

# **Politics of Teacher Evaluation<sup>1</sup>**

**Gene V Glass**  
**Arizona State University**

**Barbara A. Martinez**  
**California State University, Los Angeles**

**Abstract.** We sketch the current practice of teacher evaluation in the U.S. from our ordinary experience. We present a fragment of a case study that reveals how teachers—elementary school (K-8) teachers in this instance—view the experience of being evaluated, and how feelings of illegitimacy flow from the experience. Finally, we indicate certain considerations from political theory that bear on the problems of improving teacher evaluation.

## **The Current State of Things**

Elementary school teachers are evaluated differently from how secondary school teachers are evaluated; and secondary school teachers are evaluated differently from college professors, who further underline the differences between themselves and their public school colleagues by not even wishing to be called teachers. Nothing seems to account for these differences so clearly as does what we might loosely refer to as the politics of evaluation. We often learn something interesting about the organization and politics of education when we contrast how it is pursued at its different levels.

College professors are usually evaluated by their peers and superiors yearly for raises and less often for promotion; but in spite of what might be claimed (by the president of the college when addressing parents, for an instance), they are seldom evaluated qua teachers. It is common today for students to fill out simple forms, rating scales, at the end of a semester: "Instructor was organized," "Instructor knew the subject," "Instructor graded fairly." Typically, these student ratings count for little. The better the university, the less teaching is weighed in the balance that sways toward research and publication; and most colleges aspire to be like the better universities. Extraordinarily bad student ratings will be used to terminate an untenured faculty member if that person's research is poor; but the administration will swallow bad ratings when a strong researcher receives them.

Secondary school teachers are evaluated sporadically. Peer evaluation is non-existent in America's schools, and administrators seldom venture into a high-school teacher's classroom. Their presence would be viewed with suspicion by the teacher; the legitimacy of their place there would be questioned (silently or behind closed doors if not publicly). Administrators appear to concede that secondary school teaching involves specialized knowledge (of chemistry or mathematics), and that a specialist may be needed to recognize good teaching.

Elementary school teachers are treated substantially differently. Principals visit their class once or more each year. Indeed, principals regard these visits as a responsibility of their position. Teaching is observed, occasionally a check-list is filled in; lesson plans may be

---

<sup>1</sup> Paper commissioned by the Center for Research on Accountability and Teacher Evaluation, Western Michigan University. Presented at its Cross-cutting Theory Conference, June 4-6, 1993.

inspected. Much is written in school personnel manuals about evaluating teachers on the basis of their students' achievement test scores; but the threat is an empty one that nonetheless has the power to shape instructional styles and choice of content (Glass, 1990).

The legitimacy of the elementary school principal's presence in the classroom for the purpose of evaluating the teacher is less likely to be questioned openly. Everyone knows elementary school subject matter after all, and an instructional leader is an expert in the general techniques of effective instruction, or so it is widely believed.

Each level is distinguished by a different balance of teachers' professional autonomy on the one hand, and the exercise of administrative authority in a democratic bureaucracy on the other. At the elementary level, professional autonomy is difficult to discern and administrators are seen fulfilling the dictates of the duly elected school board to insure that teachers are competently delivering instruction to the students. At the secondary school level, teachers enjoy more autonomy to structure their classes and curriculum as they judge appropriate; administrative authority is exercised seldom and usually only in crises. College teachers enjoy autonomy granted by a three hundred-year tradition of academic freedom; no administrator dares to cross the threshold of the lecture hall.

Some have sought to reform teacher evaluation by attempting to alter the balance between these two forces. Art Wise and Tamara Gendler (1990), in *The New Handbook of Teacher Evaluation*, distinguished seven purposes and separate functions of teacher evaluation: preservice, selection, certification, "beginning," tenure, merit and school improvement. There is much that can be said about the politics of each of these separate phases of teacher evaluation (most of which would center on the politics of higher education and of the labor movement). Wise, for example, focuses on licensure and recommends a state board licensure system for teachers like that for physicians and lawyers. Such licensing might confer prestige on the profession and with prestige may come autonomy. But one might wonder whether medicine is well served by doctors or justice by lawyers. John McNeil (1981), in the *Handbook of Teacher Evaluation*, acknowledged the deep conflicts that surround this phenomenon in the school, but recommended new forms that scarcely differ from established practice and that fail to separate incompatible purposes for evaluating teachers. Armiger (1981), of the New Jersey Education Association, recommended guidelines for teacher evaluation that would give teachers more power within a system still "owned" by the bureaucracy.

It is our contention that the problems with teacher evaluation do not stem principally from the conflict between professional autonomy and bureaucracy (although these forces are apparent, they can not be changed without changing the political context), but from the perceived illegitimacy of the democratic bureaucracy in which the evaluation is embedded. Our argument will benefit from a portrayal of how teacher evaluation is experienced in the work lives of teachers.

### **Teacher Evaluation in Nocam**

Nocam Elementary School is a K-8 school of 800 students located in the heart of a city of over two million people. Its attendance area is about a third Anglo and half Hispanic, with a smattering of children of many different cultures. Nocam is one of three elementary schools in the school district. It has a full-time principal and is closely linked to the Superintendent's office through the efforts of a curriculum specialist who has assisted Nocam in a major overhaul of its language arts curriculum. "Whole language" teaching, cooperative learning and non-graded organization have come to Nocam. In the course of pursuing a larger study

focused on school reform, the second author conducted numerous interviews with Nocam teachers and administrators and observed classes, board meetings and teachers meetings. On the following pages, Nocam teachers speak of the way in which their work is evaluated by their superiors.

District policy mandated that all teachers be evaluated once a year, despite the fact that there was no merit system of pay. Teacher evaluations were conducted by the district personnel director, except in the case of "new" teachers. "New" teachers, those who had been employed by the district for less than three years, were evaluated by their immediate supervisor, the school principal. Both new and veteran Nocam teachers viewed the evaluation process as "a joke," regardless of who the evaluator was. As one teacher explained,

"...some man is going to come into my classroom, who has never been in my classroom all year and evaluate me on how good a teacher I am, by [observing] a twenty minute lesson [and]checking things off? That's impossible. I couldn't evaluate my students that way."

Teachers generally shared this opinion,

"...it's a scheduled appointment, they will be in your room at 10:30 and you have to have the handbook and the detention notices and the homework notices... they want to see homework, they want to see discipline records and it has to all be clearly posted, your discipline plan and everything..."

Few of the veteran teachers were intimidated by the evaluation process; many, however, found the process coercive and demeaning. As one teacher explained,

"I never see that personnel director except when he comes into make an appointment to do the observation... and you can't talk to him then. And then the next time that I see him is when he is handing me back my evaluation. And the thing is so arbitrary...it's 'you're an A teacher, you're a B teacher, you're a C teacher, and you fail.' I don't need to know if I'm an A teacher or a B teacher, I don't care. I care about whether or not I'm improving. They don't have enough respect for me...the amount of work that I have done and the amount of dollars that I have put in...to give me something that would actually help me improve. Instead they give me something that makes me work first of all to put something together for [the evaluator] to keep, then they want to evaluate my classes."

Few teachers found the evaluation process informative or instructive. Teachers complained that the process failed to provide them with any insights as to how to improve their teaching. Frustrated by the procedure, teachers did not feel that the exercise was meant to help them improve their teaching,

"...it's not meant to improve [teaching] although that's the letter of the law, that the teacher evaluation systems are to improve teachers...to improve instruction. But it does not do that. In fact if it does anything I think it doesn't improve my instruction because I'm ticked for two days before I have to do it and I'm ticked for two days after I have to do it."

"New" teachers were a bit more anxious about the process, perhaps because they were evaluated by their principals. Like the veteran teachers, however, new teachers found the evaluation process much more burdensome than helpful. The evaluation experience described by Ms. Clark, a "new" K-3 Project teacher, was not unusual:

"[The principal] kept saying ahead of time 'don't worry, I really need to see what's going on in your class'. I thought okey. And it was a time when I thought it was a good lesson, except for one kid, and this kid has been documented sexually abused, well he pooped in his pants during the lesson. The principal got hysterical with me, he was like 'I spent forty minutes in this room and I've seen nothing of value happening', and he left. That was my first evaluation. ... But I decided that I was going to talk to the principal [about my evaluation] and I said [to him], 'I'm a first year teacher and you can't just walk in my room and spend almost an hour and tell me nothing of value is going on. I'm not leaving until you tell me what good you saw. So he decided not to count that [evaluation] and to do it again. So for the next one I prepared the kids...."

Ms. Anderson was coached by her fellow teachers as to what this principal liked to see, the curriculum he preferred, and the practices he approved of. With this information in mind both she and her students practiced what they would do for the next evaluation:

"...we practiced, we rehearsed what we would do when [the principal] walked in and I told the kids if we did it right, we got a surprise. And [other teachers] took the worst kids [to their classes that day] so that it wouldn't be bad.

"It was awful...like the kids did these little work-sheets and they sat there and we had practiced what the work-sheets would look like, so they sat there and did them without talking. It was so awful...it was really hollow...we played the game. [The principal] told me I did a good job and I thought [to myself] 'you don't know anything.' I've learned part of the principal's game. I can do it when I have to, I've done it."

Stories about "putting on a show" for both evaluators and administrators were common among Nocom teachers:

"The kids know how to act for the administrators. We bribe them [to act a certain way] when administrators are there. Then [the administrators] leave and we go back to our real way of working and of teaching."

"...when the district people come into our class, I have to act a certain way, to put on a show.... I train the kids to act the way the administrators expect them to act...even if the way they act [and the things we do] are not developmentally appropriate."

The administrators' ability to evaluate accurately teacher's performance was questioned by many teachers. Nocom teachers were of the opinion that their administrators did not really understand the pedagogical techniques nor the theoretical underpinnings of the techniques which were the basis of the K-3 Project,

"I don't think they have any idea what [whole language or developmentally appropriate practice] look like. They think they are very supportive of whole language, but it's only as long as kids are sitting at their desks being really quiet."

"...they don't understand how children learn and they come in and what can you tell them when you talk [with them] in the classroom for five minutes. Of course it looks like chaos...but learning is going on...you have to be there for a while to really get a grip on what is happening with the children...they don't understand [developmentally appropriate practice]. [The principal] evaluates you at a desk, a file cabinet between you and the reading group, he's not listening to the kinds of questions I'm asking the children or, you, know the communication skills going on. He's watching the behavior problems, and you are always going to have some. He's counting how many crayons the kids have on the floor."

"They are not very supportive in the teaching methodology way, but more picky, and you have to do this and you have to teach from this book and you have to cover so much and you try to slip your own things in between without getting caught... You understand that [administrators] don't know much about [teaching] and you try to take it with a grain of salt."

"...there are all kinds of politics on why we are getting raked over the coals for this and that, but if a principal does not understand what you are doing in the classroom before the evaluation starts, or if they don't agree with what you are doing, how can they evaluate you fairly ..."?

In addition to their not being very well grounded in the nontraditional models used by the K-3 Project teachers, many Nocom teachers were of the opinion that administrators didn't spend enough time in classrooms to accurately assess their teaching ability:

"...they come in and what can you tell them when you talk [with them] in the classroom for five minutes...you have to be there for a while to really get a grip on what is happening with the children...."

"...I don't really think he has a perfect understanding of what it is...of what exactly it is, because he doesn't come in our classrooms and hang out for an hour or two for a few days a year."

Nocom teachers also had concerns about the evaluators' ability, or lack thereof, to provide them with practical guidance and relevant assistance:

"[The administrators] can't sit and discuss whole language theory with you... if they can't discuss the concepts with you how can they tell if what you are doing is right or wrong, and how can they help you improve upon it?"

"...sometimes the things that they ask us to do don't particularly correspond with what we are trying to accomplish [in terms of teaching]."

"...they don't understand how children learn...they don't understand developmentally appropriate practice...I can't get much guidance from them...."

"[The principal] says 'what can I do to help you', but I feel that there is nothing he can do to help me because he doesn't have any knowledge to give me...."

"I know they are busy [but] they need to spend more time in the classrooms with us. They come in once in a while, they make me nervous because they don't come in often enough for me to feel like they are friends, like I can ask them for help...I don't even know if they know what I'm doing in here."

"...They might be supportive when you are explaining [the methodology] to them, the cooperative learning and the learning centers, but then they come into our rooms and see the movement and it's 'wait a minute, you didn't say kids were going to be talking to each other and moving around the classroom, you said *cooperative learning*.'"

Nocam teachers lacked guidance and direction. Even their direct supervisors—the school principals—lacked the appropriate pedagogical theory, and therefore were not a source of assistance or guidance. As one teachers explained:

"I've really had to depend just on myself with all this. It's like I've been left out on this island, all alone... no guidance, no support, no validation... it's been pretty much a sink or swim situation...I still don't know which I'm doing..."

Nocam principals viewed their role primarily as that of "facilitator." Although these administrators encouraged their kindergarten through third grade faculty to "use whole language, cooperative learning, and developmentally appropriate practices," neither they nor the district superintendent provided teachers with concrete suggestions for implementing the techniques or improving their teaching. This task, according to the principals, was "left to the experts," who were brought in to provide in-service training throughout the school year.

The K-3 Project teachers managed to convince the district administration that it was unfair for teachers to be evaluated on their use of traditional teaching techniques when the K-3 Project relied so heavily on the use of nontraditional approaches. After three years of requesting, and largely in response to the requests of the K-3 Project coordinator, Nocam teachers were given a choice of being evaluated based on the standards of the "traditional teacher evaluation" instrument or based on the standards of a "whole language evaluation" instrument, recently developed by the district administration.

Both instruments assessed the same general categories of performance: classroom management, communication skills, instructional capabilities and materials, planning and organizational skills, compliance with school policies, and professional qualities. The whole language instrument, however, was much more extensive than the traditional instrument. The traditional evaluation instrument contained a total of five criteria per category which teachers could "exceed," "meet," or against which they could be judged "average" or "failed to meet." The whole language instrument contained twenty different criteria per category, which teachers could "exceed," "meet," or with respect to which they could be found "adequate" or "inadequate."

Technically, teachers had a choice in the evaluation matter; practically they did not. Neither the district evaluator nor the principals were adept in using the nontraditional instrument. As

a result, though a fair number of teachers requested that the new instrument be used, only one teacher was actually assessed with it. The process was described by the teacher as "a disaster."

According to this teacher, the district evaluator did not have the new forms in his evaluation package when he came to evaluate her, nor did he know what was on the forms. The teacher had to supply the evaluator with the new forms. The evaluator didn't understand the stated criteria; he had the teacher explain to him how the new criteria related to the old, so he would know what to look for during the evaluation session. The teacher described thus:

"It was just a disaster.... [The district evaluator] couldn't sit and discuss whole language theory with me.... he doesn't know a thing about developmentally appropriate practice...and cooperative learning... forget it.... he [kept] looking for my assertive discipline program. That's not my priority.... I had to explain the entire process to him, what to look for, what was appropriate and why. I'm sure he learned a lot, if he paid any attention, but for him to evaluate me, what a joke."

Whether they knew how to use the nontraditional evaluation instrument or not, the district principals avoided using it. The principal at one of the other two elementary schools in the district went as far as to tell his teachers that it was his choice which instrument was used, not theirs. And he chose to use only the traditional instrument.

One teacher decided to "check this out with the district." She was informed that the teachers did indeed have the right to choose. When she shared this information with her principal he "got upset at me for questioning his authority and he told me that he was going to talk with the district people himself. In the meantime, he used the traditional instrument to evaluate me."

The Nocam district superintendent maintained that the district allowed teachers "the freedom to use what they think is appropriate" in teaching their classes. He also believed that someone needed "to make sure that what they feel is appropriate is in line with ... our curriculum philosophy... [that it is] highly matched to what we expect kids to be tested on." A number of tools were developed to assist Nocam administrators in monitoring curriculum and instruction.

The Nocam case makes one thing clear. Even at the elementary school level where it might be expected that administrative evaluation is most defensible, it is viewed by teachers as illegitimate. Principals are seen as uninformed about curriculum and unable to spend the time to understand the circumstances of the class in such a way that they could help improve it. Bureaucratic evaluation of teaching at the secondary and college levels is seen as an affront to professional autonomy and as being even less legitimate than elementary school teacher evaluation.

It can not be argued that what was seen in Nocam is somehow an outgrowth of the special circumstances of poverty or ethnic minority culture. Similar experiences are widely spread in the American educational system and, perhaps, elsewhere. (Beery, 1992)

## **Legitimizing Teacher Evaluation**

It is our contention that the principal problem with teacher evaluation is that it is viewed as lacking legitimacy by the persons who are the object of the evaluation, the teachers themselves.

### **Consequences of Loss of Legitimacy**

Teacher evaluation viewed as illegitimate by teachers themselves generates nothing but dissembling, passivity and feelings of alienation and powerlessness. (Glass, 1990) School boards through administrators have a legitimate interest in how instruction is conducted, but it is not an overriding interest nor does it follow that their interest is served by direct participation of the principal in the evaluation of teachers.

Where legitimacy is lacking, one can expect little more than passive compliance. Is it a matter of concern that an evaluation system is imposed from the administrative hierarchy and not seen as legitimate by the teachers who are being evaluated? One line of argument answers "No." Suppose that the system imposed is so comprehensive and well designed that it encompasses most of what teachers should be expected to perform. One might argue then that it is irrelevant whether the teachers "like" what it imposes, since if they conform to its vision of what a good teacher is, they will ipso facto teach well. This argument is similar to questions debated under the topic of "teaching to the test" in educational assessment. Some maintain that if a test is good enough, then teaching to it will only result in good education. Similarly, if a teacher evaluation system is good enough (i.e., defines a good model of what a teacher is to be), then complying with it—even if that compliance is "false" or pretended in some sense—will result in the teacher being a good teacher.

The difficulty with the "teaching to the test" argument is that the kinds of test generally used in assessments are rather pale reflections of a good education. Likewise, many of the bases of teacher evaluation systems are weak or impoverished models of what good teaching is. Checklists of teaching acts or "elements" reduce teaching to a few general principles of instruction, and divert attention from concerns of curriculum. A teacher can be a good teacher under such surveillance while teaching shallow or false knowledge. Some believe that this is little concern at elementary grades since "there is no discipline" (in the academic sense) at that level. "Elementary school teachers have no discipline, they just teach"; or "teaching is their discipline." Others are shocked to hear that the teaching of reading or language or mathematics is believed by some not to raise technical and intellectual questions as complex and sophisticated as the teaching of calculus to high-school students.

### **Ways of Seeking Legitimacy**

Legitimacy can be bestowed in at least two ways: by appeals to widely accepted scientific or technological knowledge or through the appeal to the authority of legitimate political institutions or arrangements.

The attempt to legitimate the standard practice of teacher evaluation by appealing to science and technical-rationality fails for a couple of reasons. First, there is no widely respected science of teaching and learning. Common sense or practical and tacit knowledge of teaching usually succeed as well as systems that profess to be based on research. Second, most efforts to reform teacher evaluation start from an assumption that all parties with a direct interest in improved education share a consensus on what good education is. From false assumptions of consensus come technical-rational attempts to manage teachers. We begin from a different starting point. Schools are micro-political units where teachers, administrators, parents, students and even society far removed from the classroom seek to realize their interests.



These interests often conflict. Without agreement on ends, mechanical and technical solutions fail. Third, school administrators who are vested with the authority to evaluate teachers as instructors generally lack knowledge of the subject matter being taught. Their role as evaluator consequently strikes the teacher as superficial, then illegitimate. (See Scriven, 1992, for a discussion of what limits subject matter specificity does and does not place on teacher evaluation.)

An example may help illustrate how a nonspecialist's lack of subject matter knowledge can invalidate the type of evaluation that focuses on general acts of teaching. A junior high school English teacher is reviewing a lesson on nouns. He writes on the board "A noun is a Person, Place or Thing," leaving space beside each for examples. Turning to the class, he invites anyone who can illustrate a noun as the name of a person to step to the board and write. A student is congratulated for writing "singer" beside Person, as is a second student who writes "school" beside Place. The third volunteer writes "kitchen" beside Thing only to be told politely by the teacher that a kitchen is a place, not a thing. The checklist that the observer of this episode was filling out had categories only for the commonly identified important elements of teaching: previews lesson, clarifies goals, provides for active participation, reinforces correct responses, and the like. This teacher even scores points for correcting mistakes quickly. This kind of evaluation of this act of teaching misses the important point of what actually took place. One need not be a grammarian to sense that something is seriously wrong with this lesson. Of course a kitchen is a "thing," as in "I have remodeled my kitchen." And it may function as a place in other uses. Nor are Person, Place and Thing mutually exclusive and exhaustive categories. Where does the "unicorn" reside, grammatically speaking, and what about "truth"? The point is that this teacher is teaching a shallow or false point and this consideration should override all other questions. Indeed, his grasp of grammar has led him to confuse a student (probably more than one) and draw that student into a publicly embarrassing situation, where her valid understanding of language usage is labeled "wrong." Where in the evaluation of this teacher is it noted that the teacher has a responsibility to understand and continue to learn the subject being taught? Some will argue that it is impossible for a principal or a principal's deputy to know all the subject matter taught by all the teachers in the school. Indeed it is; and lacking that understanding, it is questionable to what degree the principal can serve as the guide for the teacher's efforts to become a better teacher.

If technical-rationality can not confer legitimacy on teacher evaluation, then it remains for political arrangements to do so. Modern political institutions are bureaucratic democracies, with one distinguished from the other by the balance of democracy and bureaucracy.

Appeals to the science of teaching or to technical-rational arguments about what ought to be taught and how can not hope to justify a particular form for teacher evaluation within the hierarchical bureaucracy of contemporary schools. Legitimacy for some form of teacher evaluation must be found in a new set of political arrangements that will be viewed as legitimate by the teachers themselves.

### **Seeking New Political Arrangements**

Three evaluation theorists have addressed the systemic political problems that have led to the current state of teacher evaluation.

**MacDonald.** "Democratic evaluation," as envisioned by Barry MacDonald (1974), addresses the tension between power concentration and power diffusion in liberal democracies by opting for radical power diffusion. MacDonald focuses more on the role of the evaluator than

on such aspects of the evaluation as the criteria, data and the like. He saw the evaluator as an information broker among interested parties. The evaluator stops short of making recommendations; information, ultimately owned by those from whom it is collected, is presented to those persons with legitimate interests in what is being evaluated. Decisions flow from some unspecified process of democratic discussion among interested parties.

"MacDonald's evaluation approach intentionally includes diverse interests, allows people to represent their own interests, and is based on an idea of mutual consent." House (1980, p.150) A direct, rather than a representative form of democracy, is being imagined by MacDonald. The limitations of direct democratic participation in complex, mass societies are obvious. However, one is casting a small net when the object of an evaluation is a teacher and a classroom. The range of interests to capture and bring to consensus is narrow. After all, juries reach consensus even when the stakes are high.

**Strike.** Kenneth Strike has pursued an examination of political forms and their relevance to education. He contrasts two quite different approaches to achieving democratic governance. The first is John Locke's legislative majoritarian democracy. Its operations are familiar to all; its assumptions are less obvious. Naturally free and equal humans are to be granted equal sovereignty, which is exercised by voting for representative government. With the consent of the governed, sovereignty is placed in the legislative body. Legislatures exercise sovereign control through hired managers who follow the policy direction set down by the legislative body. Exigencies will call for clearer rules and policies; in time, rules will accrue and the modern bureaucratic democracy will emerge. Citizens may not regard every rule as legitimate, but the stability of the institution rests on a wide-spread belief in the legitimacy of the process by which the representatives are first chosen and then formulate the rules. "We can see a legislature as a means to vector interests more than as a means for making and judging the merits of practical argument. Majorities may be seen as formed more by a process of combining and reconciling interests than by a process that seeks the better argument." (Strike, 1993, p. 16). Current practice in teacher evaluation is embedded in this context of legislative bureaucratic democracy. It has failed to engender among teachers a belief in its legitimacy. Most writing to date about the politics of teacher evaluation have assumed no changes in the basic nature of democratic institutions and as a result offer suggestions that tinker with the balance of power between democratic bureaucracies and professions.

An opposing view of democratic institutions grows out of the attempts of Jurgen Habermas to justify liberal democracy. Habermas argues for the legitimacy of a communitarian democracy in which social norms are justified by uncoerced argument among equals in an ideal speech community. To Habermas, a social choice is "discursively redeemed" when it has the consensus of a community of citizens and that consensus was reached in open and undominated discourse. Argument—not votes—legitimizes choices and actions for the good of the community.

Strike recognizes a Utopian character to Habermas's notion of the discursive redemption of policy choices in an ideal speech community. As a practical matter, sovereignty will have to be located in a representative body and conflicting interests "vectored" to a solution when consensus is impossible. But he can not back away from Habermas's ideas without trying to answer the question, "How might we make bureaucratic democratic institutions more Habermasian?"

**Kemmis.** While Strike may wish to cook the Habermasian omelet without breaking the Habermasian eggs, Stephen Kemmis does not hesitate to recommend the Habermasian ideal. To Kemmis, teacher evaluation would be one particular aspect of what he calls "emancipatory action research" (Carr Kemmis, 1986; Kemmis, 1993a; Kemmis, 1993b). "When schools—

teachers, students, principals and others—are forced to change on the basis of outside evaluations and the crude coercive powers of the state, however, they frequently resist, passively if not actively. And that, it seems to me, just produces still further administrative demands for surveillance, regulation and control.... I believe that the evaluation processes I have attempted to develop—as well as some of the practices associated with 'responsive', 'illuminative' and 'democratic' evaluation—did (and do) contribute to the development of less irrational, less unjust and less satisfying forms of social life. Though some of those perspectives have no particular inclination to justify themselves against the criteria of critical social theory or critical social science, in practice they do seem to offer increased opportunities for what Habermas describes as "communicative action"—action oriented towards mutual understanding and unforced agreement...." (Kemmis, 1993, pp. 46-47)

### Conclusion

Teachers view the evaluation to which they are subject as being illegitimate. They do not recognize the authority of those who perform the evaluation; they do not accept it as valid and defensible. Legitimacy can be conferred by democratizing the process of teacher evaluation, by removing it from the context of hierarchical bureaucracy in which it now resides, and by carrying it out in a new context. Some theorists offer justifications for this rearrangement of the politics of teacher evaluation. As yet, it is unclear how the reform would be played out in its essential details, e.g., who would participate in the evaluation of teachers, what information would be relevant and how it would be obtained, with whom authority would lie to call for an evaluation, and the like.

### References

- Armiger, M.L. (1981). The political realities of teacher evaluation. Chp. 16 (pp. 292-302) in Millman, J. (Ed.) *Handbook of Teacher Evaluation*. Beverly Hills, CA: SAGE Publications.
- Beery, B.F. (1992). Master teacher conceptions of relationships between teacher evaluation and excellence in teaching performance. Doctoral dissertation. Tempe: Arizona State University.
- Carr, W. Kemmis, S. (1986). *Becoming critical: education, knowledge and action research*. London: Falmer Press.
- Darling-Hammond, L., Wise, A.E. and Pease, S.R. (1983). Teacher evaluation in the organization context: A review of the literature. *Review of Educational Research*, 53, 285-328.
- Glass, G.V. (1990). Using student test scores to evaluate teachers. Chapter 14 (pp. 229-240) in Millman, J. & Darling-Hammond, L. (Eds.). *The New Handbook of Teacher Evaluation*. Newbury Park, CA: SAGE Publications.
- House, E.R. (1980). *Evaluating with validity*. Beverly Hills, CA: SAGE Publications.
- Kemmis, S. (1993a). Action research and social movement: A challenge for policy research. *Education Policy Analysis Archives*, 1(1). Retrieved May 1, 2005 from <http://epaa.asu.edu/epaa/v1n1.html> .
- Kemmis, S. (1993b). Foucault, Habermas and evaluation. *Curriculum Studies*, 1(1), 33-52.

- MacDonald, B. (1974). *Evaluation and the control of education*. Norwich, England: Centre for Applied Research in Education.
- Martinez, B.A. (1993). *How educational reform is compromised: A critical investigation*. Doctoral dissertation. Tempe: Arizona State University.
- McNeil, J.D. (1981). Politics of teacher evaluation. Chp. 15 (pp. 272-291) in Millman, J. (Ed.) *Handbook of Teacher Evaluation*. Beverly Hills, CA: SAGE Publications.
- Scriven, M. (October 1992). Should teacher evaluation be subject-matter specific? TEMP Memo #10. Kalamazoo, MI: Center for Research on Educational Accountability Teacher Evaluation, Western Michigan University.
- Scriven, M. (September, 1991). Duties of the Teacher. Memo from the Center for Research on Educational Accountability Teacher Evaluation, Western Michigan University. (Also see Scriven, M. (1988). Duty-Based Teacher Evaluation. *Journal of Personnel Evaluation in Education*.)
- Strike, K.A. (1993). Professionalism, democracy and discursive communities: Normative reflections on restructuring. *American Educational Research Journal*, 30(2), 255-275
- Strike, K.A. (1990). Is teaching a profession? How would we know? *Journal of Personnel Evaluation in Education* 4(1990), 91-117.
- Strike, K. A. (1991). The Moral Role of Education in a Liberal democratic Society. In G. Grant (Ed.), *Review of Research in Education* (Vol. 17, pp. 413-483). Washington, D.C.: American Educational Research Association. (Ed.), *Review of Research in Education* (Vol. 17, pp. 413-483). Washington, D.C.: American Educational Research Association.
- Strike, K.A. (1991). Humanizing Education: Subjective and Objective Aspects. *Studies in Philosophy of Education* 11(1), 17-30.
- Strike, K.A. (1990). The ethics of educational evaluation. Chp. 21 (pp. 356-373) in Millman, J. & Darling-Hammond, L. (Eds.). *The New Handbook of Teacher Evaluation*. Newbury Park, CA: SAGE Publications.
- Strike, K. & Bull, B. (1981). Fairness and the legal context of teacher evaluation. Chp. 17 (pp. 303-343) in Millman, J. (Ed.) *Handbook of Teacher Evaluation*. Beverly Hills, CA: SAGE Publications.
- Wise, A. Gendler, T. (1990). Governance issues in the evaluation of elementary and secondary school teachers. Chp. 22 (pp. 374-389) in Millman, J. & Darling-Hammond, L. (Eds.). *The New Handbook of Teacher Evaluation*. Newbury Park, CA: SAGE Publications.