

## CHANGING TEACHER BEHAVIOR AND IMPROVING STUDENT WRITING ACHIEVEMENT

Reauthorization of the Elementary and Secondary Education Act (ESEA) in 2001, known as No Child Left Behind (NCLB), raised accountability levels. All students, grades three through eight, must be tested in reading and mathematics by the 2005-2006 school year. Section 1001 of the NCLB recommended increased professional development (PD) to elevate the quality of instruction delivered to students. Teachers must be highly qualified in core subjects by 2005-2006 as shown by performance on rigorous licensing and certification tests. The NCLB was not the first to advocate for PD and qualified teachers. During the past 50 years several reports, such as *A Nation at Risk* (1983), urged increased PD as one method for raising student achievement. Results of almost 50 years of legislated spending on PD and its direct impact on student achievement have not yet been determined by consistent replicable, empirical studies.

### *Purpose and Hypothesis*

The NCLB legislation includes the phrase “scientifically based research” over 100 times. A stated intent is to increase the use of experimental or quasi-experimental designs to evaluate educational programs, including PD initiatives. The purpose of this empirical study was to determine if there was a measurable difference on the content and organization portion of a narrative writing test by a sample of students taught by teachers who received PD as compared to students whose teachers did not have this training and thus their students did not receive this level of instruction. The PD trained teachers how to provide instruction to students in using the criteria contained within the New Jersey Registered Holistic Scoring Rubric and a set of higher-order (Bloom, 1958) reflective questions as self-assessment and reflection devices when composing, revising, and editing narrative essays.

A lack of replicable research and compelling theoretical literature on PD supported the use of the null hypothesis: There will be no difference in the frequency of scores on a narrative writing assessment between students taught by teachers who received the PD and students of teachers who did not receive it.

### *Definitions of Terms*

Professional and staff development should have two measurable levels of impact: (a) To improve the participants’ observed teaching performance,

and (b) to improve measurable student achievement. Definitions clarify terms used in this study:

*Professional development (PD)* refers to ongoing education through which certified education professionals learn processes that relate to classroom instruction. The planned and long-term focus is on improved performance of professionals and their students.

*Staff development (SD)* is similar to PD except that the audience or target group could include paraprofessionals and non-certified personnel who work with students.

*Job-embedded staff or professional development* is continuing education that occurs in the context of the classroom and focuses on teacher behaviors related to instruction and student learning. Instructional groups are small ( $n < 15$ ). Activities include such things as action research, peer coaching, structured study groups, mentoring, and calibration exercises.

### *Literature Review*

Historically, staff development for teachers consisted of single-day in-service events or workshops with little follow-up. Joyce and Showers (1983) stated that one-day events were largely ineffective relative to classroom implementation and teachers did not include the content of the in-service activities in future lesson planning or implementation. Teachers will not automatically transfer what they learn in workshops into the classroom without assistance (Wood & Thompson, 1993). Achilles, Dickerson, Dockery-Runkel, Egelson, and Epstein (1992) identified an alternative to single-day training, "Successful in-service is not an add-on; it occurs during the day in the laboratory of the classroom.... Some type of continuing "renewal" is needed if educators are to stay current and vibrant" (p. 3). Garet, Porter, Desimone, Birman, and Yoon (2001) identified reform types of staff development as an alternative to traditional in-service. Reform staff development includes study groups, mentoring, coaching, and professional discussion. While traditional workshops tend to occur outside of the teacher's classroom and outside of the regular teaching day, reform activities tend to take place within classrooms during school hours.

Cohen and Hill (1998) examined the impact of California's instructional policy initiatives on fourth-grade students' performance on the mathematics portion of the California Learning Assessment System (CLAS). While learning about the CLAS test had a positive effect on student achievement, the researchers did not find a strong relationship between PD related to mathematics curriculum and increased student achievement as measured by the results on the CLAS. This led the authors to conclude that, "This study confirms that neither teachers' practice nor students' achievement was changed

by the professional development most California teachers had experienced. Still, very large amounts of money are spent every year on just such activities” (p. 33).

One experimental study demonstrated that PD impacted positively the achievement of students. Caulfield-Sloan (2001) divided 27 teachers into groups of 13 and 14 and provided PD in the use of higher-order questioning strategies for science instruction. The size of the participant groups for the PD was always under 15 ( $n < 15$ ). The same trainer taught the teaching strategies to all the participants. Students of teachers trained in these small-group sessions scored significantly higher ( $p < .001$ ) on an open-ended science question assessment than did students of teachers who had not been trained. Caulfield-Sloan stated, “Staff development directly influences instructional practices and pupil performance. The instructional practices of teachers do, in turn, have a significant and measurable impact on the performance of students” (p. 62).

Harwell, D’Amico, Stein, and Gatti (2000) attempted to determine the effects of PD for teachers on student achievement over a 10-year period. The researchers analyzed student reading and math achievement in grades 3-5 during the period 1988-1998 and found only one area in which PD had an impact. The PD did not influence mathematics performance. Students of teachers who reported engagement in structured discussions with colleagues and their principals about literacy instruction had higher classroom averages in reading as measured by the California Test of Basic Skills than did the students of teachers who reported little or no engagement in such activities.

### *Conceptual Framework: Change Models as a Theoretic Construct Underlying Professional Development*

Innovation and change are intimately linked. Change taken in the context of educational innovation could relate to changing behaviors, structures, thinking, or attitudes relative to education. Innovation is change but not all change can be considered innovation. There needs to be a voluntary and deliberate element present for change to be considered an innovation (Cros, 1998). Models exist to help educators facilitate and monitor the change process. For example, as shown in Table 1, Achilles, Reynolds, and Achilles (1997) suggested a change model that incorporated terms from earlier models such as Rogers (1962), Rogers and Shoemaker (1971), Berman and McLaughlin (1974), and Yankelovich (1991). One goal of PD is an observable change in practice at an individual level, such as how teachers teach. Teachers are the acceptors or rejecters of the change at the individual level and diffusion cannot occur without individual teachers accepting the innovation (Rogers & Shoemaker, 1971).

**Table 1**

*Relationship Between Communication Processes and Change Theory*

Communication aspect	Stage I: Awareness	Stage II: Trial/Evaluation	Stage III: Use
Relation of professional development (PD) to the change process	Raise awareness of the need to improve writing and the role of PD; build interest.	Trial and error by teachers. Working through the problems of implementing. Persuasion by the change agent.	Use in the classroom.
Message (Purpose)	Background information/conceptual underpinning of the PD. Understanding the PD.	Building specific skills and strategies: using rubrics and reflective questions to self-assess writing.	Transfer of the skill from PD sessions to classroom situations and various student learning styles.
Methods of transmitting message: Media	Provide literature related to student self-assessment, and conduct small group discussions.	Demonstrations, question and answer sessions, classroom observation, collaborative lesson planning activities.	Coaching, reflection, and application in classroom environment. Frequent "checking in" (classroom visits) to answer teacher questions.
Communication (audience and feedback)	Primarily one-way with some two-way to address questions.	Two-way small group and individual. Question and answer. Discussion.	Two-way and individual. Open give and take of ideas, and any implementation problems.

*Note.* Adapted from Achilles et al. (1997, pp. 132-133).

Communication is embedded in the change process and PD. Diffusion of an innovation is predicated upon someone communicating a new idea or knowledge to others in the system. Learning is a form of change. Once a person learns something new, by definition, that person changes (Achilles, 1986). The theory and model suggest that the change-agent's ability to manage the structure and medium of the communication effectively impacts the successful adoption of an innovation. Table 1 shows the communication/change process used to structure and deliver the PD provided in this study. The sender or "transmitter" is the PD provider. As a recipient moves toward acceptance of the proposed innovation that person passes through each stage. The communication elements become more personalized the closer the individual gets to independent use of the PD.

## **The Research Study**

### *Subjects and Training*

Five fourth-grade teachers and their classes (98 fourth-grade regular education students) participated. Two teachers received the treatment and three teachers comprised the control group. The researcher delivered the PD using a small group size ( $n=2$ ), job-embedded format and followed a communication/change structure (Achilles 1986; Achilles et al., 1997). The job-embedded PD was based on two premises: (a) Successful PD focuses on teacher behaviors that impact classroom instruction; (b) PD is linked to problem solving and organizational change. Each participant ( $n=2$ ) received individualized feedback and instruction tailored to meet his/her level of experience and understanding.

### *Teacher Characteristics and Comparisons*

The teacher groups were compared on the following characteristics: (a) years teaching, degrees, and certifications, (b) teaching mode as described by the teacher participants, and (c) teaching mode relative to writing instruction as defined by the teacher participants. The researcher determined the participants had not used the treatment previously in their classes. When taken as groups, the characteristics of both teacher groups were similar (see Table 2) with the exception that the experimental group had fewer years teaching experience.

**Table 2**

*Teacher Characteristics (Control = O; Experimental = X) Based on Researcher's Interviews, Teacher Self-Reports, and Personnel Records*

Teacher	Degree status	Years of experience <sup>a</sup>	Teaching mode (self reported)	Use strategies prior to study?
A (O)	B.A. Sociology & Elem. Edu. +20 credits	30 total 24 in 4 <sup>th</sup>	Presentational	No
B (O)	B.S. Elem. Edu. +30 credits	17 total 11 in 4 <sup>th</sup>	Presentational	No
C (O)	B.A. Psychology & Elem. Edu.	1 total 1 in 4 <sup>th</sup>	Presentational	No
D (X)	B.A. Psychology & Liberal Arts +15 credits	10 total 10 in 4 <sup>th</sup>	Presentational	No
E (X)	B.A. Biology & Elem. Edu.	2 total 2 in 4 <sup>th</sup>	Presentational and Environmental	No

<sup>a</sup> Average years of teaching: O-Group = 15.9 years; X-group = 6 years. Average years teaching fourth grade: O-Group = 11.9 years; X-group = 6 years.

All the teachers described their teaching modes in similar ways. The researcher categorized their teaching modes according to Hillocks' (1981) four modes of instruction: (a) Presentational, (b) environmental, (c) natural process, and (d) individualized. Based on the teachers' self-descriptions of their teaching behaviors, the researcher characterized their styles as presentation mode (Hillocks, 1986) except for X-group teacher E who described a presentation mode with aspects of environmental mode. For example, all of the teachers used the words "structured" or "teacher directed" when asked to describe their overall teaching styles. The teachers used phrases such as "step-by-step" and "follow a format" and the word "structured" when asked to describe how they taught writing.

The presentational mode of instruction is characterized by specific objectives, teacher-led lectures, the study of models or exemplars, teacher

generated writing assignments, and feedback from teachers following a written experience. The teacher dominates the talk in the classroom and directs all operations. Students are expected to follow the model presented by the teacher.

The environmental mode of instruction facilitates student reflection and self-assessment. Hillocks (1986) wrote:

(The environmental mode) brings teacher, student, and materials more nearly into balance and, in effect, takes advantage of all resources of the classroom. In this mode, the instructor plans and uses activities which result in high levels of student interaction concerning particular problems parallel to those they encounter in certain kinds of writing, e.g., generating criteria and examples to develop extended definitions of concepts or generating arguable assertions from appropriate data and predicting and countering opposing arguments. (p. 247)

### *Student Characteristics and Comparisons*

The researcher compared the characteristics of the students in the experimental and control groups based on: (a) Scores from the verbal section of the Cognitive Abilities Test, (b) eligibility for free or reduced lunch, and (c) eligibility for the school's gifted and talented programs. Taken intact and excluding the scores of the learning disabled students receiving special education services, the students in both groups were similar in their overall characteristics (see Table 3). This was expected because in the participating district the students were assigned to heterogeneous classes with equal numbers of high, medium, and low achieving students in each class.

**Table 3**

#### *Student Characteristics of the Control (O) and Experimental (X) Groups*

Teacher / Class	Group	<i>n</i> <sup>a</sup>	COGAT Verbal <sup>b</sup>	# F/R Lunch <sup>c</sup>	% F/R Lunch <sup>c</sup>	# Gifted/ Talented
A	O	20	106.1	3	15	0
B	O	21	109.0	4	19	2
C	O	21	103.3	2	9.5	0
Average	O		106.1		14.5	
D	X	17	104.3	3	17.6	1
E	X	19	108.2	3	15.7	1
Average	X		106.3		16.6	
Total		98		15		4

**Table 3** (continued)

<sup>a</sup>Total *n* including learning-disabled students receiving special education were, by classes: A) 22, B) 22, C) 22, D) 22, and E) 19.

<sup>b</sup> Mean score for the COGAT Verbal section is 100 and standard deviation is 15. The researcher excluded the scores of learning-disabled students who received special education services ( $X=4$ ;  $O=5$ ) here and in study outcomes.

<sup>c</sup> Free or Reduced lunch is abbreviated F/R for this table.

### *Design*

The researcher examined the differences in the frequencies of student responses to a narrative writing picture-prompt, rubric-scored, assessment from the group of students whose teachers received the PD treatment and from the students whose teachers did not receive the treatment. An experimental posttest only control design was used because the study occurred in an ongoing educational environment with intact classes. This design controlled efficiently for threats to validity and sources of bias (Campbell & Stanley, 1963). The small number of teachers did, however, pose a study limitation. Quantitative and qualitative steps were employed. The researcher used the non-parametric chi-square ( $X^2$ ) to analyze student-outcome data. Parametric steps included the t-test and Levene's Test for Equality of Variances.

Qualitative steps included interviews of all participants and classroom observations of the experimental-group teachers. Each experimental-group teacher was interviewed prior to beginning the study, using a prepared questionnaire, to determine his or her teaching style related to writing. Table 2 provides an overview of the teachers' characteristics based on the researcher's interviews with the teachers and access to their personnel files. The researcher conducted one observation of each teacher in the X-group during the tenth week of the study. The researcher observed the teachers implementing the content they learned during the staff development sessions. The observations occurred within three days of each other. The researcher observed that the teachers' instructional behaviors were congruent with the environmental mode of instruction.

## **Findings**

### *Changes in Teachers' Instructional Mode*

By the end of the study, the teachers in the experimental group taught differently from when they started and from the control group. The teaching styles of the experimental-group teachers changed from presentational mode

to the environmental mode (Hillocks, 1986) by the conclusion of this study. Both teachers structured the communication of their lessons similarly to that used by the researcher during the staff development.

For example, each of the lessons observed by the researcher began with the teacher interacting with the entire class and disseminating information. The communication was one-way with the teacher acting as the sender of information. Then activities occurred in both lessons in which students worked in small groups or pairs. The teachers conducted two-way small group and individual communication. Both teachers provided demonstrations of strategies and skills when appropriate. The demonstrations occurred with the small groups and individuals. The teachers encouraged students to share strategies and information. The researcher observed each of the X-group teachers using inductive questioning strategies to review the concepts of self-assessment instead of giving directives. The teachers provided wait-time for the students to reflect and recall strategies. They facilitated student-led reviews of the self-assessment strategies. Each teacher's questioning review lasted between seven and ten minutes. Then each teacher implemented a 'think-pair-share' strategy to facilitate the student-led revision sessions. Both teachers used similar phrases and terminology to facilitate the activity. Teacher D said, "Now you have to decide what changes to make to your work. Think about the types of (reflective) questions you are going to ask yourself when revising." Teacher E said, "You need to share with your partner the types of questions you might ask yourself as you revise your papers." The teachers' terminology and methods of structuring the activities were consistent with the environmental mode of instruction.

Teacher D said on several occasions, "Tell your neighbor your idea." This appeared to be an attempt by Teacher D to facilitate student-led diffusion of information. The teachers focused their efforts on individual students who appeared to have difficulty implementing self-assessment strategies. They provided individual instruction when needed. Both teachers closed their lessons by interacting with the entire class.

Each teacher kept a reflective log during the study. Teacher responses in the reflective logs indicated a shift in instructional thought and attitudes away from the presentational mode and toward the environmental mode. Teacher D wrote, "The teaching of reflective questioning allows a writer to take control of his own writing. The 'old' way of handing back a paper, or even conferencing with a student, and telling him what needs to be fixed doesn't (help the student) internalize why it needs to be corrected. It's done because the teacher says it should be."

Comments from Teacher E included aspects of teacher self-assessment and reflection. Teacher E wrote:

My awareness of how children learn is evolving. I recognize the importance of understanding and utilizing new strategies and techniques to address the needs of the students.... In the past, my instruction tended to be linear and concentrated on direct product outcomes.... Though students wrote more frequently (before the staff development training) they were generally unable to make significant revisions to their own work or the work of their peers. Their work was brought to "showcase" final product as a result of an editing conference with an adult. I believed that speaking with students about their work would help them improve it. I did not recognize that this process usurped students' ownership of their writing, nor did I recognize that I was neglecting to provide students with opportunities to develop the skills necessary to the revision process.... The control I exerted over my students' writing interfered with student development of metacognitive or self-directed abilities.

The researcher observed the use of an instructional mode, congruent to the teachers' reflective comments, during in-class observations completed for this study. Formal data were not collected to be included in the teachers' professional folders as the researcher and teachers agreed that classroom observations would be informal and not evaluative.

Caulfield-Sloan (2001) stated that staff development directly influenced the instructional practices of the teachers in her study. The shift in the teachers' instructional mode may demonstrate a link between PD and a change in teaching behavior in this case, although the perceived change in instructional practices may be influenced by factors not accounted for in this study. The findings suggest that PD can be a factor used to change teachers' instructional behaviors when implemented in a deliberate and planned manner. This empirical conclusion supports teacher self-reported results from other studies on the topic (e.g., Garet et al., 2001). Change in teacher behavior can lead to a change in student outcomes. In this case the change was student learning and improved narrative writing as measured by the New Jersey Registered Holistic Scoring Rubric.

### *Student Self-Assessment Inter-Rater Agreement*

Students in the experimental group self-assessed three of their writing pieces during the study. The teachers also scored each paper. An inter-rater agreement average was calculated by dividing the number of papers to which the teachers and students awarded the same score by the total number of papers scored by the teachers. The initial scoring session produced an agreement of 62% between the teachers and students. By the third scoring

attempt, students in the experimental groups achieved agreement with their teachers at a rate of 72%, a positive gain.

### *Student Writing Scores and Frequencies in Proficiency Categories*

The literature suggested that teaching students to use scoring rubrics and reflective questions should impact their ability to enhance the content/organization of their writing but have little if any effect on the mechanics. The teachers assessed the students from both groups at the conclusion of the 14-week study. The assessment was a narrative writing assignment based on a picture prompt. An independent test scoring company scored the writing samples and reported the results to the researcher. Content/organization scores and frequencies of both groups were compared. The content/organization scores were based on the New Jersey Registered Holistic Scoring Rubric. The rubric contains six score points ranging from 1-6 and is segmented into four categories: (a) Content/organization, (b) usage, (c) sentence structure, and (d) mechanics. A score can be assigned for each of the four sections. The independent scoring company contracted to score the writing samples in this study assigned scores of 1-6 relative to the content/organization section of the rubric based on the criteria found on the rubric.

The mean content/organization scores of the experimental group ( $n=36$ ) was 3.39 with a standard deviation of .99 and the mean of the control group ( $n=62$ ) was 2.97 with a standard deviation of 1.04. The researcher conducted a test for homogeneity of variances because the groups were not the same size. Levene's Test for Equality of Variances produced an F score of .229, providing a non-significant ( $p<.63$ ) outcome (see Table 4). The t-test for independent samples showed the difference in means was significant ( $p<.10$ ) with 96 degrees of freedom and approached the traditional  $p<.05$  level ( $p<.052$ ).

**Table 4**

#### *Quantitative Data Output Describing Content / Organization Scores for X and O Groups*

Group	<i>n</i>	Mean	Standard deviation (SD)	t-score	<i>p</i>
X	36	3.39	.99		
O	62	2.97	1.04		
Total	98			1.964	.052

*Note.* X= Experimental Group, O = Control Group.

Effect Size = Mean of X-Group – Mean of O-Group / SD of O-Group

$3.39 - 2.97 / 1.04$  or  $.42 / 1.04 = .40$  Effect Size

However, because the class outcomes clearly could be dependent upon class composition and teacher instruction, the researcher chose as the predominant indication of student achievement the frequencies of students who achieved various proficiency levels as reported. Thus, the researcher performed a chi-square analysis to determine if the differences in score frequencies for the categories were statistically significant at  $p < .05$ . The researcher compacted the score categories from six to three for two reasons. First, the New Jersey Elementary School Proficiency Assessment reports achievement to students, parents, and school officials based on three categories, Partially Proficient, Proficient, and Advanced Proficient. Second, the sample size and six score-point categories caused several  $X^2$  cells to have fewer than the required five cases. The score frequencies appear in Table 5. Compacting six categories into three has the potential to reduce the variances between the groups and lessen the possibility of obtaining significant differences.

**Table 5**

*Frequency of Compacted\* Rubric Scores\*\* for Experimental (X) and Control (O) Groups*

Experimental group			Control group			% Difference X-O
Score	n	%	Score	n	%	
1 Partially	4	11.1	1	19	30.6	-19.5
2 Proficient	27	75	2	40	64.5	+10.5
3 Advanced	5	13.9	3	3	4.8	+9.1
Total	36	100	62	99.9		

\* To achieve the compacted frequencies, the researcher combined the non-compacted frequencies from categories 1 + 2 = 1 Partially; 3 + 4 = 2 Proficient; 5 + 6 = 3 Advanced.

\*\* Scores represent content and organization aspects of the scoring-rubric.

A value for chi-square of 6.354 at two degrees of freedom ( $df$ ) was calculated. This value exceeded the critical value for chi-square of 5.991 and was statistically significant ( $p < .05$ )  $p = .042$  indicating that the differences in the frequencies between the groups were not by chance. Table 6 summarizes the results of the chi-square analysis. The null hypothesis is rejected. An effect size value of .40 was computed using the means and standard deviation

shown in Table 4.

**Table 6**

*Chi-square ( $X^2$ ) Analysis of Compacted Rubric Scores from the X and O Groups*

Observed and expected frequencies	Frequencies by category <sup>a</sup> (1-3)			Total row frequencies
	Category 1	Category 2	Category 3	
Observed X	4	27	5	36
Expected X	8.4	24.6	2.9	
Observed O	19	40	3	62
Expected O	14.6	42.4	5.1	
Total column	23	67	8	98

Note.  $X^2=6.354$   $df=2$   $p<.05$   $p=.042$  ( $p\leq 5.991$  for 2  $df$ )

<sup>a</sup> Categories were 1 = Partially Proficient, 2 = Proficient, and 3 = Advanced Proficient.

## Data Analysis

### *PD and Student Achievement*

Study results indicated a statistically significant difference ( $p<.05$ ) in the frequency of rubric scores favoring the experimental over the control group. Students in classes taught by the experimental-group teachers performed better on the narrative writing assessment than did students taught by the teachers in the control-group. Findings in this study parallel Caulfield-Sloan's (2001) findings that students taught by teachers trained to ask higher-order questions performed significantly better on an open-ended science question assessment when compared with students taught by teachers who did not receive the staff development. The PD used in the present study and in the Caulfield-Sloan study influenced student outcomes positively.

### *Staff Development Group Size*

The size of the staff development group could have been a positive factor in this training model. The experimental-group teachers received PD in a small group ( $n=2$ ). Teachers, and not just students, may benefit from structured, small-group, and individual instruction. An important aspect of the com-

munication/change model used in this study was two-way, small-group and individual communication. Based on this theoretic issue, and the continuing lack of substantive empirical findings on the effect of other PD delivery modes on teacher behavior and on student outcomes, one could conclude that small group size for PD delivery, as in this study and Caulfield-Sloan's, is one factor that enables teacher change and improved student outcomes.

### *Student Self-Assessment*

The students were able to apply the criteria taught by the teachers to their writing and make positive revisions to their work. The level of student/teacher agreement achieved by the experimental-group students by the end of this study demonstrated that the students responded to the writing instruction, were able to think critically about their work, and made positive revisions. The students and teachers agreed about the criteria of quality writing. Students characterized by the school district as average or below were able to self-assess, reflect, internalize, and apply the criteria.

### **Implications**

This study is relevant to district-level staff development planners and educators who must provide high quality, effective PD and evaluate its impact on student achievement. The research surrounding the ability of PD programs to impact student achievement positively as measured traditionally has been inconclusive. The results of this study suggest that carefully planned, small-group PD implemented with an awareness of change processes can be empirically tested and (a) change the instructional practices of teachers and (b) positively impact student achievement. Those responsible for planning PD may be assisted in designing and conducting effective activities by being aware of and using change models and theories and using research to validate observable and measurable benefits. Effective PD should be planned and implemented according to research related to change processes, participant group size, and proven instructional practices.

### **References**

Achilles, C.M. (1986, November). *New perspectives on change and the change process: Teaching and learning as viewed through communication elements and change processes*. Mimeo. Paper presented to Administrators and Supervisors Council, Milwaukee, WI.

- Achilles, C.M., Dickerson, C., Dockery-Runkel, L., Egelson, P., & Epstein, M. (1992). *Practical school improvement: The Mary Reynolds Babcock Project at Moore School*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Achilles, C.M., Reynolds, J.S., & Achilles, S.H. (1997). *Problem analysis: Responding to school complexity*. Larchmont, NY: Eye on Education.
- Berman, P., & McLaughlin, M. (1974). *Federal programs supporting educational change: Vol. 1. A model of educational change*. Santa Monica, CA: Rand Corporation.
- Bloom, B. (Ed.). (1956). *Taxonomy of educational objectives: The classification of educational goals*. (Handbook I). New York; Toronto: Longmans, Green.
- Campbell, D.T., & Stanley, J.C. (1963). Experimental and quasi-experimental designs for research on teaching. In N.L. Gage (Ed.), *Handbook of research on teaching* (pp. 171-246). Chicago: Rand McNally.
- Caulfield-Sloan, M. (2001). *The effects of staff development of teachers in the use of higher order questioning strategies on third grade students' rubric science assessment*. Unpublished doctoral dissertation, Seton Hall University, South Orange, NJ.
- Cohen, D.K., & Hill, H.C., (1998). *Instructional policy and classroom performance: The mathematics reform in California*. (CPRE Research Report Series RR-39). Philadelphia, PA: University of Pennsylvania, Consortium for Policy Research in Education.
- Cros, F. (1998). L'innovation en education et formation: vers la construction d'un objet de recherché? In Organization for Economic Co-operation and Development (Ed.), *Innovating Schools* (pp. 55-79). Centre for Educational Research and Innovation, France.
- Garet, M.S., Porter, A.C., Desimone, L., Birman, B.F., & Yoon, K.S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38(4), 915-945.
- Harwell, M., D'Amico, L., Stein, M., & Gatti, G. (2000). *The effects of teachers' professional development on student achievement in community school district #2*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Hillocks, G. (1986). *Research on written composition*. Urbana, IL: ERIC Clearinghouse on Reading and Communication Skills, National Institute on Education.
- Joyce, B., & Showers, J. (1983). *Power in staff development through research on training*. Arlington, VA: Association for Supervision and Curriculum Development.

- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: Author.
- Rogers, E.M. (1962). *Diffusion of innovations*. New York: The Free Press.
- Rogers, E.M., & Shoemaker, F. (1971). *Communication of innovations*. New York: The Free Press.
- Wood, F., & Thompson, S. (1993). Assumptions about staff development based on research and best practice. *Journal of Staff Development*, 14(4) 52-57.
- Yankelovich, D. (1991). *Coming to public judgment: Making democracy work in a complex world*. Syracuse, NY: Syracuse University Press.

**Christopher H. Tienken is a K-12 Director of Curriculum and Instruction in the Plumsted Township School District, New Egypt, New Jersey.**

**Charles M. Achilles is a half-time professor in the Department of Educational Leadership, Management and Policy at Seton Hall University, South Orange, New Jersey; and a half-time professor in the Department of Educational Leadership at Eastern Michigan University, Ypsilanti, Michigan.**