The common assumption that teacher evaluation has a positive impact on the improvement of teaching is challenged. Critical challenges to the assumed contribution of teacher evaluation to learning originate from at least three areas of scholarly work: (1) revisionist/Marxist literature, which describes educational structures including teacher evaluation as mechanisms for reproducing the social and economic class distinctions of society; (2) literature on educational organizations and teaching, which examines relationships between the management and service sectors; and (3) literature on the educational and social factors that contribute to the development of creativity. While teacher evaluation may promote teacher behavior that is oriented toward organizational goals, there is little evidence that it contributes to an improvement in learning. It appears that teacher evaluation may, indeed, be detrimental to creative teaching. (TJH)
CREATIVITY/EVALUATION

TEACHER EVALUATION: REINFORCING MECHANICAL INSTRUCTION

Miles T. Bryant  
Dept. of Educational Administration  
Seaton Hall 1209  
University of Nebraska at Lincoln  
Lincoln, NE 68588-0638  
Fall 1987

Paper presented at the annual meeting of the American Evaluation Association, October 15-17, 1987, Boston, MA.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY Miles Bryant  
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

2
The common assumption that teacher evaluation has a positive impact on the improvement of teaching is challenged. Marxist critiques of schooling and selected organizational theory are two literatures used as a base for questioning the efficacy of teacher evaluation practices. Then recent work on the impact of evaluation on creativity is utilized to extend the challenge to teacher evaluation proponents. The author argues that teacher evaluation may be inherently detrimental to creative teaching.
What's really angering about instructions of this sort is that they imply there's only one way to put this rotisserie together—their way. And that presumption wipes out all of the creativity. Actually there are hundreds of ways to put the rotisserie together and when they make you follow just one way without showing you the overall problem the instructions become hard to follow in such a way as not to make mistakes. You lose feeling for the work. And not only that, it's very unlikely that they have told you the best way. (R. Pirsig, *Zen and the Art of Motorcycle Maintenance*, p. 147)

The Basis Challenging to Teacher Evaluation

What Pirsig says about packaged instructions for handypersons could easily be applied to teacher evaluation. In spite of whatever good intentions exist to guide the practice, there is a logical tendency to prescribe and to limit diversity of approach in favor of a particular way of evaluating teachers and teaching. But, educators often seem indifferent to the possible consequences of such managerial mechanisms as teacher evaluation.

"There seems little need to offer an extensive justification for the existence of teacher evaluation. Among educators it is, in fact, one of the few areas in which there is agreement." This sentence opens a book on teacher evaluation by Thomas McGreal that has been widely
distributed to approximately 75,000 administrators and supervisors about the country. Bolton cited by McGreal in the opening paragraph of this book offers the common shibboleth that the purpose of evaluation is "to safeguard and improve the quality of instruction received by students."

But while teacher evaluation may promote teacher behavior which is oriented toward organizational goals, there is little evidence that it contributes to an improvement in learning. In fact, it may well be an obstacle to the very ends it seeks to serve. McGreal's position is conservative at best. To hold that because educators agree on the need for teacher evaluation there is no need of further examination into its impact approximates a form of sophistry. Furthermore, the assumption that teacher evaluation is to safeguard and improve instruction is one needing examination.

Wendell Berry writes:

If critical intelligence has a use, it is to prevent the coagulation of opinion in social and political cliques. (W. Berry, Disciplines and Hope)

Uncritical acceptance of educational practices yields that very coagulation of opinion that troubles Berry. McGreal's assumption that wide acceptance obviates criticism...
illustrates this coagulation of opinion. "Any field that seeks to make programmatic and conceptual headway must stand open to even quite basic criticism and change if is to me more than a pretender to rationality." Because we understand the degree to which our knowledge is bound in particular paradigms, we are even more in need of a critical evaluation of educational practices.

Challenges to Teacher Evaluation

Critical challenges to the assumed contribution of teacher evaluation to learning come from at least three areas of scholarly work:

1) a revisionist/ Marxist literature which describes educational structures including teacher evaluation as mechanisms for reproducing the social and economic class distinctions of society;

2) a literature on educational organizations and teaching which examines relationships between the management and service sectors;

3) a literature on the factors which contribute to the development of creativity.

A brief summary of how each of the first two literatures can be construed as a challenge to teacher evaluation follows. Then the literature on the educational and social/psychological factors relating to evaluation that promote or impede creativity in schooling will be explored. The author argues that the likely negative impact of
evaluation on creativity in the learning process provides a compelling challenge to proponents who would expand the evaluation of teaching.

The Revisionist/Marxist Challenge

The common portrayal of teacher evaluation as a mechanism for safeguarding teaching and improving instruction has been implicitly challenged in various Marxist interpretations of schooling. In some cases, these challenges to the educational efficacy of teacher evaluation are direct and in some cases implied. School organizations are an outgrowth of class structure in U.S. society and the control mechanisms in these organizations serve as aids in the reproduction of class distinctions. Teacher evaluation under such a Marxist lens does safeguard instruction, but it does so in order to ensure that the needs of the dominant class are met.

When teachers are closely supervised, curricular or teacher decision-making is relegated to the "how" and the "when." The "what"—the choices that are made about what educational experiences are likely to result in the most favorable learning for a child—may be determined far from the classroom when teachers are closely supervised. Teacher evaluation may "improve the quality of instruction" but it may do so by monitoring the implementation of the
state (or more commonly the district) curricula by the teaching force—an example of hegemony in action.

One of the "striking features of teaching as an occupation is its inseparability from the organizational context of the school." No longer do teachers sell their services to students independent of the state as they did in the days of the "itenerant" 18th century teacher. This monopoly held by school organizations over "teaching" is reinforced by such practices as accreditation, certification, and teacher evaluation. The latter provides for controls over teachers in order to assure that the aims and goals of the state are pursued. The concept of hegemony in Marxist literature holds that the goals of the state mirror those of the powerful classes in society.

It should be emphasized that this is no simple scenario. The means for serving the interests of a capital class are interwoven into the structure of the educational system as it has evolved. For example, during the early years of the twentieth century, schools developed "tracking" in order to meet the needs of children. However, the needs of lower class children were conceptualized as different from those of upper class children. Given different needs, a student's preparation was different. A superintendent at the turn of the century noted:

Until very recently the schools have offered
equal opportunity for all to receive one kind of education, but what will make them democratic is to provide opportunity for all to receive such education as will fit them equally well for their particular life work.

(Quoted by Bowles Schooling in a Corporate Society, Martin Carnoy (ed.) p. 46)

The schools took up the institutional burden of defining for individual students the nature of their "particular life work." Socio-economic status colored that definition heavily. Thus, class distinctions were institutionalized in the school system.

A way of increasing the probability that education was meeting these class based needs of children was to subject teachers to various efficiency scales. In fact, while teacher evaluation was conducted informally by school boards in the 19th century, it was the development of rating scales during rise of the efficiency experts when teacher evaluation by check list became common. Here was a mechanism to socialize teachers, produce compliance, and orient the learning process toward externally defined ends.

Apple provides another understanding of the dynamic whereby evaluation contributes to compliance. He suggests that a teaching workforce composed primarily of women resisted external incursions into classroom. "Women teachers were not passive in the face of the class and gender
Apple notes that teachers (largely women) via their political organization into unions "fought to have a much greater say in ...and how and by whom their work was to be evaluated." Teacher evaluation mitigates the possibility that gender (female) might have an adverse impact on the program in the schools by subjecting the teachers to administrative (largely male) supervision.

This revisionist/Marxist literature does issue a challenge to teacher evaluation. Teacher evaluation is not just a means of attempting to improve instruction. It is a means exerting control over the classroom in order to enforce a particular type of instruction with a particular type of curricular content.

Organizational Challenges

The interdisciplinary study of organizations also leads to challenges to teacher evaluation. For example, evaluation is a mechanism by which supervisors inspect and control the work of workers. From an organizational perspective, this is a legitimate function of management. Evaluation ensures that there is a coordinated understanding of organizational goals and that work is aimed at accomplishing those goals with some modicum of efficiency.

Researchers have noticed that in educational organizations the relationship between the institutional
goals of a school administration and the instructional activities of a classroom were often loosely related. This phenomenon was labeled "loose coupling." Weick, who made this concept well-known, did not define a loose connectedness between departments or organizational levels as detrimental. In fact, he suggested that loose coupling might be advantageous for the organization especially when the organization confronts a segmented and diverse environment. Because organizational levels (teachers for example) are free to vary independently as a result of their loose coupling, more adjustment to local demands is possible. To teachers of a developmental persuasion, a loose coupling of classroom teacher and administrative office was probably desirable. The attention of the teacher, freed from the demands of the district, would turn more directly to the specific needs of students.

Loose coupling became not just a descriptive term in the early phases of the present school reform movement. For some policy makers, loose coupling became something to root out of the structure of schools. Reform proponents needed to tighten the linkages between administration and classroom, a position most vividly dramatized by the creation of the principal or administrator as an instructional leader. By requiring that the principal be
extensively involved in the class and in the evaluation of teachers, the loose coupling was tightened. Getting control over the classroom and opening up the classroom door to administrative watchfulness has been one of the educational policy themes of the eighties. A significant means of attempting to obtain that control has been through teacher evaluation.

But, organizational literature suggests that under certain conditions, the particular work of an organization (its technology) needs to be protected from outside influence or the work will not be done as well as it otherwise might. This insight has particular relevance for teacher evaluation. When external forces intrude on the work process, the means carefully devised by the workers to attain ends are often upset. For those conducting the work of teaching, the increased inspection of the classroom may well have interrupted the relationship between instructional means and instructional ends. More precisely, when the determination of appropriate instructional methods and ends is made outside the classroom, the capacity of teachers to develop appropriate means and ends for students is diminished. Furthermore, if the work of teaching does not allow teachers to determine appropriate means and ends for the instruction of children at given moments in time, teaching becomes quite mechanical...
Certain kinds of organizations perform certain kinds of work. One conclusion that can be gleaned from the organizations literature is that evaluation should be unique to different organizational types. The control mechanisms exerted over the technological level (the work level) need to be thought through carefully. Evaluation models drawn for other organizational types will not likely work well for education.

The Challenge from the Work or Creativity

In 1962 a Ruth Strang reported that "in a group of high school teachers, the correlation between ingenuity test scores and overall ratings of teaching ability by principals or superiors was .38." In other words, teachers who scored highly on variables one might associate with inventiveness, openness, flexibility and creativity did not receive correspondingly high evaluations from principals. Bridges found that teachers with "independent-mindedness" as a trait were "generally antipathetic to the principal."

Such studies are illustrative of a wide-ranging exploration of creativity in education that occupied practitioners and researchers in the early sixties.
Broadening definitions of school success as America became more pluralistic allowed educators to turn their attention to developing talents in children other than skill acquisition. One form of this movement took was to emphasize the need to identify kinds of creative talent in students and to nurture it.

This theme of developing creativity in children provided much fodder for damaging critiques of educational practice. Holt's *Why Children Fail*, Postman and Weingartner's *Teaching as a Subversive Activity* and Silberman's *Crisis in the Classroom* illustrate the temper of school criticism in this era. Similar critiques continue in the present reform era. Sizer's *Horace's Compromise* and Cuban's work on the prevalence of teacher centered instruction have some kinship with these earlier criticism. Sizer writes that we should be "giving teachers and students full room to take advantage of the variety among them..." Cuban writes that the "core repertoire of teacher-centered instructional practices finds students listening to lectures, completing worksheets or homework at their desks, reciting from textbooks, and seldom asking questions. Such work requires little application of concepts, little imagination, and little serious inquiry."

A continuing criticism has been that schooling in
America stifles the natural learning impulses of youth. In brief, it is alleged that many public classrooms are not and have not been places where creative learning occurs. While there are numerous reasons advanced to explain such an uninspired approach to learning, one contributing factor may logically be the disincentives that are given teachers via the evaluation system.

Before assessing the possible contribution of teacher evaluation to the diminishment of creative teaching, it is germane to suggest some of the general conditions that suppress creativity in schools. The literature on creative development from the early sixties provides some indication of how expanded evaluation may hinder creativity.

For example, scholars argued that the "success-orientation" of American public education undermined substantive challenge in school classrooms. Students are given only that which will lead to success. Failure is viewed as a condition that will lead to defeat and a poor self-image. Thus, teachers find encouragement in the system to create minimal challenges for students that have a high probability of leading to failure.

In turn, much of the literature on the development of creativity underscores the essential element of risk taking if the "new" and "different" is to emerge. Indeed, from one philosophical point of view, risk is associated with
personal growth. Kierkegaard held that "to dare is to lose one's footing momentarily. To not dare is to lose oneself." J. Adams writes:

Most of us have grown up rewarded when we produce the 'right' answer and punished if we make a mistake....and are taught....to avoid risk whenever possible.  
(James Adams, Conceptual Blockbusting, p.53)

In schools, children are sheltered from risk. This success-orientation remains perhaps even more pervasive today than it was in the sixties.

In fact, there is cause to argue that the tendency of the public school teacher to orchestrate success has grown. Some popular new teacher evaluation programs are oriented strongly toward the psychology that the learning experience must be geared to the developmental level of the child, i.e. guarantee the child success in small, digestible steps. Those readers familiar with formative evaluation programs will recognize this emphasis in such phrases as "learner readiness," "correct level of difficulty," and "checking for understanding." It is one thing to inform one's knowledge of teaching with models of instruction from which one might choose. However, it is another to push teachers toward programs that ensure that each learning experience is a successful one. As a consequence, the risk-taking character
associated with the development of creative responses (on the part of teacher or child) is diminished.

Other inhibitors of creative development suggested by proponents of creative teaching are peer-orientation which produces pressure for conformity, the push for on task behavior which diminishes exploratory questions, and the cultural labeling of schooling as work which diminishes the playful aspect of creativity.

Educational critics have held that such socially derived values delimit creative development because these values are transmitted from the larger society through the school. Scholars have not, however, paid much attention to the impacts of the institutionally derived practices that might influence creativity in schooling. Teacher evaluation is one of those practices.

The work of Teresa Amabile tests the broad hypothesis that in addition to the personal characteristics of individuals there are social/psychological conditions that impede or are conducive to creativity. Those social conditions that exert an external motivational force on the creative activity tend to diminish value and outcome. Permissive conditions that release internal motivational forces tend to yield a more creative outcome.

Amabile treats evaluation appropriately as a form of
external motivation. She notes that there is little "empirical research on the effects of external evaluation on creativity." And what work there is does not distinguish between the type of task that is being evaluated, a necessary dimension in Amabile's construct.

Accordingly, two types of tasks are defined:

1) The algorithmic task is one in which the path to a solution is "clear and straightforward." Some portion of the work that teachers do with children in schools would fit into this category. Children are asked to do straightforward tasks, algorithmic in nature—"complete the math exercises and check your answers;" "write a paragraph and underline the topic sentence." Or, in November when the second grader is given a cut-out of a turkey and told to color the bird, the task is algorithmic—i.e., get the crayons and color in between the lines. Many classroom tasks are externally imposed and algorithmic in nature.

2) The heuristic task is one for which algorithms have yet to be developed. The goals of the task may be clear or unclear but the pathway to its solution is not. "Write a poem;" "design your own Utopian society." Or, in November when the second grader is given a pile of found objects and told to create a three dimensional piece that resembles a turkey, the task is heuristic. Many tasks are externally imposed.
but heuristic in nature.

Amabile presents research evidence from McGraw supporting the notion that evaluation or extrinsic motivation "enhances performance on algorithmic tasks." There is also evidence holding that when the task is heuristic, extrinsic motivation diminishes the creative performance. In a number of experiments, Amabile has created experimental settings in which subjects produced "works of art" (often collages) under different conditions, i.e. evaluation conditions which ranged from an experimental group given no clue that the work would be judged, to another group expecting an evaluation of technical merit, to a third group expecting an evaluation of creative worth. For the most part, work done in these types of experimental settings was judged (by judges with interrater reliability) as more creative when there is no expectation of evaluation. Thus, for heuristic tasks, the expectation of evaluation resulted in less creative output.

Amabile reports a related strand of research examining the impact of reward on intrinsic motivation. Since the manipulation of rewards to motivate student work habits has gained much credence in recent years, this research demands serious attention. Amabile finds that the offering of a reward tends to decrease intrinsic interest in the
task. The task becomes perceived of not as something intrinsically worth pursuit but as an end to the reward. Engagement in the activity diminishes because the activity has been devalued. It should be noted that teachers are encouraged utilize some clear and obvious method of manipulating rewards so that students know exactly what is and is not appropriate behavior.

One cannot help but conclude that if even a small portion of Amabile's work proves generalizable to educational settings the growing use of external evaluation will diminish creativity in the teacher and in the classroom. There is no question that some portion of classroom teaching needs to present children with algorithmic tasks. Cognitive growth and mastery requires that some work will lead students toward known solutions along known solution paths. In Bloom's taxonomy, these tasks are characterized by the two lesser levels of cognition: comprehension, application. But much of teaching and learning needs to be heuristic if children are to develop those very qualities of problem solving and creativity currently in high demand.

To the degree that schools increase the evaluation of instruction, they push teachers and children toward a learning dynamic that is algorithmic in nature. There are consequences implied in such a direction. Learning becomes
work; there is little joy in the process since the discovery factor has been removed; all students get to the desired ends in quite similar ways thus enforcing a conformity that is far from natural or probably healthy; significant portions of the student population become disenchanted with school; and probably most significantly, students are not provided with abstract models by which they may reflect critically on their world.

One can not lay the shortcomings of education upon the shoulders of teacher evaluation. However, because it is an entrenched practice, there is a need to know what happens to different kinds of teaching as a result of evaluation. The growing recognition that teachers must be part of the decision-making process and must be empowered may be the harbinger of a new collaborative organizational structure where evaluation is used with discretion. There is even a need to know if anything awful might happen to teaching if evaluation were discarded in the cases of engaged, competent, and successful teachers.

"To help somebody learn to think," speaks Grady Cassidy, a character in Gwaltney's *The Dissenter*, "is to lose your power to make that person a knife of your own will." Teacher evaluation may have the ultimate purpose of forcing teachers to socialize students in particular ways,
not to help students "learn to think." Yet, if Americans truly want public education to help all students learn to think, we need to do something differently than we are doing. One action we can take is to examine the consequences of school structure.
FOOTNOTES


2 Officials at the Association for Supervision and Curriculum Development estimate that since 1983 about 75,000 copies of McGreal's book have been disseminated to school administrators and curriculum directors. To put this number in some perspective, there are around 15,000 school districts large and small in the United States.

3 Op. Cit., p. vii

4 Stallings and Krasavage report weak effects of clinical supervision on instructional practice (1985); Smith (1986) and Slavin (1987) find no documentation that clinical supervision, a popular formative teacher evaluation model, improves instructional practices of teachers.


10 There are other interpretations which provide indirect understandings of the reasons and consequences of close teacher supervision. David Tyack's *One Best System* and Raymond Callahan's *Education and the Cult of Efficiency* are excellent examples of historiographies that can be used to place teacher evaluation within the ideological frameworks that shaped educational change.


12 See Michael Apple, *Teachers and Texts*, for a close examination of how class and state functions interact.


15 Ibid.


17 Ibid, p. 76.


22 An example of the negative impact resulting from the close scrutiny of a teacher's competence is reported in a
CREATIVITY/EVALUATION


28 Some would argue that teachers dismiss the teacher evaluation systems as an empty ritual. However, given the increased administrative attention to evaluation the inconsequential nature of teacher evaluation can no longer be assumed.

29 Calvin Taylor, Creativity: Progress and Potential, p. 98.


31 A review of the Hunter Model or the Direct/Explicit Teaching Model will provide many examples of the degree to which the success-orientation guides approved contemporary pedagogy.

32 "What's Noteworthy on Teaching," a recent publication from the Mid-Continent Regional Education Laboratory, lists seven models of instruction: 1) Learning to Learn; 2) Active Mathematics Teaching; 3) Beginning Teacher Evaluation Study; 4) Explicit Teaching; 5) the 4Mat Model; 6) The Hunter Model; and 7) Mastery Learning.

CREATIVITY/EVALUATION


35 Ibid., p. 5.

36 Ibid., p. 100.

37 Ibid., p. 33.

38 Ibid., p. 33.

39 Ibid., p. 103.

40 Ibid., pps. 103-117. Amabile operationalizes the concept of creativity with the following definition: "A responses will be judged creative to the extent that a) it is both a novel and appropriate, useful, correct or valuable to the task at hand, and b) the task is heuristic rather than algorithmic." (p. 33)

41 Lee Canter's *Assertive Discipline*, a popular approach to pupil management and control, illustrates this point.

42 Interest in reversing the conception of the teacher's role as the mechanical engineer of a pre-determined classroom experience appears in reports such as the Homes Group Report and the Carnegie Commission Report on Teaching. See also B. Joyce and B. Showers, "Power In Staff Development Through Research on Training," Alexandria, VA: ASCD, 1983. This paper presents a strong argument for empowerment through knowledge and skill acquisition.

43 See for example Susan Ohanian's article on basal readers in *The Atlantic*, Sept., 1983. In the article, Ohanian writes that textbook publishers "tend to cut and trim and simplify, to tame and domesticate what is powerful, florid, and wild in the way writers use language." The same may well be true for formative teacher evaluation programs.