature

The American public indicated in a 1979 Gallup survey that they believed the best way to improve the quality of public schools was to improve the quality of classroom instruction. In response to the demand for improved teacher performance, many states and individual school districts have initiated policies which affect the certification, hiring, and tenure of teachers (Gudridge, 1980; Viaanderen, 1980). In many states this has led to legislation which has, at times, prescribed the teacher evaluation instruments to be used (Beckham, 1981).

The goal of evaluation of teacher effectiveness at the preservice level is two-fold: degree-granting institutions and certifica-

tion agencies want some form of quality control, and also wish to diagnose and remediate preservice teachers' weaknesses, in order to improve classroom performance. Darling-Hammond, Wise, and Pease (1983) stated these concepts in terms of detecting, preventing, and correcting incompetencies. Detecting involves the development and application of instruments designed to measure teacher knowledge or behavior. Preventing focuses on the development of effective teacher education programs. Correcting involves intervention utilizing results of evaluation measures.

Research in teacher evaluation has focused on the identification of factors which influence teacher performance and on the establishment of valid and reliable measures of these factors. Previous research has centered on studies of intelligence, academic ability, personality characteristics, the teacher's ability to manipulate and/or control the educational environment, time on task, and teaching knowledge. Review of the literature reveals:

(1) Studies investigating the relationship between teachers' intelligence, college grade point average, college board scores and teacher effectiveness reveal small positive correlations (Hawn, 1981; Hellfritzch, 1945; LaDuke, 1945; Lovelace, Collins & DeSanti, 1984; Rostker, 1945). Some preliminary studies also indicate that a measure of writing may prove a better predictor of classroom performance than other measures of academic achievement.

(2) Prediction of teacher performance based on teachers' attitudes, flexibility, interests, problem-solving ability, creativity, and professional commitment is ineffective. King (1981) reviewed the literature and found no set of skills, interests, or abilities was able to discriminate between effective and ineffective teachers. Cronbach (1967), Hunt (1975), and Joyce and Weil (1972) found little, if any, correlation between teacher characteristics and effectiveness based on the match of teacher and student traits.

(3) Ellett, Capie, Okey, and Johnson (1978) found the teacher's ability to manipulate and control the classroom environment is a factor related to teacher performance. Ellett, Capie, and Johnson (1980) urged that research be conducted in this little-explored area.

(4) Academic engagement, or time and quality of time on task, appears to be an important factor in evaluating teacher effectiveness. The teacher who is able to keep students on task, involved in meaningful work, is more likely to be successful and affect subsequent learning and achievement (Lomax & Cooley, 1979). More specifically, Anderson and Scott (1978) determined that certain types of academic activities were more likely to produce quality engagement than others. Despite this, few studies have been conducted which related specific tasks to academic engagement (Anderson, 1976; Capie & Tobin, 1980).

(5) Though it seems logical that the greater the teacher's knowledge in the teaching field, the greater the chance of that person being an effective teacher, little research has been published which directly relates teaching knowledge to success as a teacher (Schalock, 1979). The studies reported in the literature have almost exclusively used grade point average or NTE scores as measures of teaching knowledge. Results of these studies reveal low correlation coefficients, at best.

## Methodology

*Determining interrater reliability.* Eight investigators representing seven major Louisiana universities were trained to observe student teachers with the TPAI in the fall of 1983 by Dr. Charles Johnson, Professor, University of Georgia, one of the senior authors of the instrument. All eight raters were college professors, who as a group, had acquired 70 years experience in observing student teachers. These raters spent three days learning to administer the TPAI. At the end of the training, they examined a sample portfolio, observed a videotaped lesson, and rated the teacher's performance on three instruments of the TPAI: *Teaching Plans and Materials*, *Classroom Performance* and *Interpersonal Skills*. Interrater reliability was computed by Dr. Lee Hoffman, Assistant Director, Bureau of Evaluation, Louisiana State Department of Education, and was found to be .875 for each instrument.

Higher interrater reliability is noted when observers can actually go into a classroom and conduct the observation, since the video camera never tapes exactly what a rater would observe during actual classroom instruction. The interrater reliability figures obtained agree with those found in previous studies, which indicate that, with training, agreement rates of 85 to 90 percent are common (Capie, Ellett & Johnson, 1979).

*Sampling procedures.* Ninety college seniors who had completed all prerequisites for student teaching, and who had enrolled in the student teaching practicum, were chosen at random from the population of all seniors student teaching during the academic year 1983-1984 at seven Louisiana universities. Half the students were observed during the fall semester, half during the spring semester. All students participating in this study were assured of total anonymity.

*Data collection.* The eight investigators observed the 90 student teachers three times each, with the exception of one student teacher who resigned from student teaching after the rater had completed two of the three observations. Some 269 observations were conducted, totaling 672 hours of documented clinical observation. Raters also collected the following demographic data and achievement scores on each subject: ACT and NTE scores, cumulative grade point average at the end of the semester prior to student teaching, race, sex, age, and certification level (elementary, secondary, or K-12).

## Results of Data Analyses

Descriptive statistics were used to analyze demographic data. Of the 90 students participating in the study, 74 were female and 16, male. Of the females, 11 were black, one was Hispanic, and 62, white. Four of the males were black, one was Hispanic, and 11 were white. Mean age was 24.26 years, with a mode of 22 years, and a range of 20-46 years. Mean cumulative grade point average was 2.98, with a standard deviation of 0.44 on a four-point grading scale. Of the 90 subjects, 38 were transfer students. Forty-eight desired elementary certification, while 39 would apply for secondary teaching certificates. Two students were enrolled in programs which would enable them to be certified K-12.

The investigators found it difficult in some cases to obtain students' ACT and NTE scores. Fifty-nine of the 89 subjects had scores available from the ACT, while 78 took the Core Battery on the NTE. Only 57 completed the Area Examina-

tion in their field of certification. Students' scores on the ACT ranged from 15.63 (Math) to 19.44 (Science), with a Composite Score of 17.35 (s.d. = 5.17). National percentile scores on the NTE were: General Knowledge, 45th percentile; Professional Knowledge, 38th percentile; Communication Skills, 48th percentile.

On the three instruments of the TPAI, students are considered to have demonstrated minimal competency on any of the 45 indicators if they obtain a score of three on the five-point scale used for each indicator. (Note: In order to observe the best possible instruction, all observations were scheduled at least one week in advance. Also, the scores on the indicators were averaged across the three observations, in order to determine students' usual classroom performance.) As a group, the 89 subjects in this study scored at the minimal level of competency on only one indicator on the *Teaching Plans and Materials Instrument*. They achieved minimum competency on 15 of 20 indicators on the *Classroom Procedures Instrument*, and met the criteria for minimum competency on all 10 indicators on the *Interpersonal Skills Instrument*.

Pearson product moment correlation coefficients were computed to determine the extent of the relationship between:

(1) students' ACT scores and grade point average, sex, race, age, and level of certification. As a group, these correlations were very low ( $r = .57$  and below). No strong relationships were observed.

(2) NTE and ACT scores. High correlations were noted between subjects' ACT Composite Score and the NTE General Knowledge and NTE Communication Skills subtests ( $r = .83$ ,  $r = .75$ ). These tests appear to be measuring the same factors.

(3) NTE scores and cumulative grade point average, sex, race, age and certification level. Again, as with the ACT scores, correlations were weak, with  $r$ 's ranging from .02 to .42. No strong correlations were noted.

(4) student teachers' scores on subtests of the NTE and TPAI. Results of data analyses revealed no strong correlation coefficients between the NTE and TPAI, with  $r$ 's ranging from .09 to .33.

Correlated t-tests were computed to determine the degree of change in the effectiveness of the 89 student teachers from the first to the third observation. Results of the analyses reveal no statistically-significant difference in the 89 subjects' performance from the first to the third observation on any one of the three TPAI instruments. In other words, student teachers exhibited little change in planning for instruction, actual classroom performance, and interpersonal skills from the beginning to the end of the student teaching practicum. As previously noted, the subjects were deficient in all skills measured by the *Teaching Plans and Materials Instrument*, and in 15 of the 20 indicators on the *Classroom Procedures Instrument*, while they were judged minimally competent on all indicators of the *Interpersonal Skills Instrument*.

### Results and Conclusions

Results of this study obtained through statistical analyses of data collected are summarized below:

(1) As a group, the 89 student teachers met or exceed the minimum competency level on 26 (58%) of the behavioral indicators on the three TPAI instruments used in this study.

(2) Correlation coefficients computed between subjects' ACT scores and their cumulative grade point average, sex, race, age, and level of certification are very low, and reveal no educationally significant relationships.

(3) Analyses of correlations between NTE and ACT subtests reveal strong correlations between the subjects' ACT composite scores and their scores on the NTE General Knowledge and Professional Knowledge subtests.

(4) Results of correlation coefficients computed between NTE scores and students' cumulative grade point average, sex, race, age, and level of certification suggest that students' performance on the NTE cannot be predicted on the basis of any of these variables.

(5) Correlations between student teachers' academic achievement on the NTE and their actual classroom performance as measured by the TPAI are very low. No robust correlations were noted in any of the analyses.

(6) No statistically-significant differences were noted between the 89 subjects' performance on the three instruments of the TPAI from the first to the third observation.

Conclusions drawn from the results of the data analyses include:

(1) Student teachers' performance on the revised NTE is a poor predictor of actual classroom performance, as measured by the TPAI. While the NTE serves as a paper-and-pencil test of teacher knowledge, it will not reveal how well student teachers will actually perform in the classroom.

(2) Student teacher performance, as measured by the TPAI, was highest in the area of interpersonal skills (classroom management and discipline), low in the area of classroom performance, and lowest in the area of planning for instruction. The NTE, used in conjunction with the TPAI or a similar instrument with high validity and reliability, could perhaps provide a more accurate description of student teachers' abilities. Such a combination of measures of teacher effectiveness would enable certification agencies and local school districts to more effectively select teachers, which should ensure higher academic performance among school children.

(3) The quantity and quality of supervision of classroom performance of student teachers, regardless of the source, is vital to the development of teacher effectiveness. Student teachers enrolled in this study were critical of the quality and quantity of supervision they received. Training college and classroom teachers who supervise student teachers in clinical supervision with an instrument similar to the TPAI might result in an increase in student teachers' effectiveness.

(4) The high correlations between certain subtests of the ACT and revised NTE suggest that these instruments may be measuring the same variables. Rather than overtesting students, perhaps teacher training institutions might consider exempting students who score high on the ACT from similar subtests on the NTE, saving students needless expense.

Results and conclusions from this study are generalizable only to similar populations of student teachers enrolled in teacher training programs have characteristics in common with those seven which participated in this study.

## REFERENCES

- Anderson, L.W. (1976). An empirical investigation of individual differences in time to learn. *Journal of Educational Psychology, 68*, 226-233.
- Anderson, L.W. & Scott, L.L. (1978). The relationship among teaching methods, student characteristics, and student involvement in learning. *Journal of Teacher Education, 29*, 52-57.
- Beckham, J.C. (1981). *Legal aspects of teacher evaluation*. Topeka, KA: National Organization on Legal Problems of Education.
- Capie, W., Ellett, C.D., & Johnson, C.E. (1979). *Selected investigations of the reliability of the Teacher Performance Assessment Instruments*. Athens, GA: Teacher Assessment Project.
- Capie, W., Johnson, C., Anderson, J., Ellett, C.D., & Okey, J. (1979). *Teacher Performance Assessment Instruments*. Athens, GA: Teacher Assessment Project.
- Capie, W. & Tobin, K. (1980, March). *The effects of formal reasoning ability, locus of controls and classroom process variables on science achievement*. Paper presented at the meeting of the Eastern Educational Research Association, Norfolk.
- Cronbach, L.J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist, 30*, 116-127.
- Darling-Hammond, L., Wise, Q.E., & Pease, S.R. (1983). Teacher evaluation in the organizational context: A review of the literature. *Review of Educational Research, 53*, 285-328.
- Ellett, C.D., Capie, W., & Johnson, C.E. (1980, March). *Teacher performance and pupil perceptions of the social environment of learning*. Paper presented at the meeting of the Eastern Educational Research Association, Norfolk.
- Ellett, C.D., Capie, W., Okey, J.R., & Johnson, C.E. (1978). *Validating the Teacher Performance Assessment Instrument against student perception of school climate/learning characteristics*. Athens, GA: Teacher Assessment Project.
- Gudridge, B.M. (1980). *Teacher competency: Problems and solutions*. Arlington, VA: American Association of School Administrators.
- Hawn, H.C. (1981). *Implications of the Georgia TPAI for evaluating preservice teacher education programs at the University of Georgia*. Unpublished manuscript.
- Hellfritzch, A.G. (1945). A factor analysis of teacher abilities. *Journal of Experimental Education, 14*, 166-199.
- Hunt, D.E. (1975). Person-environment interaction: A challenge found wanting before it was tried. *Review of Educational Research, 45*, 209-230.
- Joyce, B.R., & Weil, M. (1972). *Models of teaching*. Englewood Cliffs, NJ: Prentice Hall.
- King, J.A. (1981). Beyond classroom walls: Indirect measures of teacher competence. In J. Millman (Ed.), *Handbook of teacher evaluation*. Beverly Hills, CA: Sage Publications.
- LaDuke, D.V. (1945). The measurement of teaching ability. *Journal of Experimental Education, 14*, 75-100.
- Lomax, R.G., & Cooley, W.W. (1979, April). *Student achievement/instructional time relationships*. Paper presented at the meeting of the American Educational Research Association, San Francisco.
- Lovelace, T.L., Collins, M., & DeSanti, R. (1984). Predicting and evaluating student teachers' performance in reading instruction. In G. McNinch (Ed.), *Fourth Yearbook of the American Reading Forum*, Athens, GA: American Reading Forum.
- Rostker, L.E. (1945). The measurement of teaching ability. *Journal of Experimental Education, 14*, 5-51.
- Schalock, D. (1979). Research in teacher selection. In D.C. Berlinger (Ed.), *Review of Research in Education*. New York, NY: American Educational Research Association.
- Viaanderen, R. (1980). *Trends in competency-based teacher certification*. Denver, CO: Education Commission of the States.