Education Policy Analysis Archives

Volume 6 Number 20 **November 3, 1998** **ISSN 1068-2341**

A peer-reviewed scholarly electronic journal.
Editor: Gene V Glass Glass@ASU.EDU.
College of Education
Arizona State University,Tempe AZ 85287-2411

Critical Evaluation for Education Reform

Gisele A. Waters
Auburn University

Abstract

The school reform movement has done little to provide an accurate analysis of the production of inequality or the reproduction of social injustice in the public schools or the larger social order. The ideology that influences this movement has often prevented the realization of any notion of an egalitarian ideal, the elimination of inequality, or the improvement of those who are least well-off. I ask educators and evaluators of education reform efforts to reconsider critically their roles in social science research, to reclaim the battleground of public school reform by focusing on the democratic purpose of public schooling, and the institutional problems in educational programs and practice that often inhibit action toward this ideal. The first part of this article includes an extensive argument explaining the "why" of critical evaluation. The theoretical literature on inquiry in science and social science, the ideology of critical theory, critical social psychology, and Freirean pedagogy are consulted as additional tools for augmenting the practice, policies, and responsibilities of evaluators in education. I review three contemporary perspectives of evaluation in order to begin rethinking the purposes and functions that evaluation serves in education. It also demonstrates how mainstream and contemporary evaluations can be used to serve a particular set of social and political values. The second part of this article begins a preliminary journey toward describing the "how" of critical evaluation. Critical evaluators can fight for social justice by combining the merit criteria of state and federal public education law, and the methods of an adversary oriented evaluation in order to transform educational environments that serve the future potentials of all children. Therefore education involves the practice of freedom, the means by which men and women deal critically and creatively with reality and discover how to participate in the transformation of their world (Freire, 1985).

The Argument for Critical Evaluation of Education Reform

Part I: The "Why" of Critical Evaluation

Schools are inextricably linked to the communities they serve through social, political, economic, and cultural interests. To better comprehend public education, the socio-cultural, political, and hierarchical relationships that transpire within the school as well as within the community must be linked to the broader political and economic issues of society at large (Ogbu, & Matute-Bianchi, in press). To begin to realize the possibility for reforming public education, and to begin fighting for social justice in education, especially for those children who are disadvantaged, we must first re-examine the historical nature of the problems of education and the communities in which these schools exist (Noll, 1997).

Education Reform

The Conceptualization of Educational Evaluation as Practical Educational Research

Many attempts have been made in recent years to clarify the meaning of evaluation and expose the distinction between evaluation and other related concepts such as measurement or research. The literature contains many approaches regarding the conceptualization of evaluation and the determination of its countenance in education (Nevo, 1986). According to Nevo (1986), many of these approaches have been unduly referred to as "models" (for example, the CIPP Model, the Discrepancy Model, the Responsive Model, or the Goal-Free Model) in spite of the fact that none of them includes a sufficient degree of complexity and completeness that might be suggested by the term "model." For the benefit of those of us who lost their way among the various models and approaches, I simply suggest taking a holistic approach to considering educational evaluation as an extended arm of practical educational research.

Education is a field like medicine in that its name simultaneously refers to a practice and to a field of disciplinary inquiry (Scriven, 1986). Scriven stated that the paradigm of research in the area of the philosophy of education, to take one example, is surely the paradigm of philosophical research in any area. But that leaves open the area of research that we normally think of as the domain of scientific research in medicine or education. Traditionally, we have tended to suppose that in this area of medical or educational research the correct model is that of the related sciences. That is, for example, educational research has modeled itself on social science research. Similarly educational evaluation has modeled itself as an offspring of educational research. In medical research that approach has brought some problems because it seems to lead to results that conflict with the practical wisdom of physicians and the economic realities of the patients. The same can be seen in education with the refined development of IQ tests, norm- referenced testing, and token economies for classroom management.

Scriven (1986) wrote that the conventional "scientific paradigm" way of dealing with these type of problems is not the business of science, they are value issues, and must be sorted out by the citizenry. Instead, he proposes a paradigm for practical educational research which subsumes educational evaluation, and which includes ethics, political feasibility, a set of practical alternatives, and an overall practical significance. Educational research is not, as he is suggesting, to be defined as all research that in any way involves the concepts related to education, because that's too broad (it includes learning theory), but as research that contributes to the facilitation of education, just as medical research should not be defined as all research that involves concepts related to medicine, since that brings in all physiological research, but simply as research contributing to health.

The research on classroom teaching, educational programming, school management, and classroom achievement have mostly been designed on the "quest for knowledge" model (traditional scientific) rather than on the "improvement of practice" model. Scriven's main point for educational evaluation stresses the acknowledgment that evaluation research in schools can be a far more complex business than just a quest for knowledge, a quest for classification, explanation, generalization, causation, and/or prediction.

In reviewing some of the theoretical literature of inquiry in science and the social sciences, it is hard to avoid the impression that there is a reluctance to confront the issues of power, democracy, inequality, ethics, politics, and pragmatics in educational research, evaluation, and in mainstream social science. Scriven proposed that one cannot reconcile the widespread support for the doctrine of a value-free social science with the

continued, inescapable practice of evaluation by social scientists, of the work and worth of students, peers, and selves, except by invoking a kind of phobia which makes them blind to the contradiction between their doctrine and their practice. This phobia, Scriven called "value phobia," has blocked us for nearly a century from addressing explicitly the methodology of evaluation and the systematic evaluation of our own practices in social science research (1986, p.62). With this in mind, I explain how the theoretical literature of inquiry in science and the social sciences can contribute to justifying the inclusion of such values as social justice within an expanded framework of critical evaluation of education reform.

Consideration of Inquiry in Science and the Social Sciences

Social Justice and the Distribution of Education

Considerations of social justice are applied in the distribution of virtually every social good. This is so much the case that, in the eyes of some, social justice simply has to be proclaimed (for example in political programs) to henceforth characterize the relations between people. In educational policy, arguments derived from social justice played a role even before World War II and were fought over by political parties, teachers' unions, left-wing intellectuals, and "pedagogical entrepreneurs" (Wesselingh, 1997). For them, the phenomenon of unequal participation was indeed a social problem, a phenomenon of social injustice. It does not take much effort to see that predominantly economic considerations have prompted the rapid expansion of equal- opportunities research. Opinions about the just provision of educational opportunities combined with economic need, have given the impetus to this research (Wesselingh, 1997).

Indeed, the Fall 1998 edition of Educational Evaluation and Policy Analysis includes one of the latest studies completed using Hierarchical Linear Modeling (HLM) on the relationship between students' opportunity to learn (OTL) and their science achievement. In this study, Jia Wang concludes that content exposure was the most significant predictor of students' written test scores, and the quality of instructional delivery was the most significant predictor of the hands-on test scores. In support for these types of conclusions, Berliner and Biddle (1997) clearly argue that opportunity to learn is the most significant predictor of academic achievement. These authors would be content to know that "scientific methods" such as HLM techniques are pushing the analysis of OTL variables at two level of instructional processes: the classroom level and the student level.

On the cusp of a new millennium, we are searching for answers not in the homes, economic backgrounds, and individual disadvantages of our students of public education, it would seem that we are finally beginning to look at the quality of instruction variables that exist in schooling processes instead of "blaming the victim." Can we begin to ask why and how our school systems are failing our children, instead of why and how these children are failing our school systems" If schools are to be held accountable for the equitable delivery of educational opportunities and if social justice is to take place within the halls of academic opportunity, the core of the education performance indicator systems should include school and classroom information.

According to Winfield (1993), there are two main reasons for obtaining OTL information. First and foremost, teacher and school factors need to be taken into consideration in explaining students' achievement. Teacher and school variables directly and indirectly influence student learning and student performance. Second, the new

performance-based assessments make the collection of OTL information crucial (National Council on Education Standards and Testing (NCEST), 1992). The performance-based assessments require higher order thinking skills. This may put students from low socio-economic status groups at a disadvantage. Studies have shown that minorities, especially African American and Hispanics, are more likely to be put into classroom with less learning opportunity even when ability is taken into account (Gross, 1993). If future research on achievement continues to disregard OTL variables, the achievement gap between majority and minority will continue to increase and a lack of educational opportunity will continue to expand (Arreaga-Mayer & Greenwood, 1986; Madaus, West, Harmon, Lomax, & Viator, 1992).

> Education as a good to be distributed gets the character of a good that provides access to other goods. The key power of schooling is based on the fact that education serves as a criterion for the distribution of all kinds of other material and immaterial goods. The consequence of this development is an instrumentalization of education. It evolves into an outstanding example of an instrument of mobility in a society where now qualification and rapidly growing demands for qualification create the space for moving up and, to a lesser extent, moving down the social ladder. This is at least the idea; the question of how education actually performs or is able to perform its role as a social agency of distribution for various social groups is of course not answered (Vervoort, 1975, p. 104).

For various groups this question still challenges our daily lives as critical evaluators, leaders and researchers of social justice in education.

As generally acknowledged since the traditional bourgeois ideas of the Enlightenment, the only valid criterion for determining who deserves which education is achievement. Achievement as a criterion for selection stems from egalitarian principles and is generally accepted in education as a just criterion. By now we know that this distribution model has let to serious forms of social inequality. The assumption that in schools everybody has equal opportunities to perform and thereby has a fair chance to take part in the subsequent competition on the labor market, has proven to be a misconception (Wesselingh, 1997). Education thus functions as an instrument for the reproduction of social inequality and thereby reflects the irony of a principle derived from egalitarian Enlightenment philosophy.

Social Justice and Education

Walzer's *Spheres of Justice*, published in 1983, can be seen as a reaction to John Rawls's *A Theory of Justice*, published in 1971. Walzer's objective was to provide an interpretation of what we contemporary Americans see as the essence of such concepts as equality and justice. His book makes clear that a discourse on the selection criteria for such an important social good as education is now needed more than ever. Reflection on this topic should not be left to politicians and policy makers for in that case considerations outside the sphere of justice will tend to dominate. Educational scientists, sociologists of education, and educational evaluators in particular, should definitely be more concerned with issues of social justice in education. Social justice is one of the most important values that we should hope to secure in critical evaluation studies of educational reform.

One of the goals of this article, besides arguing that critical evaluation is needed in

order to begin fighting for social justice in education, is to recommend an open and purposeful discourse about social justice in the reform of American public education between the "public intellectuals" (Giroux, 1997, p.263), otherwise known here as the social scientists, educational researchers, evaluators, and practitioners, a discourse about social justice in the reform of American public education. The participation in discourse that values a moral imperative and a political commitment to social justice in the evaluation of education reform is crucial to understanding the ideology of a critical evaluator.

Reaching Beyond the Incommensurable Perspectives

When it comes to dealing with such issues as social justice in education in a way that recognizes its moral complexity and political nature, the social sciences have "incommensurable perspectives" based in various traditions which have had different ideas about the individual and his/her society (Wertsch, 1998). These views have been updated but often at the cost of further fragmentation in the social sciences. The work of the new "public intellectuals" is to translate and connect, the ideologies and contributions of Aristotle's Realism, Plato's Idealism, Comtean Positivism, the Vienna Circle of Scholars and their Logical Positivism, Constructivism, Postmodernism, Critical Social Theory, and Feminist Theory.

The immeasurable challenge of the future is to look through diplomatic eyes without the "terministic screens" (Burke, 1966) of our specializations and disciplines that impair our vision. We could begin to address the phenomenon of public schooling, its reform, and its evaluation within a politically honest analysis. By refocusing our individual and collective powers into a moral and political analysis, critical evaluation of education reform in the next century can begin to regain the democratic imperatives or possibilities of public institutions. Exercising this moral and political "judgement" in evaluation of education reform, as a social responsibility in public practice, requires instrumental courage and conscience.

Analytic Primacy

Also this article aspires to begin a discourse beyond what politicians, educators, and philosophers have debated for centuries, the extent to which education should develop the individual or serve the needs of the state and society. The fact that this debate seems to go on and on with no principled resolution in sight suggests that deeper issues may be at stake. Namely, it suggests that academic dispute over what has "analytic primacy" (Wertsch, 1998, p.9), the individual or society, may reflect an underlying debate, a debate that cannot be resolved through rational argument. I am recommending that evaluators of education reform lift the blinders of methodological habit, move beyond their rational arguments, and discover how their own morals and politics are partly reflected in their professional decisions. With this in mind we live in times of increasing uncertainty as to how to reform public education. Part of the success of education reform will depend on those who have the power to affect social change, who have control over the knowledge base, who judge the worth and merit of educational programs, and what kinds of morals and politics are profoundly ingrained within their minds, spirits, and hearts.

Social Justice and Public Practice

For the most part, educational research and evaluation have remained both moral and political innocents in theory, practice, and policy. Part of this political innocence is derived from self reproducing ideologies and scientific paradigms that have explicitly or implicitly neglected moral and political issues. The conception of social justice, as considered here, is not a privilege for some (meritocractic) but rather a birthright for all (democratic) (Sirotnik, 1990). The value of social justice forms the foundation for working towards the restoration of a moral and political theory in the evaluation of public education reform, as part of a social responsibility in public practice, and as a part of confronting the moral and political purposes of social inquiry and research.

The contributions of Wertsch, (1998), Giroux (1997), Prilleltensky (1994), Tsoi Hoshmand, (1994), Howard (1985), Kohlberg (1984), Rawls (1971), Habermas (1971), and Kuhn (1970) are offered as significant commissions to support the reconsideration of our individual and professional decisions in education, by deliberating on our own morals and politics. Reflections and deliberations on our own values, beliefs, passions, and the reasoning for our professional decisions are mostly done outside of the confines of our "professional lives." Thus we are left with the interesting and paradoxical conclusion that what "ought" to be the most central in the evaluation of our schooling of American children, the moral and political reasoning, becomes inevitably peripheral to our public practice (Miller & Safer, 1993). In terms of articulating in-depth moral and political positions related to evaluation in educational reform, these considerations and decisions are vital to building and transforming schools that are struggling to achieve democratic ideals.

Between the Potential and the Present

Issues such as equality, democracy, race, gender, class, and poverty are certainly integrated through variable means into the contemporary scholarship of educational psychology, research, and evaluation. However, these issues and their historical, political, moral, and economic meanings are rarely discussed in a comfortable forum naturally or agreeably in the impregnable halls of academia. Therefore, the silent space between the potential in education and the present crisis in public education is successfully and safely insulated decade after decade. As a result, inquiry and discourse between "public intellectuals" remain fixed in a non- political environment without values, beliefs, and passions. This environment within an "ideology of neutrality" became internalized in the consciousness of most researchers following the establishment of the modern university. The links between the political agendas and research were, and often remain, blurred by the legitimating function of social and educational research. This can be seen in many educational evaluation studies that accept the objectives of pedagogical programs and are organized to "explain" how the objectives were reached.

Redefinition of Identity and Purpose

No problem is more difficult and complex in the social sciences than that of determining how morals and political values are embedded within the research methodologies that we employ and the "academic" decisions that we make (Cronbach & Associates, 1980; Hamnett, Kumar, Porter, & Singh, 1984; Fetterman, 1988; and Sirotnik, 1990; Maruyama & Deno, 1992). That morals and political values should exist in research is no longer denied (Warren, 1963; Fetterman, 1981; Freire, 1985; Apple & Beyer, 1988; Habermas, 1990; Prilleltensky, 1994; Giroux, 1997; Kanpol, 1997;

Wertsch, 1998). In terms of educational evaluation, the ideas found in this article, reconfirm the conviction made by Sirotnik (1990) that the practice of evaluation is part of the political authority structure of society, and that evaluation as an aid to public decision making entails conceptions of democracy and social justice, even when these conceptions are not immediately apparent.

House (1993) wrote that evaluation receives its authority not only from its presumed "scientific method" but also from government endorsement itself. Within the analysis of evaluation in advanced capitalist societies, House reviewed how governments face serious problems in governing such a multicultural "amorphous mass of people" (1993, p.vii) and how evaluation is both political and scientific authority applied to practical decision and actions, particularly public decisions and actions. He went on to explain how governments are capable of making decisions based on their own political authority, but that it is easier to govern based on voluntary acceptance by the populace attained through scientific persuasion, particularly when the populace is pluralistic and increasingly non-traditional. In addition, House expanded the notion of political and scientific authority by redefining formal evaluation as a new from of cultural authority. Cultural authority can be manifested in the probability that descriptions of reality and judgements of value will prevail as valid, an increasingly difficult accomplishment in societies with disparate value systems (House, 1993). Current literature in evaluation confirms that evaluation as a social activity is becoming increasingly self-conscious about its own identity and purpose in the larger social order (Cronbach & Associates, 1980; Guba & Lincoln, 1989; Rossi & Freeman, 1993; Patton, 1994; Scriven & Kramer, 1995; Chelimsky & Shadish, 1997).

Critical Evaluation

Critical evaluation of education reform involves the practice of completing empirical, historical, public and social work by employing explicit theories of justice (House, 1976,1980) that require serious commitment, persistence, courage, conscience, and conviction in order to restructure and transform education environments. Hence, as a social practice, evaluation involves an inescapable ethic of public and social responsibility that extends well beyond the immediate clientele by focusing on the democratic purpose of schooling. Social justice in evaluation, then, concerns the manner in which various interests are served. Critical evaluation should serve the interests not only of stakeholders, sponsors, or the reformers, but of the larger society and of various groups within society, particularly those most affected by the educational programs under review. One of the aims of this article is to reiterate that institutions of higher education must be seen as deeply moral and political spaces in which evaluators, indeed intellectuals, assert themselves not merely as professional academics but as citizens, whose knowledge and actions presuppose specific visions of public life, community, and moral accountability (Giroux, 1997).

A Political Theory

I propose here that critical evaluation represents a kind of political theory that integrates explicitly the value of social justice into the practices, policies, and responsibilities of evaluation of educational reform. Moreover, the political theory of critical evaluation can be defined as the implicit and explicit social and professional ethics of evaluation, and the moral and political consequences of these ethics, which could reconstruct and reconsider the power relations in academia and public education.

One of the reasons to begin a journey into a critical political perspective in educational evaluation is to arrive at an account, a kind of "translation at the crossroads" (Wertsch, 1998, p.7), that would make it possible to link, but not reduce, one perspective of "science" to another. Another reason is to begin addressing explicitly the methodology of evaluation and systematically evaluate our own practices in social science research (Scriven, 1986).

The task is to reflect, to discourse, and to collaborate with each other, between and within disciplines, to dialogue about the human condition, especially the conditions of inequalities that public institutions perpetuate in our democratic society. In order to talk and listen to one another about social justice in education our "knowledge base" and our morals and politics should be integrated into an ideology of hope and sincere cooperation for a better future for children through education reform.

Overview

A characterization of a critical evaluator will be advanced shortly. The role divisions of academic versus service orientations existing in evaluation today are described. The ideology of a critical theory of education, and critical social psychology will then be reviewed in order to consider augmenting traditional positivist perspectives of evaluation. Afterwards I give a brief summary of evaluation in general. Three perspectives of evaluation and their purposes are explained, in order to illuminate the more traditional positivist approaches in prevalent current evaluation literature and to describe a spectrum of responsibility, purpose, and definition within the discipline of evaluation. The three perspectives on the spectrum are those of accountability, knowledge, and development.

Next, the limitations of contemporary and critical evaluation and how these approaches may implicitly serve a particular set of social and political values is forwarded. Integration of critical evaluation into a changing society, Fetterman's silent scientific revolution, the ideas of practicing critical evaluation, the neutrality of schools, and change in American schools are also presented. Subsequently this article conceptualizes one important process that an evaluator must experience in the context of Freirean pedagogy, so that a critical evaluator can begin the special role of critical evaluation in educational reform. The implications of critical social thought for evaluation in educational reform are then proposed. Finally, the second part of this article begins by describing the interdisciplinary methods and procedures of the "how" of critical evaluation, by introducing the integration of American public school law as enhanced by collaborative consultation and the adversary-advocate oriented evaluation model.

The Critical Evaluator

Ernest R. House was the first prominent evaluation theorist to advocate valuing based on principles of social justice (Patton, 1997). He has consistently voiced concerns for democratizing decision making in that context, he has analyzed the ways in which evaluation inevitably becomes a political tool in that it affects "who gets what." As mentioned earlier, education itself, as well as educational evaluation can enhance fair and just distribution of benefits or it can distort such distributions and contribute to inequality. In considering judgements on programs, the social justice evaluator, the critical evaluator, is guided by such principles and values as equality, fairness, and concern for the common welfare (Sirotnik, 1990).

Kenneth Sirotnik and Jeannie Oakes collaborated in this same endeavor by considering the epistemological connections between critical theory and evaluation. To be specific, they stated that if one accepts the proposition that inquiry is never value free and accepts social justice as the ethical starting and ending points for moral argument, then the accumulated body of work done by Freire (1973), Habermas (1971), and others points the way toward a useful epistemological synthesis, one that they called critical inquiry, that is evaluative by its very nature (Sirotnik & Oakes, 1990). By no means is critical evaluation a new idea. Regardless, the argument for fighting for social justice with critical evaluation of education reform is not a trivial one, but it is an argument that I have extended with much interdisciplinary literature and paradigmatic considerations.

Michael Quinn Patton (1997) wrote that social justice and other similar principles change the role of the evaluator from the traditional judge of merit or worth to a social change agent. Many evaluators surveyed by Cousins, Donahue, and Bloom (1996) were hostile to or at least ambivalent about whether evaluation, particularly a type of critical evaluation, can or should help bring about social justice. Certainly, evaluators undertaking such an approach need to be comfortable with and committed to it, and such an activist agenda must be explicitly recognized, negotiated with, and formally approved by primary intended users. From Michael Quinn Patton's utilization focused perspective, using evaluation to mobilize for social action and support social justice are options on the menu of evaluation process uses (1997).

In this article, part of the argument is that wherever one places oneself on the spectrum of evaluation responsibility, purpose, and definition; the evaluator can earnestly acknowledge the powerful critical role that he or she may interpret in placing judgement or giving merit to one of the most profound social activities in our lives, that of educating our students and our children. This role as a critical evaluator can be found anywhere on the spectrum. As typically happens with most spectrums the outlier situation is pretty rare. A critical evaluator can produce empirically traditional research designs in combination with critical social ideology, as long as one maintains a critical stance towards methods, practice, and policy that addresses the more difficult questions about the institutional problems in educational programs, those of democracy, power, and inequality. Patton (1994) also advocated the use of "mind shifts back-and forth between paradigms within one evaluation setting."

Most of the time, in most environments represented on the spectrum, "scientific" positivist traditions about knowledge and postmodern critical social constructions of knowledge are almost bound together, and evaluators must therefore always be prepared to confront them both (Young, 1990). In *Ethics, Politics, and International Social Science Research*, Hamnett, Kumar, Porter, and Singh (1984) compared and described the aforementioned theoretical presuppositions such as that of positivist constructions of knowledge and that of critical theories of knowledge. A significant point here is that a critical evaluator can utilize the necessary tools and methods within shifting research paradigms and changing concepts of knowledge construction, in order to augment practices and policies which are continuously participating in a discourse that values a moral imperative and a political commitment to social justice in the evaluation of education reform. This understanding of a moral imperative and a political commitment in educational evaluation is crucial in establishing explicitly the ideology of a critical evaluator and in making one's analytical biases clear.

The following paragraph provides a synopsis of Sirotnik's and Oakes' review of the three faces of inquiry and analysis (1990). Most educational researchers and evaluators have been schooled in the tradition of the scientific method and the hypothetico-deductive paradigm borrowed, presumably, from the physical sciences. But

there are at least two other separate and general orientations for systematic inquiry having strong philosophical roots and demonstrable utility for the social sciences. The more familiar is the whole class of naturalistic methodologies. The second major departure from the empirical analytical tradition is less well known and much more separable, namely, the critique of knowledge. Its roots are also in the hermeneutical tradition. But as a philosophy of inquiry, it represents a significant extension of interpretive inquiry. Inquiry and analysis does not happen in a normative vacuum, as they so eloquently stated.

Sirotnik and Oakes (1990) also suspected that "an epistemological trap can be created through assuming necessary and sufficient connections between method and the political content of cognitive interests. Conducting empirical analytic inquiry, for example, does not necessarily imply a hidden agenda of domination. On the other hand, a hidden agenda of domination cannot in principle survive an inquiry based on critical theory" (p.45). I agree with these authors that this, indeed, points the way out of the trap, a truly practical unification of the three faces of inquiry requires the self correcting epistemological stance that is made to order in a critical perspective. At the same time evaluation must consider what kind of orientations are created in practice when these epistemological and empirical stances are postured.

Academic Versus Service Organizations

One of the most basic role divisions in the profession today is between academic and service oriented evaluators, a division identified by Shadish and Epstein (1987) when they surveyed a stratified sample of the members of the Evaluation Network and the Evaluation Research Society, the two organizations now merged as the American Evaluation Association. The authors inquired about a variety of issues related to evaluators' values and practices. They found that responses clustered around two contrasting views of evaluation. Academic evaluators tend to be at universities and emphasize the research purposes of evaluation, traditional standards of methodological rigor, summative outcome studies, and contributions to social science theory (Patton, 1997). Service evaluators tend to be independent consultants or internal evaluators and emphasize serving the stakeholders' needs, program improvement, qualitative methods, and assisting with program decisions (Patton, 1997).

According to Shadish and Epstein, " The general discrepancy between service oriented and academically oriented evaluators seems warranted on both theoretical and empirical grounds" (1987, p.560). The profession of evaluation remains very much split along, these lines, but with new twists and perhaps, deeper antagonisms (Patton, 1997). Patton goes on to explain how the "schism" erupted openly, and perhaps deepened, in the early 1990's, when morality entered into the evaluation arena much more explicitly, and the American Evaluation Association elected successive presidents who represented two quite divergent perspectives.

Yvonna Lincoln (1991), in her 1990 presidential address, advocated what Patton would call an activist role for evaluators, one that goes beyond just being competent applied researchers who employ traditional scientific methods to study programs, the academic perspective. She closed her speech by asserting that "my message is a moral one." The following year, the American Evaluation Association president was Lee Sechrest, who by his own definition represented the traditional, academic view of evaluation. He objected to Lincoln's metaphorical call for a new generation of evaluators. "I ask myself," Sechrest (1992) mused, "Where in our makeup are the origins of this new creation so unlike us.... I sense a very real and large generational gap" (p.2).

From this contemporary discourse in what the role divisions personify in evaluation, one can tell that critical evaluators may be characterized as divergent or even marginal in their theoretical and empirical presuppositions. Here lies the embedded professional challenge of remaining open to pluralistic and cosmopolitan approaches which adapt evaluation practice to new situations, mainly the situation of public education institutions which are failing a growing disproportionate amount of disadvantaged children thereby reproducing social and symbolic inequalities. Ultimately, there is no one way to conduct an evaluation. This insight is crucial. The design of a particular evaluation depends on the people involved and their situation.

Ideologies of Critical Theory, and Critical Social Psychology

Traditionally social science and social psychology we are told, is a vocation of scientific method, a devotion to truth that should not be compromised by the researcher's idiosyncracies or other external forces and should not be unduly affected by the social context in which the researcher operates (Hamnett et al., 1984). In the realm of the natural sciences, statements often appear to be reaffirming this stance. For instance, in practice there is very little to distinguish Soviet and U.S. nuclear physics. Changes in theoretical presuppositions in one country are rapidly translated to others.

Social science and social psychology, however, do not have the canons of theoretical perspective, verification, or even of data collection found in natural science (Hamnett et al., 1984). Hamnett and his co-authors state that theoretically, the sociology of knowledge has demonstrated how science (including the concepts, methods, and procedures embodied) presupposes historically relative values, interests, and ideologies. The taken for granted notion of the methodological neutrality of scientific method has been undermined by theorists of many persuasions including that of critical theorists and critical social psychologists (Wexler, 1983). I agree with Wexler when he writes that conventional wisdom and common sense concedes that science is influenced by human values and the political contexts of its expedition. This is why the evaluator of education reform cannot posture a neutral, purely objective point of view on the object of his research, especially in the reforming of such a social contract as education.

The writings of critical theory developed from the Institute for Social Research in Frankfurt. The critical theorists are concerned with the role of values and ideology as "part of the conceptual framework which defines what it is to have, i.e., scientific knowledge about some phenomenon" (Sabia & Wallulis, 1983). Such a focus raises important questions concerning social science research, ethics, and inevitably the practice of evaluation in education. Critical theorists state that it would be incorrect to claim that positivist doctrine is responsible for the unreflexive state of the research ethics and politics debate in social science; the social, historical, economic, and political context of research is of overwhelming importance (Sabia & Wallulis, 1983).

How one views the role of social research, its relations to political practice, and how one assesses responsibilities, relationships, and appropriate conduct should be explicitly negotiated up front with potential clients in terms of one's underlying assumptions and ideological presuppositions. Moreover, critical research methodology is distinctive from other approaches in that it traces the origin of our concept of validity back to everyday human interaction. This is true, at least, for the specific brand of critical methodology I advocate, which draws heavily from Habermas's work on validity (Habermas, 1981, 1987). The later discourse of this critical evaluation perspective, which can be embedded in a positivist scientific method, does not assume the posture of rejection or exclusion, but rather will serve to provide an additive component to

constructing knowledge and representing it with critical and conscious eyes.

I repeat what Lewis Carroll's Alice would have said, "things are not what they seem." There is a difference between listening to the goals of reformers, and listening attentively to the underlying assumptions forwarded by education reform efforts, and consequently holding the reformers responsible for living up to their social ideals and their program mission statements, mainly those mission statements that become framed cultural symbols of what a program or a school represents. These framed paper certificates, these mission statements, are usually strategically placed in the front office of every public school and meticulously published in brochures summarizing the goals and objectives that school districts represent to welcome potential inhabitants of the communities they serve. If we can understand the central role played by validity claims in normal human communication (symbolic or otherwise), we will then be able to formulate the special requirements that a critical evaluator conducting formal inquiries into social processes must employ to produce a trustworthy account. In critical evaluation, the validity claims made by the evaluator do not differ in nature from validity claims made by all people in everyday contexts.

Critical social psychology draws from the critical theory of the Frankfurt School and the theoretical traditions of Marxism (Wexler, 1983). Philip Wexler (1983) augmented and amplified what he perceived as developing tendencies in social relations and in social psychological processes. Like Philip Wexler's expression of a need for augmentation, I am asking those who study, practice, and use evaluation in education to broaden and amplify their view of the applications and functions of evaluation with an eye to the future. The evaluator could be responsible for reaching beyond mainstream philosophy and practice in evaluation because the transforming of education and the reforming of such a significant social activity requires an exceptionally conscious human being. Like critical social psychology, a critical evaluator requires a theory which can comprehend and facilitate social change movements.

Next I shall give an overview of evaluation in general, its development, and then review three perspectives of evaluation and their purposes, in order to illuminate the more traditional positivist perspectives in prevalent current evaluation literature. These three perspectives again are those of accountability, knowledge, and development. By looking at these three perspectives and their positions along a spectrum, I argue that the evaluator must go beyond those delineated perspectives in mainstream evaluation theory, policy and practice, in order to take a more critical posture toward both education and the very process of thinking about education.

Evaluation

Evaluation as an academic discipline, a profession, and a government function has only developed in the past four decades in the United States and in several other industrially developed nations. In many nations, however, evaluation is in its infancy as a standardized pursuit; and certainly on a global scale, evaluation is only beginning to enter the scene (Chelimsky & Shadish, 1997). There is comfort in knowing, as previously mentioned, that current literature in evaluation confirms that evaluation as a social activity is becoming increasingly self- conscious about its own identity and purpose in the larger social order and is beginning to systematically evaluate its own methodology, utilization, values, and politics (Cronbach & Associates, 1980; Guba & Lincoln, 1989; Rossi & Freeman, 1993; Patton, 1994; Scriven & Kramer, 1995; Chelimsky & Shadish, 1997; House, 1993; Scriven, 1991). I would agree with Chelimsky and Shadish (1997) when they propose that evaluators, in whatever field of

evaluation they may be, are likely to find themselves, at least sometimes, at odds with the political actors, systems, and processes in their own backyards, that rally against a free flow of information and collaborative action which endangers the status quo.

Between 1965 and 1990 the methodology, philosophy, and politics of evaluation changed substantially, partly in response to the structural transformations in an advanced capitalistic society (House, 1993; Scriven 1991). The strongest stimulus to the development of evaluation was Lyndon Johnson's Great Society legislation, which, though not capable of changing U.S. society as a whole, certainly transformed educational and social research. At Senator Robert Kennedy's insistence, the Elementary and Secondary Act in 1965 mandated evaluation of programs for disadvantaged students, and this spread to all social programs and beyond (McLaughlin, 1975). House (1993) reviewed clearly in lay terms how evaluation moved from monolithic to pluralist conceptions, to multiple methods, multiple measures, multiple criteria, multiple perspectives, multiple audiences, and even multiple interests.

Methodologically, evaluation moved from a primary emphasis on quantitative methods, in which the standardized achievement test employed in a randomized experimental control group design was most highly regarded, to a more permissive atmosphere in which qualitative research methods were acceptable. Mixed data collection methods are advocated now in a spirit of methodological ecumenism (House, 1993). The following three perspectives describe more thoroughly the way that evaluation is characterized in contemporary evaluation circles.

Examples of Purpose and Perspectives in Evaluation (Chelimsky & Shadish, 1997)

Below find a review of the definitions and characterizations that Chelimsky and Shadish write about in *Evaluation for the 21st Century*. They offer an inexhaustible listing of possible purposes for evaluation. These purposes include the following: (a) to measure and account for the results of public policy, and programs, (b) to determine the efficiency of programs and their component processes, (c) to gain explanatory insight into social and other public problems, (d) to understand how organizations learn, (e) to strengthen institutions and improve managerial performance, (f) to increase agency responsiveness to the public, (g) to reform governments through the free flow of evaluative information, and (h) to expand results or efficiency measurement from that of local or national interventions to that of global interventions such as reducing poverty and hunger or reversing patterns of environmental degradation. All of these purposes are, of course, worthwhile and legitimate reasons for conducting evaluations, but they differ with regard to the questions they address and the kinds of methods needed to answer these questions.

Chelimsky and Shadish propose that these different purposes, along with the questions they seek to answer, seem to fall naturally into three general perspectives:

- evaluation for accountability (e.g., the measurement of results or efficiency);
- evaluation for knowledge (e.g., the acquisition of a more profound understanding in some specific area or field); and
- evaluation for development (e.g., the provision of evaluative help to help strengthen institutions).

The methods of these three perspective are not mutually exclusive. Though they do represent notable differences on a variety of dimensions. Each may be needed at

particular times or policy points and not others (e.g., evaluation for knowledge may need to precede accountability). Chelimsky and Shadish (1997) write that they appear to have considerable explanatory power with regard to the current tension in the evaluation field. (See Table 1 for further details.) Table 1, an adapted chart from Chelimsky and Shadish's book (1997, p.21), shows the following three different perspectives and their respective positions along five dimensions.

Table 1
Three perspectives and their positions along five dimensions, adapted from Chelimsky and Shadish (1997, p.21)

DIMENSIONS	ACCOUNTABILITY PERSPECTIVE	KNOWLEDGE PERSPECTIVE	DEVELOPMENTAL PERSPECTIVE
PURPOSE	to measure results or value for funds expended: to determine costs, to assess efficiency	to generate insights about public problems, policies, programs, & processes, to develop new methods and to critique old ones	to strengthen institutions to build agency or organizational capability in some evaluative area
TYPICAL USES	policy use, debate and negotiation, agency reform, public use	enlightenment use, policy, research and replication, education, knowledge base construction	institutional or agency use as part of the evaluative process, public and policy use
EVALUATOR ROLE	distant	distant or close depending on evaluation design and method	close, the evaluator is a "critical friend" or may be part of a team
ADVOCACY	unacceptable	unacceptable, but now being debated	often inevitable, but correctable through independent outside review
POSITION UNDER POLICY DEBATE	can be strong (depending on leadership)	can be strong (if consolidated and dissemination channels exist)	uncertain (based on independence and control)

The Accountability Perspective

From the standpoints of auditors, government sponsors of evaluation studies, donors to international organizations, and many others, evaluation is done to establish accountability. This involves the provision of information to decision makers, whether they are in the public or private sector. Specific cause and effect questions about the results in an accountability perspective might be: What happened to poverty levels among the very poor as a result of development assistance provided" Did an educational intervention or program produce more "effective" learning for all learners? Has teacher training increased student achievement?

Sometimes, questions about the results from an accountability perspective may involve merely documentation of whether or not anything has changed after something new has been tried. Normally, however, the ability to say that something is in fact a "result" hinges on the ability to establish that it came about because of something else. Many methods are used to answer these kinds of accountability questions including: randomized designs, quasi-experimental designs, mixed multi-level designs, mixed qualitative/quantitative designs, case studies, process studies, and research synthesis designs.

The Knowledge Perspective

In the view of many researchers working independently in universities and other evaluators in scientific institutions, evaluation is done to generate understanding and explanation. Chelimsky and Shadish (1997) stated that the specific questions may not be especially important to analyze here, given that it is the evaluator who decides what will be asked and answered, and the topic generally follows from the researcher's prior work. They explained that the evaluations associated with individual academic researchers, or those of research teams, will be more likely to continue in depth cumulative inquiry into particular areas or sectors of research than to be concerned with applying systematic research methods to a variety of sectors, as with accountability and developmental evaluations.

The larger purpose of the knowledge perspective is to increase understanding about the factors underlying public problems, about the "fit" between these factors and the policy or program solutions proposed, and about the theory and logic (or lack thereof) that lie behind an implemented intervention. "These evaluations may employ any of the methods discussed above, separately or in conjunction with each other, but the purpose of knowledge gain leads logically to the use of the strongest designs as well as the greatest clarity possible in explication and documentation of methods to facilitate replication or later use in research synthesis and policy formulation" (1997, p.14).

The Developmental Perspective

For government reformers, public managers, and others, evaluation is done to improve institutional performance. It serves as a flexible tool that works: (a) to improve the design of projects, (b) to measure and recommend changes in organization activities, (c) to develop the indicators and performance targets needed to improve institutional effectiveness and responsiveness, (d) to monitor, in an ongoing way, how projects are being implemented across a number of different sites, and/or (e) to find out how beneficiaries feel about an agency and its programs. To some accountability or knowledge perspective evaluators, developmental evaluators may seem more like evaluation "consultants" than evaluators, but those who do developmental work are convinced that building evaluation capability is as important an evaluation function as evaluation itself and that indeed, in some cases, evaluation cannot be done without capacity building.

Specific questions asked of evaluators in a developmental perspective might include the following: What is the best research evidence with respect to formulating a new program or modifying an old one? How can projects be structured so that they produce evidence on the value of the intervention being tested? What is the most appropriate agenda for the agency? Both process and outcome designs may be used in a developmental perspective, depending on the evaluation question posed. In addition to

the methods mentioned earlier, the formative methods used in the developmental perspective include the following: monitoring, empowerment evaluation, cluster evaluation, performance measurement, and research synthesis of both qualitative and quantitative methods. A developmental evaluator becomes part of the design team helping to shape what's happening both processes and outcomes, in an evolving, rapidly changing environment of constant interaction, feedback, and change. Using mixed methods and multiple criteria in this perspective are productions of some of the many current trends in the practice of evaluation.

Demonstrating a Particular Set of Social and Political Values

Although evaluation has developed as a discipline, a profession, and as a government function in the past four decades by building on its scientific positivist traditions and by systematically evaluating its own existence in the larger social order, this particle emphasizes continual growth and augmentation of its practices, policies, and responsibilities through a "conscientization" of evaluation's socio-political reality. Over the years evaluation has come to be seen as political. Michael Quinn Patton, at the National Evaluation Conference in Youngstown State University held in September 1998, summarized 12 recent trends in evaluation. One of them being the increasing political sophistication and acknowledgment of the role of values and morals in evaluation practice. There can be no doubt, that evaluation is influenced partly by political forces, and in turn, has political effects. Whose interests are served and how interests are represented in an evaluation are now very critical concerns in a society with increasing disparate value systems.

In the earlier days, it was assumed that the interests of all parties were properly reflected in the traditional outcome measures, but this assumption came to be questioned, and it was recognized that different groups might have different interests and might be differentially affected by the educational program and its evaluation (House, 1993). "Stakeholders" (those who had a stake in the program under review) became a common concept, and representing stakeholder views in the evaluation became an accepted practice.

The stakeholder concept is based on the prevailing pluralist-elitist-equilibrium theory of democracy, which disclaims any normative judgements and which holds that the current system of competing parties and pressure groups performs the democratic function of equalizing the diverse and shifting political demands (MacPherson, 1987). It is perceived that describing what others value is the stance best suited to the political context in which evaluators operate, because decision making depends on the values held by relevant policy makers and stakeholders. Presumably, these parties will use the findings to make informed decisions. Neither the government nor the evaluator is supposed to intervene to support any particular interests but rather only to provide information that is value- neutral and interest-neutral. The interests of various groups somehow dissolve into the values of decision makers and stakeholders.

However, one must note that today's professional evaluators sometimes become evaluators by default. We represent an eclectic and diverse combination of various professional, academic, and research areas. Shadish and Epstein (1987) found that 31% of the respondents in their survey described their primary professional identity as that of "evaluator" (p. 560). Others thought of themselves first as a psychologist, sociologist, economist, educator, and so on, with identity of evaluator secondary. When both Charles Murray (1983,1984) and Michele Fine (1983b, 1988) have been successful evaluators representing a particular set of social and political values and interests, one has to

acknowledge the diverse socio-political reality in which evaluators actually find themselves in practice.

In two highly visible stakeholder evaluations funded by the federal government, those of Cities-in-Schools and Jesse Jackson's PUSH/Excel program, the evaluations worked against the interests of the program participants and the inner-city students which the programs were supposed to serve, thus calling into question the justice of these evaluations (House, 1988; Stake, 1986). The results of the PUSH/Excel evaluation were used not only to discredit the program but also to question Jesse Jackson's ability to manage large enterprises during ensuing presidential campaigns. In truth, the stakeholder model was never implemented (House, 1988; Stake, 1986). Charles Murray, the evaluator in both cases, substituted a technocratic model of evaluation and expressed his disdain for the stakeholder concept in his article *Stakeholders as Deck Chairs* (1983). Although the stakeholder approach seems firmly entrenched, there is disagreement about how to implement it. In reality, stakeholders do not have equal power to influence and utilize the evaluation, nor do they have equal protection from the evaluation.

These types of problem in evaluation led into a discussion of misuse of findings. The fact that so much standardized achievement testing is reported to the public in general and its interpretation left to the media or government officials makes misuse particularly salient (House, 1993). In fact, the professional standards for evaluation developed by a committee led by Stufflebeam, devoted considerable space to issues of misuse, but the context in which evaluation results are presented does not lend itself to the employment of such standards, even though the standards are widely accepted in the evaluation community itself. How misuse of findings and disparate interests can be curtailed is by no means clear. The professional standards for evaluation developed by the Joint Committee dramatically reflected the ways in which the practice of evaluation had matured in 1981. In 1994, revised standards were published following an extensive review spanning several years.

While some changes were made in the 30 individual standards, the overarching framework of four primary criteria (utility, feasibility, propriety, and accuracy) remain unchanged. However, the profession of evaluation has not yet developed to the point of reflecting a common core of practices and principals as demonstrated by the original professions, divinity, law, and medicine (House, 1993). We must pay attention to the fact that certification programs and higher education programs in evaluation and evaluation research are a very recent development in the discipline (Chalimsky & Shadish, 1997). For a deeper understanding of how the original professions compare with evaluation as a profession, refer to House's (1993) book, *Professional Evaluation* .

Limitations of Contemporary Evaluation and a Reflection on American Public School Law

There are limitations to contemporary and critical evaluation frameworks. The problem of addressing multiple values and interests and how they should be represented in an "equitable" evaluation can take one directly into the realm of social justice and the recognition of the assumptions, character, and consequences of conventional forms of educational evaluation and American public school law. The problem of evaluation representing a particular set of political and social values (i.e., a broadly conservative set) also raises some serious questions about evaluation in general. Although the socio-political reality of multiple stakeholders and evaluators who have legitimate values and sometimes conflicting interests is recognized, how these values and interests are legitimized will become one of the most important challenges for educational

evaluation in the future, especially for critical evaluation of education reform. How to synthesize, resolve, and adjudicate all these multiple multiples in our increasing multicultural and amorphous society remains a formidable question, as indeed it does for the larger society.

One thing we do know is that the socio-political reality in evaluation of public programs, both in education and health, often works in favor of higher income groups and against equity despite the stated objectives (Birdsall & Hecht, 1995; Paul, 1991; Fine, 1983). When we look at the political structures and the broad organization of society, resource allocation and subsequent delivery of services and programs tend to be skewed in favor of those who have more "voice" (Fine, 1983; Fine & Weis, 1993). In many cases, powerful stakeholders or groups, which are able to effectively demonstrate their interest in receiving social services and "effective" or "successful" social programs, manage to get the lion's share of the resources and the funds. It is no secret that the United States of America is one of the last Western industrialized nations to base their educational financing system on that taxation of largely differentiated property values. This financial arrangement alone should illuminate some of the deeper issues at stake in the evaluation of public education environments.

American public school law and its case history has demonstrated time and again that there are very few instances where citizens have been able to prove that state school finance systems result in revenue disparities which violate the Equal Protection Clause of the Fourteenth Amendment. In 1973, in the case of San Antonio Independent School District v. Rodriguez, Mr. Justice Powell, delivered the opinion of the Supreme Court. He said, "to the extent that the Texas system of school financing results in unequal expenditures between children who happen to reside in different districts, we cannot say that such disparities are the product of a system that is so irrational as to be invidiously discriminatory...."

If disparate allocation of governmental benefits can be justified on the basis of reasonable classification or the interests involved are not fundamental, then statutes will be regarded as constitutional (Alexander & Alexander, 1992). The court in the Rodriguez case basically ruled that a state legislature can heap benefits on some wealthy school districts and deprive others of fiscal resources and not offend the federal Equal Protection Clause. Thus representing the educational interests of disenfranchised stakeholders, even within the American public school law domain, can be confounded with many inherently unequal and disparate value systems.

In other instances, our social service institutions, such as education and health, are able to shape the systems to serve their own personal and professional goals at the expense of equitable delivery (Paul, 1991). Problems created by the limited voice of politically weak or disenfranchised stakeholders are exacerbated in educational evaluation, when combined with direct provision of services in virtual public monopolies of the "best teachers," the allocation of "best practices" in education, and the provision of high quality curriculum and professional development training which are centralized in higher socio-economic communities. Ultimately, citizens have limited capacity to improve the public education they are provided through participating, informing, and making recommendations. This is especially true of lower socio-economic community stakeholders which have traditionally been limited in their capacity to have their "voice" heard without legal representation (Fine, 1993; Oakes & Guiton, 1995).

Historically, when interests have been ignored and educational procedures have been violated, lower socio-economic communities, minorities, exceptional populations, and limited English proficient citizenry have had to turn to the legal system for any kind

of adjudication (Paul, 1991; Haring, McCormick, & Haring, 1990; Oakes & Guiton, 1995). Similarly, in terms of fighting for social justice in education, evaluation of education reform efforts could benefit from addressing some of the principals in American public school law. This idea will be further developed in Part II below. However, for the time being, contemporary evaluation which was invented to solve social problems, can be afflicted with many of the problems it was meant to solve.

Another limitation of critical evaluation in education reform pertains to its inherent questioning of the institutional character of education. By producing educational criticism and value judgements of institutional programs and personnel, in conjunction with the ideologies of critical theory of education, critical social psychology, and Freirean pedagogy, critical evaluators risk certain professional isolation from the mainstream. The socio-political reality in which one can survive as an evaluation professional of education reform becomes integrated into a world with those individuals that agree with your views, particularly those who agree with your views on social justice and in general the democratic purposes of public schooling. As critical evaluators conduct evaluations to address the elimination of inequality and the improvement of those who are least well off, they will come into conflict and threaten established authority. Any method of evaluation that claims to be nonobjective and value-laden will be marginalized. Society expects evaluation to be based on scientific authority. However, I expect the notion of what is scientific to be substantially redefined. The concepts of objectivity, scientific methodology, and validity will be recast to accommodate different evaluation approaches (House, 1993).

Integration of Critical Evaluation into a Changing Society

Evaluation continues to become ever more methodologically diverse. Evaluation in general draws from the theoretical foundations of many fields and is multi-disciplinary and multi-faceted in nature (Chelimsky & Shadish, 1997). It is by now well established that the full array of social science methods belongs in the evaluator's methodological tool kit, including tools from psychology, statistics, education, sociology, political science, anthropology, and economics (Cronbach & Associates, 1980). When the critical logical and analysis tools given to us by critical theorists and social psychologists are included into an evaluation design, the role that evaluators play in judging the worth of educational reform efforts is elaborated. Chelimsky and Shadish (1997) supported the notion that it is often uncomfortable to stir oneself from familiar cultural, ideological, topical, conceptual, and methodological niches.

However uncomfortable or reactionary one may feel to the content of this article, there is a message: it is that evaluations of educational reform efforts in the next century can and probably will be far more powerful and influential than they are today. This is because of the ever increasing complexity of social, economic, technological, political, and cultural tensions which are questioning the very integrity and purpose for public education as a whole (Giroux & Aronowitz, 1991). The growing populations with disparate value systems and socio-economic levels and the increasing minority populations in this country will demand to participate more legitimately in the reformation of their own education. Consequently evaluation will have to redefine its identity, its purpose and practices.

Lee Cronbach, in 1980, advanced the position that the theory of evaluation has to be as much a theory of political interaction as a theory of how to determine facts or how knowledge is constructed (Cronbach et al., 1980). Even so, 18 years later, we still do not seem to understand political processes very well, especially their dynamic nature. This

gap in understanding and consciousness is especially true for evaluators in the field of education where we are determining "facts" and constructing knowledge about educational programs designed to improve teaching and learning in the public school domain. We can begin bridging this gap in consciousness to understand the political nature of evaluation by looking at our own ideologies as evaluators. Critical thought, indeed, criticism, is essential to enable us to act in ethically and politically just, to say nothing of intellectually honest, ways. Critical thought entails questioning, reflection and thoughtful interaction with the information and body of knowledge at hand. Education then becomes an active and constructive process of continual critical growth (Dewey, 1944).

Fundamentally, I am recommending here that evaluators, as leaders of educational reform efforts, become more critical and vigilant about the questions they are contracted to answer and about the more profound functions of education programs and practices under the rubric of a critical theory of education (Giroux, 1983b; Young, 1990; Apple & Beane, 1995; Apple 1996, Apple & Carlson, in press). In addition, these same evaluators could integrate the logic of traditional psychology with the logic of critical social psychology to begin the rethinking of education as a social project and a social process. The purpose of this rethinking is to expand on positivist traditions of considering an incremental perspective on methodological and research design issues in evaluation, into a more open critical ideology of practice and policies (Fetterman, 1988; Maruyama & Deno, 1992). These schools of thought, approaches, and particular issues should not be eliminated. We should consider these issues together with the notion that evaluation of education places us in a particularly sensitive arena within the confounds of social and human science (Fetterman & Pitman, 1986; Fetterman, 1988).

Silent Scientific Revolution

Fetterman (1988) argued that there is a silent scientific revolution in evaluation and that educational evaluation is experiencing a change in direction. A critical component of this change is a shift in the paradigms underlying the method and aim of research (Lincoln, 1986). David Fetterman further suggests that a marked shift is taking place in the professional allegiance of evaluators. This shift in allegiance, he says, is not a simple linear development. As summarized in Fetterman's book (1988), this shift goes beyond perceiving evaluation as a set of chronological transformations that travel from traditional positivist approaches to dominant qualitative forms of evaluation, including ethnography, naturalistic inquiry, generic pragmatic (sociological) inquiry, connoisseurship/criticism, and phenomenography. Rather he illustrates some significant moments of metamorphosis, revealing the process of shifting allegiance to a circular and interactive paradigmatic perspective.

Similarly, I call on evaluators to lift the blinders of methodological habit, to increase the ideological options and backgrounds available to them, to go beyond any single discipline, and to build on tradition by engaging the wisdom of critical social thought. This article is simply describing a possible interplay between the sciences and between the contemporary perspectives in evaluation.

Whether using the perspectives of accountability, knowledge or development, or any combination thereof, additional questions could be examined as the evaluation/research design is imposed on the school culture and setting (Maruyama & Deno, 1992). Critical evaluators of education reform could also listen to emerging questions that are integral to the improvement and restructuring of social projects and social processes, by attending to their own consciousness and motivations (Young,

1990). Later I will review Paolo Freire's construction of *conscientization* (Hamnett, Kumar, Porter, & Singh, 1984, p.100) to describe this experience as necessary for critical evaluators of education reform.

School and university researchers/evaluators who are taking on the challenge of restructuring schools and school systems in urban areas are involved essentially in the transformation of existing bureaucracies, bureaucracies that have had the power to control what is taught and how schools are run (Kretovics & Nussel, 1994). Clearly American education is organized in a bureaucratic form. Kretovics and Nussel confirm that at any level, national, state, or local, the traditional pyramidal, hierarchical arrangement is in effect. Proposals for reform of public schools and their evaluations must consider how the bureaucratic functionaries might respond. Since bureaucracy is "an institutionalized method of organizing social conduct in the interest of administrative efficiency," the issue of response is a genuine concern (Kretovics & Nussel, 1994).

Practicing Critical Evaluation

In the public school domain, genuine concern is adequate but more critical thought and action are needed within the world of educational bureaucracy. One way of practicing critical thought and action for critical evaluation would be to negotiate these ideological and theoretical presuppositions up front with one's clients and then deliberately confront issues of institutionalized power, democracy, and inequality in the educational programs and reform efforts. One can do this by organizing specific research designs and relationships centered around the concept of listening to the multiple voices in education and its programs. Fine and Weis (1993) witnessed and wrote about the practices and consequences of silencing in public schools. I do not think that evaluators are far from becoming partners in these implicit practices. Battling the dynamics of power and privilege that nurture, sustain, and legitimate silencing in education is the first purposeful step that a critical evaluator can take to interpret his powerful role as a transformative agent for social change. Creating flexible, authentic, and reflexive relationships with the stakeholders and with the existing bureaucracies during the process of evaluation is a second step that the critical evaluator can take towards completing a critical evaluation (Schon, 1983).

If innovative and well meaning educational programs or educational reform efforts are developed to improve the education of all students in public schools, then the evaluator of these programs has a very special and conscious role in creating opportunities for authentic discourse about these difficult issues that go beyond the successes or failures (outcomes) of children within the structural and organizational components of educational practice. The role of the evaluator and the ability to communicate and address the challenging issues such as democracy, power, and inequality to clients in the field of education, especially in the future, will be essential to transforming social activity for social change. Michelle Fine and Lois Weis (1993) included the following quote in their book:

> It is a false dichotomy which suggests that academics and/or intellectuals can only speak to one another, that we cannot hope to speak with the masses. What is true is that we make choices, that we choose voices to hear and voices to silence. If I do not speak in a language that can be understood, then there is little chance for dialogue. We must be ever vigilant. It is important that we know who we are, who we are speaking to, who we most

want to hear us, who we most long to move, motivate, and touch with our words (p.2).

The Neutrality of Schools (Social Darwinism Revisited)

Jeannie Oakes (1986) stated that in their general indictment of schools, the authors of the evaluation studies and reform reports do not attach particular importance to the fact that schools fail to serve all students equally well. Certain topics like institutionalized power, democracy, and inequality are not explicitly addressed because there is a "silenced" understanding of the status quo in educational practice. Consequently, the evaluators and reformers in the commentary made by Jeannie Oakes do not consider as targets of information or understanding the school content and processes that limit school achievement for poor and minority students. Schools, in general, are often seen as essentially neutral, and the reforms are presented as color-blind and affluence blind. Jeannie Oakes (1986, 1995) further argued that current reform efforts do not address the unequal quality of school facilities, programs, materials, counseling, expectations, and instruction. No interest is shown, for example, in the unequal distribution of competent teachers. Neither do they address school organizational changes likely to equalize access to high quality educational contexts such as desegregation, the elimination of tracking, and the reconceptualizing of vocational education programs.

Thus by extracting the logic of critical theory and critical social psychology, I would extend the "meta-evaluation" done by Oakes, in saying that the evaluators of these reform efforts are additionally hard pressed to face squarely the "silent" demons lurking behind the institutional practices in public education. Even though a common issue is made of increasing the skills and knowledge of teachers, the assumption is that teachers simply need to get better at what they've always done. Also there is an assumption that all the evaluator has to do is to evaluate how the teacher is teaching and whether the outcomes are effective learning. There is little or no mention of the need for teachers to be more knowledgeable about how poverty, racism, and limited expectations affect the educational treatment of poor and minority children (Levine, 1971; Coleman, 1981; Fine, 1983,1994). Indeed there is no direct mention and acknowledgment of these issues on any explicit level within the hierarchical structure and bureaucracy in education (Levine, 1971; Coleman, 1981; Fine, 1983,1994).

Subsequently, mainstream evaluation of these reform efforts in teaching practices and educational programs misses a crucial part of the picture about how schools are functioning for all children. If we as evaluators do not ask deliberate questions about institutionalized power, democracy, quality of instruction, and inequality within the public school domain, during the process of evaluation, then we become one more vehicle that perpetuates an already neutral state of mind about the world of education and its goals for society. While many faults are found with schools, unfairness is not one of them. In addition, the omission of these concerns and "silent" demons in evaluation and education reform efforts makes clear the prevailing conviction that schools, as they are now, are neutral places (Coleman, 1975; Oakes, 1986; Fine, 1994).

Change in American Schools

Although there is a perception that change needs to occur in virtually all American school districts, including those serving the wealthiest suburbs, the success of the reform movement will be measured ultimately by its impact on our largest most troubled public

school environments. For it is in our largest cities and our most rural districts that the job of the schools is most difficult, given the often overwhelming social and economic circumstances of students living in desperately impoverished neighborhoods (Oakes & Sirotnik, 1986). These are the neighborhoods most in need of transformed schools, and it is in these neighborhood schools that the evaluator can choose to undertake his exceptional role of being a vehicle for change and transformation.

Jonathan Kozol in *Savage Inequalities* (1991) took readers inside schools in poor neighborhoods and forced them to see the places impoverished children are compelled to go. Kozol (1991) commented on more than the physical, economic, and social inequalities among different types of school, those with affluent children, and those with children from poor homes. He addressed the very "ethos" of a school as maintained by the social-class position of the students. Theodore Sizer in *Horace's Compromise* (1984) also characterized this difference between schools quite modestly:

> Among schools there was one important difference, which followed from a single variable only: the social class of the student body. If the school principally served poor adolescents, its character, if not its structure, varied from sister schools for the more affluent. It got so I could say with some justification to school principals, tell me about the income of your student's families and I'll describe to you your school. (p.6)

Critical educators such as Michael W. Apple, Henry A. Giroux, Paolo Freire, Jeannie Oakes, Gloria Ladson-Billings, and Maxine Greene would probably agree that evaluation and research in impoverished neighborhood schools presents the critical evaluator with an exceptional challenge in social responsibility. Hence, these impoverished neighborhoods, where educational reform proponents advocate change, improvement and restructuring of schools, could be the environments that create wonderful opportunities for evaluators to maintain a critical stance toward theory, research, practice, and social policy.

Freirean Pedagogy

The statement "All men are created equal" is one that resounds throughout American history. The words are found in the Declaration of Independence and Lincoln's Gettysburg Address; they are also paraphrased and applied in numerous settings. For educators and educational evaluators, it has meant that American schools are charged with offering every child equality of educational opportunity. This concept of equality of educational opportunity is one that has been implicit in most educational practices throughout the period of public education in the nineteenth and twentieth centuries (Coleman, 1981). However, no white suburb in America would long tolerate the low academic achievement taken for granted in the urban, or rural public schools attended largely by African- Americans, Hispanics, and poor children.

In big cities all over the United States, minority students by the tens of thousands leave school each year, some as dropouts, some as graduates, utterly unprepared to participate in and contribute to a democratic society (Oakes & Sirotnik, 1986). They lack the skills that will allow them to obtain gainful employment, and they are devoid of the preparation that will lead to success in further education. Paolo Freire would characterize this lack of skills and preparation as the "inability to act upon and transform one's world" (Hamnett et al., 1984). Consequently he would say that the democratic society failed to move this person toward the ever-new possibilities of a fuller and richer

life individually and collectively through the auspices of public education (Hamnett et al., 1984).

Paolo Freire is most often recognized for his literary and practical works as an educator. His study and conduct in this field have produced radically new philosophical and political insights. His basic assumption is that people are seen to be always in the process of developing. He says that the characteristic of the human species is its repeatedly demonstrated capacity for transcending what is merely given, what is purely determined (Hamnett et al., 1984). From Freire's point of view, education, or any form of activity directed at social change, can never be neutral; it can only be used to dominate or liberate people. Although this dichotomy is limited in my opinion, these extremes serve their purpose in explaining unique ideological commitments to social change, especially as social science researchers and evaluators in education. I proposed here that evaluation of public educational programs, as a form of activity directed at social change, should follow Freire's recommendation for *conscientization*:

> Conscientization refers to the processes in which men, not as recipients, but as knowing subjects, achieve a deepening awareness both of the socio-political reality which shapes their lives and their capacity to transform that reality (Freire, 1970b).

This notion conveys the realization that nobody can help or assist others without their participation; otherwise the helper is led only to treat people as objects vulnerable to control and manipulation from outside (Freire, 1973). Here we can reflect upon what such a perspective would require in evaluation of public educational programs. Conscientization is at least one experience that critical evaluators should pass through in order to become educational leaders and change agents for educational reform.

Implications for Evaluators of Education Reform

Undoubtedly the purposes, methods and functions of evaluation would change if one was to adhere to the philosophical and ideological underpinnings of critical theory, critical social psychology, and Freirean pedagogy. The question remains: would a critical evaluator actually go beyond traditional methodological concerns to design his policy and practice to deliberately address difficult and possibly uncomfortable issues such as institutionalized power, democracy, and inequality in education? Courage, persistence and conviction are presented here as three crucial elements that will consistently be needed for critical evaluation of educational reform. In addition to these three elements a critical evaluator could benefit from continual reflection about one's own changing beliefs and landmark experiences.

The need for courage, persistence, and conviction seems fairly obvious but somehow we do not seem to talk about these character traits explicitly. Speaking out, in situations that may include numerous political and bureaucratic agendas, all with different viewpoints and axes to grind, and also insisting on the right to independence in speaking out, takes a strong stomach. Even in the political and cultural environments occurring toward the middle of the spectrum, the normal skepticism of the evaluator is unwelcome amongst the pervasive enthusiasm for one program or another. But as we move down the spectrum toward differing ideologies, doubting the conventional wisdom becomes such an offensive tactic as to deconstruct credibility and solid reputations.

It also takes fortitude or conviction and strong resistance not to succumb to political or bureaucratic blinders of one sort or another. In my experience with the higher

echelons of public education evaluation, both as a teacher and as a district based advisor of educational practice, these blinders lure evaluators into wanting to become political and "institutional" players on the national scene. There is an insidious temptation to avoid ideological and philosophical battles to the promise of glorious career rewards as compensation for obedience. This possible temptation is one of the reasons why there needs to be extensive research started in discovering the implicit and explicit social and professional ethics of different types of evaluators, especially evaluators in education reform. It takes persistence and courage to refuse sponsors the answers they want to hear, and it takes conviction and certainly conscience to ask deeper more resounding questions. Goethe said, "Possessions lost, nothing lost. Principles lost, something lost. Courage lost, everything lost" (quoted in "Visions of Public Service," 1986, p.12).

A Beginning to the Methods of Critical Evaluation

Part II: The "How" of Critical Evaluation

National policymakers, educational leaders, "public intellectuals", and children in disadvantaged situations can benefit from critical evaluation, but not in the same ways and not with the same evaluator roles. Neither more nor less activism, in my judgement, is morally superior. Various degrees of activism involve different ways to practice as an evaluator, often in different arenas. Indeed, how activist to be, involves consideration of an evaluation's purpose, decisions about intended users and uses of evaluation, and the evaluator's own values and commitments, all of which need to be made explicit. The challenge will be to create appreciation and space for such diversity among both those within and outside the profession who have a single and narrow view of evaluation and its practice. The debate will and should, go on, for that is how we discover implications and ramifications of diverse approaches, but I hope and foresee no desire to turn back the clock to a single dominant perspective.

By now, there should be no doubt as to the rationale for making a space for critical evaluation in the reform of public education. Because of the complexity of the task of reconceptualizing the evaluation process toward a process that contains an explicit normative social goal, that of social justice, and a process that is designed for purposes of fundamental change, the arguments in this section will only *begin* to delineate a preliminary path toward a methodology for critical evaluation. However, a more detailed and experienced methodology for critical evaluation would require further conceptual and empirical investigation and time. Essentially the utilization of American public school law, both state and federal statues, are combined with the adversary oriented evaluation model in order to propose briefly that these statutes can serve as merit criteria for determining the value and worth of educational programs. Critical evaluation will be augmented by commissioning the principles and rules of American public school law as additional references. Lastly, the conclusion elaborates on the roles and responsibilities of an evaluator in order to highlight the significance of our commitment and vision.

Adversary Oriented Evaluation (AOE)

Adversary Oriented Evaluation is a rubric encompassing a collection of divergent evaluation practices. In its broadest sense, the term refers to all evaluations in which there is planned opposition in the points of view of different evaluators or evaluation teams, and a planned effort to generate opposing points of view within an overall

evaluation. In 1965, Guba suggested that educational evaluation might use aspects of the legal paradigm. I am suggesting not only to use certain aspects of the legal paradigm, but also to use the state and federal statutes as merit criteria for determining the worth and value of educational programs, especially those instructional programs that serve disadvantaged students.

Next, Worthen, Sanders, and Fitzpatrick (1997) presented a provocative rationale for such an approach. If trials and hearings were useful in judging truth claims concerning patents, products, crimes, civil disobedience, and if human testimony were judged acceptable for determining life or death, as in the judicial system, then might not legal proceedings and public education law be a useful metaphor for educational program evaluation? Might there be merit in educational evaluation "trials," in taking and cross-examining human testimony, and in using the concept of advocacy to ensure that evaluation fairly examined both sides of issues?

The first self-conscious effort to follow a particular adversary paradigm was made in the early 1970's by Owens. Designed to test the usefulness of a modified judicial model, the evaluation focused on a hypothetical school curriculum and included pretrial conferences, cases presented by the "defense" and "prosecution," a hearing officer, a "jury" panel of educators, charges and rebuttals, direct questioning and redirected questions, and summaries by the prosecution and defense (Worthen et al., 1997). The reports (Owens, 1973) were intriguing to the community of evaluators and led to further conceptual and empirical work on the adversary approach. For further explanation of the development, applications, strengths, and limitations of this kind of approach see Worthen, Sanders, and Fitzpatrick (1997).

Several approaches that qualify as adversary oriented do not employ hearing processes. Kourilsky and Baker (1976) described an adversary model in which two teams prepared, respectively, affirmative and negative appraisals of that which was evaluated (the preparation stage); met to present the views to one another, cross-examining and critiquing one another's contentions on pre- specified criteria (the confrontation stage); and engaged in open-ended discussions until reconciliation of views was attained and translated into written recommendations in a single report. Levine (1974) proposed that a resident adversary or critic might be assigned to the research project to challenge each bit of information collected, searching for other plausible explanations. The Stake and Gjerde (1974) strategy of having two evaluators prepare separate reports summing up opposing positions for and against the program is yet another variant of the adversarial approach that does not depend on a hearing format. These proposals are all consistent with what Worthen et al. (1997) also called "critical evaluation."

Donmoyer (n.d) proposed a "deliberative" approach to evaluation, which focused on assessing and balancing alternative conceptions of reality and the differing value positions underlying these conceptions. "Because deliberative evaluation is primarily concerned with fostering understanding of alternative conception of reality," the evaluator's role is "to foster interaction and facilitate communication among representatives of various stakeholder groups...." (p.9-10). Donmoyer saw different world-views as the cause of underlying disputes, which could be resolved by open presentation of alternative views in some type of forum.

Worthen et al. (1997) reviewed three general approaches to adversary evaluation: (1) adaptations of the legal paradigm and other "two-view" adversary hearings, (2) adaptations of quasi- legal and other adversary hearings where more than two opposing views are considered, and (3) use of debate and other forensic structures in adversary evaluation. The third type is particularly interesting for critical evaluation purposes of

establishing merit criteria using the public education laws and codes that can serve as partial "anchors" or references for determining the quality of instructional and educational program delivery. The following is a practical representation of how the education laws and codes can be used as partial "anchors" or references.

For example, if the instructional effectiveness of programs such as bilingual education or special education was to be evaluated at a predominantly Hispanic low socio-economic elementary school in Texas, the critical evaluator could turn to the Texas Law School Bulletin (1996) for crucial information on the state's public education laws and codes that applied to the "Educational Programs" (Chapter 29, Subchapters A & B). A critical evaluation could include an investigation of the history of eligibility, assessment, enrollment, and placement into the bilingual and special education programs as defined in the Texas Law School Bulletin. Similar to the study completed by Jia Wang (1998), as mentioned previously in this article, the evaluation design would also include investigating the quality of instructional delivery, content coverage, content exposure, and content emphasis (opportunity to learn variables as described by Jia Wang, 1998).

In some instances, if the educational development of certain disadvantaged students, such as their language proficiency and academic achievement or failure were called into question, the evaluation team could review carefully the student's educational history by comparing it to the eligibility criteria, assessments, enrollment, and instructional placement education codes as set out by the Law Bulletin. These education codes could be the "anchors", the starting points or references to further the understanding of current and past campus and district based educational practices that involve high risk decision making. Education code 29.056, Enrollments of Students in Program is an example of this kind of "anchor" or reference:

> The agency shall establish standardized criteria for the identification, assessment, and classification of student of limited English proficiency eligible for entry into the program or exit from the program. The student's parent must approve a student's entry into the program, exit from the program, or placement in the program. The school district or parent may appeal the decision under Section 29.064 (p. 120).

Again, the laws and codes can be used as additional references for the evaluators to place classroom instruction, the school, the program, or the school district, in context to legal precedent and required administration. Because a public school is a governmental agency, its conduct is circumscribed by precedents of public administrative law supplemented by those legal and historical traditions surrounding educational organization that is state established, yet locally administered. In this setting legal and educational structural issues must be considered that define the powers to operate, control, and manage the schools (Alexander & Alexander, 1992).

In analyzing the American educational system and comparing it to central state systems of education in foreign countries, one is struck by the diversity of authority under which the American public schools are governed. As a federal and not a national system, the government of the United States comprises a union of states united under one central government. The particular form of American federalism creates a unique educational system, which is governed by laws of fifty states with component parts amounting to several thousand local school district operating units. Through all of this organizational multiformity, and indeed complexity, runs a legal basis on which the entire system is founded, those generally prescribed by our constitutional system.

The critical position of education in a democratic society is self-evident. Over the years the courts have come to conclude that society is best served by an educational system that teaches "through wide exposure to that robust exchange of ideas which discovers truth out of multitude of tongues. Thus because of the importance of the schools and because this robust exchange of ideas is vital to the educational process, the perpetuation of that exchange is, at all levels of the educational system, a special concern of the First Amendment" (Alexander & Alexander, 1992, p.229). No school can function appropriately as a place for the exchange of ideas unless both students and faculty enjoy an atmosphere conducive to debate and scholarly inquiry.

With this in mind, the reform of public education which includes the improvement of educational programs for those children who are least well off, should remain open to alternative views and divergent conceptions of evaluation. Critical evaluation can begin to provide an accurate analysis of the production of inequality or the reproduction of social injustice in the public schools. The ideology of critical evaluation can begin to influence a movement toward the realization of an egalitarian ideal and the elimination of inequality. I have asked educators and evaluators of education reform efforts to reconsider critically their roles in social science research, to reclaim the battleground of public school reform by focusing on the democratic purpose of public schooling, and the institutional problems in educational programs and practice that often inhibit action toward this ideal.

Conclusion

Irrespective of the many social, economic, technological, cultural, and political problems that face our American communities, the public schools exist for the purpose of educating all children. Teachers are a part of the never-ending struggle to create conditions in which learning takes place and provide the best educational opportunities in a given situation. As evaluators rendering judgement on educational programs, and giving merit or not giving merit to the educational repertoires and learning outcomes of teachers; we also become inextricably linked to the process of either perpetuating an already neutral disconnected reality of education or critically examining and observing a wide range of crucial issues, structures, and problems in contemporary education. As evaluators of education programs and teaching, we cannot ignore that we become a part of the never-ending struggle to make judgment calls about a social activity which creates the conditions or obstacles for social mobility.

The central task of the current reform movement in education is nothing less than building and transforming schools that are struggling to achieve democratic ideals (Fine, 1994). While schools can be described as potentially a site of extraordinary democracy, the processes and outcomes of schools deeply reproduce and promote the very social inequities they are said to equalize (Fine, 1983). This circumstance imposes onto the roles of educational leaders and critical evaluators a social responsibility, one that demands sincere conscience and deliberate action. Evaluators and researchers, who in the past have been content to describe dispassionately what schools are doing and how they are functioning, are actually involved in and committed to a collaborative view of knowledge creation. These characters in social change should not struggle to find a voice that sensitively captures both the insider's and outsider's view of reality. When characters, such as evaluators of educational reform, gain the conscience and purposefulness of their critical role, no relationship is left untouched or unchanged.

In conclusion, evaluation is a powerful social force that has evolved only

> recently in advanced capitalistic societies, a new institution that promises to be a major influence over the long term. Its influence can be both good and bad. In either case, society before formal evaluation is not the same as society afterward. Exactly what shape the practice, institution, profession, and discipline will take in the future is impossible to predict. What is clear is that its fate will be bound to the government and the economic structure and determined in part by its own history and traditions. Some of its destiny lies within the control of the evaluators themselves; some does not (House, 1993, p.172).

Note. Paper presented at the National Evaluation Conference, Youngstown State University, Youngstown, Ohio, September, 1998

References

Alexander, K., & Alexander, M. D. (1992). *American public school law*. (3rd ed.). St. Paul, MN: West Publishing Company.

Apple, M. W. (1996). *Cultural power and education*. New York: Teachers College Press.

Apple, M., and Beyer, L. (Eds.) (1988). *The curriculum: Problems, politics, and possibilities*. Albany: State University of New York Press.

Apple, M. W. & Beane, J. A. (Eds.) (1995). *Democratic education*. Washington, D.C.: Association for Supervision and Curriculum Development.

Apple, M. W. & Carlson, D. (Eds.) (in press). *Critical educational theory in unsettling times*. Boulder, CO: Westview Press.

Arreaga-Mayer, C., & Greenwood, C.R. (1986). Environmental variables affecting the school achievement of culturally and linguistically different learners: An instructional perspective. NABE: *The Journal for the National Association for Bilingual Education, 10* (2), 113-135.

Barth, R. S. (1986). On sheep and goats and school reform. *Phi Delta Kappan, 68* (4): 293-296.

Berliner, D.C., & Biddle B. J. (1997). *The manufactured crisis: Myths, fraud, and the attack on America's public schools*. White Plains, NY: Longman Publishers USA.

Birdsall, N., & Hecht R. (1995). *Swimming against the tide: Strategies for improving health.* Working Paper No. 305, Washington, D. C.: Inter-American Development Bank, Office of the Chief Economist.

Burke, K. (1966). *Language as symbolic action: Essays on life, literature, and method*. Berkeley: University of California Press.

Coleman, J. (1975). Racial segregation in the schools: New research with policy implications. *Phi Delta Kappan, 57*, October 1975, 75-78.

Coleman, J. (1981). Quality and equality in American education: Public and catholic

schools. *Phi Delta Kappan, 63*, 159-164.

Chelimsky, E., & Shadish, W.R. (Eds.) (1997). *Evaluation for the 21st century: A handbook*. Thousand Oaks, CA: SAGE Publications.

Cousins, J. B, Donahue, J. J., & Bloom, G. (1996). *Understanding collaborative evaluation: Results from a survey of North American evaluators*. Unpublished paper submitted for publication, University of Ottawa.

Cronbach, L. J., & Associates. (1980). *Toward reform of program evaluation*. San Francisco, CA: Jossey-Bass.

Dewey, J. (1944). *Democracy and education*. New York: The Free Press.

Donmoyer, R. (n.d.). *Evaluation as deliberation: Theoretical and empirical explorations*. (Grant No. G 810083). Ohio State University, Columbus, National Institute of Education.

Fetterman, D. (1981). Blaming the victim: The problem of evaluation design and federal involvement, and reinforcing world views in education. *Human Organization, 40* (1), 67-77.

Fetterman, D. (Ed.). (1988). *Qualitative approaches to evaluation in education: The silent scientific revolution*. New York, NY: Praeger Publishers.

Fetterman, D. & Pitman, M. A. (1986). *Educational evaluation: Ethnography in theory, practice, and politics*. Beverly Hills, CA: SAGE Publications.

Fine, M. (1983). Perspectives on inequality: Voices from urban schools. In L. Bickman, ed. *Applied Social Psychology Annual IV*. Beverly Hills: SAGE Publications.

Fine, M. (1983b). Dropping out of high school: The ideology of school and work. *Journal of Education, 165*, 259-272.

Fine, M. (1988). De-institutionalizing educational inequity. In Council of Chief State School Officers. (Eds.). *At risk youth: Policy and Research*. New York: Harcourt Brace Jovanovich.

Fine, M. (Ed.). (1994). *Chartering urban school reform: Reflections on public high schools in the midst of change*. New York: Teachers College Press.

Fine, M., & Weis, L. (Eds.). (1993). *Beyond silenced voices: Class, race, and gender in United States schools*. Albany, NY: State University of New York Press.

Freire, P. (1970b). Cultural action and conscientization. *Harvard Educational Review, 40* (3), 452- 477.

Freire, P. (1973). *Education for critical consciousness*. New York: Seabury.

Freire, P. (1985). *The politics of education*. South Hadley, MA: Bergin and Garvey.

Giroux, H. A. (1997). *Pedagogy and the politics of hope: Theory, culture, and schooling*. Boulder, CO: Westview Press.

Giroux, H. A. (1983). *Theory and resistance in education: A pedagogy for the opposition*. South Hadley, MA: Bergin and Garvey.

Giroux, H. A. (1983b). *Critical theory and educational practice*. Geelong: Deakin University Press.

Giroux, H. A. (1988). *Teachers as intellectuals: Toward a critical pedagogy of learning*. South Hadley, MA: Bergin and Garvey.

Giroux, H. A. & Aronowitz. S. (1991). *Postmodern education: Politics, culture, and social criticism*. Minneapolis, MN: University of Minnesota Press.

Gross, S. (1993). Early mathematics performance and achievement: Results of a study within a large suburban school system. *Journal of Negro Education, 62* (3), 269-287.

Guba, E. G. (1965). *Evaluation in field studies*. Address at evaluation conference sponsored by the Ohio State Department of Education, Columbus.

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: SAGE Publications.

Habermas, J. (1971). *Knowledge and human interest*. Translation. Boston: Beacon Press.

Habermas, J. (1981). *The theory of communicative action, volume 1: Reason and the rationalization of society*. Boston: Beacon.

Habermas, J. (1987). *The theory of communicative action, volume 2: Lifeworld and system, a critique of functionalist reason*. Boston: Beacon.

Habermas, J. (1990). *Moral consciousness and communicative action*. Cambridge: MIT Press.

Hamnett, M. P., Kumar, K., Porter, D. J., & Singh, A. (1984). *Ethics, politics, and international social science research: From critique to praxis*. East-West Center: University of Hawaii Press.

House, E. R. (1976). Justice in evaluation. In G.V. Glass (Ed.), *Evaluation studies review annual* (Vol. 1, pp.75-100). Beverly Hills, CA: SAGE Publications.

House, E. R. (1980). *Evaluating with validity*. Beverly Hills, CA: SAGE Publications.

House, E. R. (1988). *Jesse Jackson and the politics of charisma: The rise and fall of the PUSH/Excel program*. Boulder, CO: Westview.

House, E. R. (1990). Methodology and justice. In K. A. Sirotnik (Ed.), *New directions for program evaluation, Vol.45*, (pp. 23-36). San Francisco: Jossey-Bass.

House, E. R. (1993). *Professional evaluation*. Newbury Park, CA: SAGE Publications.

Howard, G. S. (1985). The role of values in the science of psychology. *American Psychologist, 40*, 255-265.

Kanpol, B. (1997). *Issues and trends in critical pedagogy*. Cresskill, NJ: Hampton Press.

Kohlberg, L. (1984). *Essays on moral development: The psychology of moral development*. Vol.2. New York: Harper & Row.

Kourilsky, M. & Baker, E. (1976). An experimental comparison of interaction, advocacy, and adversary evaluation. *Center on Evaluation, Development, and Research (CEDR) Quarterly, 9*, 4-8.

Kozol, J. (1991). *Savage Inequalities*. New York: Crown Publishers.

Kretovics, J., Farber, K., & Armaline, W. (1991). Reform from the bottom up: Empowering teachers to transform schools. *Phi Delta Kappan, 73* (4), 295-299.

Kretovics, J., & Nussel, E. J. (1994). *Transforming urban education*. Needham Heights, MA: Allyn and Bacon.

Kuhn, T. (1970). *The structure of scientific revolutions*. 2nd Ed. Chicago: University of Chicago Press.

Levine, D. U. (1971). Concepts of bureaucracy in urban school reform. *Phi Delta Kappan, 52*, 329-333.

Levine, M. (1974). Scientific method and the adversary model. Some preliminary thoughts. *American Psychologist, 29*, 661-677.

Lincoln, Y. S. (Ed.). (1986). *Organizational theory and inquiry: The paradigm revolution*. Newbury Park, CA: SAGE Publications.

Lincoln, Y. S. (1991). The arts and sciences of program evaluation. *Evaluation Practice, 12* (1), 1-7.

Madaus, G. F., West, M.M., Harmon, M.C., Lomax, R G., & Viator, K.A. (1992). *The influence of testing on teaching math and science in grades 4-12*. Boston: Boston College Center for the Study of Testing, Evaluation, and Education Policy.

Maruyama, G. M., & Deno, S. (1992). *Research in educational settings*. Newbury Park, CA: SAGE Publications.

McCormick, L., Haring, N. G., & Haring, T.G. (1990). *Exceptional children and youth*. (6th ed.). New York, NY: Macmillan Publishing Company.

McLaughlin, M. W. (1975). *Evaluation and reform: The elementary and secondary education act of 1965*. Cambridge, MA: Ballinger.

MacPherson, C. B. (1987). *The rise and fall of economic justice*. Oxford, UK: Oxford University Press.

Miller, S. I., and Safer, L. A. (1993). Evidence, ethics, and social policy dilemmas. *Education Policy Analysis Archives, 1* (9) (Available online at http://epaa.asu.edu)

Murray, C. A. (1983). Stakeholders as deck chairs. In A. Bryk (Ed.), *New directions for*

program evaluation: Vol. 17. Stakeholder-based evaluation (pp.58-61). San Francisco: Jossey- Bass.

Murray, C. A. (1984). *Losing ground: American social policy, 1950-1980*. New York: Basic Books.

National Council on Education Standards and Testing (NCEST). (1992). *Raising standards for American education: A report to Congress, the secretary of education, the national education goals panel, and the American people*. Washington, D.C.: National Council on Education Standards and Testing.

Nevo, D. (1986). The conceptualization of educational evaluation: An analytical review of the literature. In E. R. House (Ed.), *New directions in educational evaluation*.(pp. 15-29). London, UK: Falmer Press.

Noll, J. W. (1997). *Taking Sides: Clashing views on controversial education issues*. (9th ed). Guilford, CT: McGraw-Hill.

Oakes, J. (1986). Tracking, inequality, and the rhetoric of reform: Why schools don't change. *Journal of Education, 168* (1), 60-79.

Oakes, J. & Sirotnik, K. A. (Eds.). (1986). *Critical perspectives on the organization and improvement of schooling.* Hingham, MA: Kluwer-Nijhoff.

Oakes, J., & Guiton, G. (1995). Opportunity to learn and conceptions of educational equality. *Educational Evaluation and Policy Analysis, 17* (3), 323-336.

Ogbu, J. U., & Matute-Bianchi, M. E. (in press). Understanding socio-cultural factors: Knowledge, identity, and school adjustment. In California State Department of Education (Ed.), *Socio-cultural factors and minority student achievement*. Sacramento: Author.

Owens, T. R. (1973). Educational evaluation by adversary proceeding. In E. R. House (Ed.), *School evaluation: The politics and process*. Berkeley, CA: McCutchan.

Patton, M. Q. (1994). Developmental evaluation. *Evaluation Practice, 15*, 311-319.

Patton, M. Q. (1997). *Utilization Focused Evaluation* (3rd ed.). Thousand Oaks, CA: SAGE Publications.

Paul, S. (1991). *Accountability in public services: Exit, voice, and capture*. Washington, D.C.: The World Bank.

Prilleltensky, I. (1994). *The morals and politics of psychological discourse and the status quo*. Albany: State University of New York Press.

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Belknap Press.

Rossi, P. H, & Freeman, H. E. (1993). *Evaluation: A systemic approach* (5th ed.). Newbury Park, CA: SAGE Publications.

Sabia, D. R., Wallulis, J. (Eds.). (1983). *Changing social science: Critical theory and*

other critical perspectives. Albany, NY: State University of New York Press.

Schön, D. A. (1983). *The reflexive practitioner: How professionals think in action*. New York: Basic Books.

Scriven, M. (1986). Evaluation as a paradigm for educational research. In E. House (Ed.), *New directions in educational evaluation*. (pp.53-67). Philadelphia, PA: Farmer Press.

Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Newbury Park, CA: SAGE Publications.

Scriven, M., & Kramer, J. (January 1995). Risks, rights, and responsibilities in evaluation. *Australasian Journal of Evaluation 15*, 15-19.

Sechrest, L. (1992). Roots: Back to our first generations. *Evaluation Practice, 13* (1), 1-7.

Shadish, W. R. Jr., & Epstein, R. (1987). Patterns of program evaluation practice among members of the evaluation research society and evaluation network. *Evaluation Review, 11*, 555-590.

Shapiro, H. S., & Purpel, D. E. (1998). *Critical social issues in American education.* (2nd Ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Sirotnik, K. A. (Spring, 1990). Evaluation and social justice: Issues in public education. In K. A. Sirotnik (Ed.), *New directions for program evaluation, Vol.45*. San Francisco: Jossey-Bass.

Sirotnik, K. A. & Oakes, J. (Spring, 1990). Evaluation as Critical Inquiry: School improvement as a case in point. In K. A. Sirotnik (Ed.), *New directions for program evaluation, Vol.45*. Evaluation and social justice: Issues in public education. (pp. 37-59). San Francisco: Jossey-Bass.

Sizer, T. (1984). *Horace's compromise: The dilemma of the American high school.* Boston, MA: Houghton Mifflin.

Stake, R. E. (1986). *Quieting reform: Social science and social action in an urban youth reform*. Champaign: University of Illinois Press.

Stake, R. E., & Gjerde, C. (1974). An evaluation of T- CITY, The Twin City Institute for Talented Youth. In R. H. P. Kraft, L. M. Smith, R. A. Pohland, C. J. Brauner, & C.Gjerde (Eds.), *Four evaluation examples: Anthropologist, economic, narrative, and portrayal* (AERA Monograph Series on Curriculum Evaluation No. 7). Chicago: Rand McNally.

Tsoi Hoshmand, L. L. (1994). *Orientation to inquiry in a reflective professional psychology*. Albany: State University of New York Press.

Texas Education Agency. (1996). *Texas School Law Bulletin*. Austin, TX: West Publishing Company.

Vervoort, C. E. (1975). Onderwijs en maatschappelijke ongelijkheid als verdelingsprobleem (Education and social inequality as a distribution problem). In P. van der Kley & A. A. Wesselingh (Eds.). *Onderwijs en maatschappelijke ongelijkheid.* Boekaflevering Mens en Maatschappij, Jrj, 50. Rotterdam: UPR.

Visions of public service. (1986, Fall/Winter). *JFK School of Government Bulletin.*

Warren, R. L. (1963). *Social research consultation*. New York: Russel Sage.

Wertsch, J. V. (1998). *Mind as action*. Oxford: Oxford University Press.

Wesselingh, A. (1997). Spheres of justice: The case of education. *International Studies in Sociology of Education*, 7(2), 181-194.

Wexler, P. (1983). *Critical social psychology*. Boston, MA: Routledge & Kegan Paul.

Winfield, L. F. (1993). Investigating test content and curriculum content overlap to assess opportunity to learn. *Journal of Negro Education, 62* (3), 288-310.

Worthen, B. R., Sanders, J. R., & Fitzpatrick, J. L. (1997). *Program evaluation: Alternative approaches and practical guidelines*. (2nd ed.). White Plains, NY: Longman Publishers.

Young, R. E. (1990). *A critical theory of education: Habermas and our children's future.* New York: Teachers College Press

Note: I would like to thank Dr. Patricia Whang and Dr. Cynthia Reed for their advice and support during the writing of this article.

About the Author

Gisele A. Waters

Educational Foundations Leadership and Technology
3084 Haley Center
College of Education
Auburn University
Auburn, AL 36849

Email: waterga@mail.auburn.edu

Education:

Student studying for Ph.D. in Educational Psychology
M.Ed. Special Education/Educational Psychology, 1996. University of Houston
B.A. Economics, 1990. University of Texas at Austin
Certification: Texas; Generic Special Education (K-12),

Bilingual and ESL Education, Elementary Education

Working in Texas as a public school teacher for six years in bilingual/ESL

education classrooms, and special education resource significantly influenced my critical analysis of the educational process. During that time, I held delegate positions of leadership on district and campus advisory teams, sharing the responsibility in the development and implementation of district and campus level improvement plans. The obstacles encountered to implement research based best practices carefully positioned my observations and questions about teaching and learning. Ultimately, my heart lies with teachers and children and the institutional pressures that affect them. Today my teaching at the undergraduate level reflects an instructional approach that frames my passions for issues of social justice, democracy, power, voice, and equity in schools and schooling. My methodological interests lie in the cultivation of a critical social science, a science intended to empower those involved to change as well as to understand their world.

The World Wide Web address for the *Education Policy Analysis Archives* is http://olam.ed.asu.edu/epaa

General questions about appropriateness of topics or particular articles may be addressed to the Editor, Gene V Glass, glass@asu.edu or reach him at College of Education, Arizona State University, Tempe, AZ 85287-2411. (602-965-2692). The Book Review Editor is Walter E. Shepherd: shepherd@asu.edu . The Commentary Editor is Casey D. Cobb: casey@olam.ed.asu.edu .

EPAA Editorial Board