RECLAIMING EVALUATION
EVALUATION'S RESPONSE TO URGENT PUBLIC POLICY ISSUES

Oregon Program Evaluators Network
Portland, Oregon
Keynote Address, October 2004

Jennifer C. Greene
University of Illinois, Urbana-Champaign

These are troubled times, perhaps more so than others, perhaps not. It doesn't really matter, as comparisons are not the point. Intrinsically, these are troubled times. Our world is terrorized by religious, ethnic, and nationalistic warfare. And our world is terrorized by the huge numbers of people who live in unimaginable poverty – without clean water, adequate food, stable housing, basic safety, much less education, health care, or meaningful employment. Our earth is scourged by chemicals, by the effluvia of modern industry, by burn-and-slash agriculture conducted by farmers have no other way to feed their families. Our earth also suffers from neglect, from mindsets preoccupied with short-term profits instead of long-term sustainability, from a privileging of people over other living creatures. And our nation is fractured in so many damaging ways – most profoundly by race, ethnicity, and class. We live in multiple America's – many of us materially comfortable and intellectually rewarded with interesting jobs, yet many others living paycheck to paycheck or living without paychecks with other family members or in homeless shelters or in abandoned cars or in a ragged sleeping bag under a highway overpass. Many of us own homes in lovely neighborhoods, with stately tree-lined streets, municipal parks, and neighborhood schools or in spacious apartments with security guards and access to elite private schools, yet many others rent rooms in asphalt jungles or parched farmlands, with a few broken swings in playgrounds of schools forsaken by

just about everyone. The divisions fracturing America today, and indeed our global community as well, are visibly and uncompromisingly demarcated by color and class.

- When was the last time you saw a group of young African-American males walking toward you on a city street, and you panicked and rushed across traffic to the other side of the street?
- When was the last time you stood in a service line – like at a post office – frustrated and exasperated, as a limited-English speaking person struggled to get her business done at the counter while everyone else waited?
- When was the last time you wondered – perhaps secretly in the privacy of your shower, or perhaps more publicly – about the fairness of some affirmative action policies?
- When was the last time you nervously hunkered down in your car, rolling up all the windows and locking all the doors as you entered a run-down part of a city that looked like it was rife with drugs and crime and criminals?

These and other challenges of our contemporary world are manifest in multiple spaces of our lives – in private life spaces characterized by mistrust and suspicion; in public spaces dominated by 30-second sound bytes, lots of blaming of others, and few instances of people stepping up and taking responsibility for their actions and for the principles and beliefs that underlie them. Contemporary challenges are also manifest in relational or social spaces filled with fear of the "other" and with divisive alliances among the "same." In my view, it is this private mistrust, this public ducking of responsibility, and this relational fear and divisiveness – it is these underlying moral-

ethical dimensions of human action and interaction -- that are so seriously frayed and troubled today, and thus compel our urgent attention and intelligence.

And how are these challenges and issues manifest in the field of evaluation? How are evaluators responding to these private, public and relational challenges so evident in global, national, and local troubles? Where are we located? Where should we be?

How contemporary challenges and urgent public issues are manifest in evaluation

I will respond to these important questions by focusing on three ways in which urgent public issues currently intersect with the field of evaluation: the current infatuation with accountability (related to mistrust and suspicion), evaluation's relationships to government decision making (related to who is responsible, anyway?), and challenges of diversity and difference in our practice (related to fear and divisiveness). These are inter-connected, to be sure, but each engages somewhat different facets of today's challenges and of evaluation's multiple countenances in response to these challenges.

Let me start with some snapshots of contemporary evaluation practice

These snapshots convey what is evaluation is currently like, in many though of course not all contexts. My talk will critique these current practices and offer some alternatives, alternatives that I believe are more responsibly responsive to urgent public issues.

At the local Even Start office …

Mary Williams gave a deep sigh as she stared out of her office window. Today in the intensely blue sky, beautiful white billowy clouds tumbled by on a quiet breeze, coloring the sparsely populated countryside with a checkerboard of light and shadow. Mary had just returned from another visit to Sally Jo Hitchcock, who was the mother in one of Mary's caseload of Even Start families. Mary really liked the whole-family literacy emphasis of the Even Start program. It made so much more sense to help both parents and young children gain in literacy skills, instead of concentrating on just one or the other. But, in Sally Jo's case, while Mary could accomplish some storybook reading with the two pre-schoolers in the family, she couldn't even begin to engage Sally Jo in any literacy activities. Every time Mary visited, Sally Jo was having some kind of crisis – usually to do with getting food or taking one of the children to the hospital. (The hospital emergency room was the family's doctor.) Amidst these crises, Mary really couldn't do anything except try to help Sally Jo in the moment – provide resources and support to address the immediate crisis and also try to stave off future ones. With families like Sally Jo's, Mary felt more like a social worker than an educator. In truth, that was OK with her. She liked helping people with whatever the occasion demanded. Yet, it was almost time for the required 6-month literacy assessment for Sally Jo, and Mary knew that this mother wouldn't show any progress on this assessment. In fact, she might even refuse to take it!! And Mary's supervisor wouldn't like this at all! "Oh dear," murmured Mary to herself, "how do I meaningfully help Sally Jo and her family while also doing the literacy testing that I am supposed to do? What's the right thing to do in this situation?"

In a middle school teacher's lounge … (adapted from Crockett, in preparation)

The faculty of the Union Township Middle School were discussing their professional development agenda for the year, chattering excitedly about a new possibility – that of adapting the Japanese "lesson study" idea as a central part of this agenda. In lesson study, small groups of teachers work collaboratively to improve identified challenges in their teaching – for example, how do students learn about reactions in chemistry or about graphing in math. Teachers identify a possible change in their practices, try it out, observe each other's classrooms, gather data on student performance and attitudes, and then meet to reflect on what was learned. Teachers then revise again their instructional ideas and strategies and repeat the whole process. Many teachers in Union are attracted to this possibility, as they highly value opportunities to learn from each other, and the teachers' discussion is animated and engaged. But, the school principal abruptly interrupts this discussion to wonder aloud, "Do you really think that lesson study will help raise students' test scores?" Upon hearing this, teachers slump visibly in their seats, reminded of the extraordinary pressure they are experiencing for higher test scores. This pressure, recently codified in the *NCLB* legislation, has noticeably sapped some of the life out of the school, certainly in the experiences of the teaching staff. Then the principal says, "Let's go back and reconsider learning more about test coaching skills and strategies for our professional development plan. *I* think that has to be our priority at this time."

At a downtown café ... (adapted from Hellenger, 2004)

Boyd Johnson was really excited about this new program for juvenile offenders in Willowbrook, a mid-sized city in the Midwest. Instead of serving time in the county juvenile jail, youth who commit minor offenses are able to participate in a program of intensely-supervised community service and education. The program is designed to significantly reduce the current overcrowding of the county juvenile facility as well as interrupt and redirect the negative life trajectory of many of these youth. Boyd had been a probation officer in the juvenile justice system for about 10 years. For the past year, he had been given the chance to work with this new program. While having a double mocha latte in his favorite downtown café, Boyd was reading over a first-year evaluation report about the program. And he was disturbed by some of what he read, but actually more by what he didn't read. From Boyd's point of view, some of the biggest challenges with the program so far had to do with prejudicial attitudes of some people in the community toward the kids in the program and with the kids getting along with one another. Although they were all in the system together, the east side and west side kids brought very different background experiences and future possibilities to the program. For Boyd, getting these kids to know something about each other's lives was a really exciting potentiality of the program. But, none of this was even mentioned in the report!

In the name of science … evaluation in the US Department of Education, 2004 …

People say, our schools are sad
Kids don’t learn, the teachers are bad.

Parents don’t care, parents don’t know.
Kids don’t learn, test scores are low.

So how do we help? What do we do?
My kid is in school, what about you?

Should I volunteer?
Try not to fear?
Try instead to trust?
No, says Russ (Whitehurst, director of IES)

What you need to do is ….

Randomization
Experimentation
You can forget about … imagination
Show the door to … deliberation
Shove in your locker … socialization
But give a big hand for … memorization
Bring ‘em on … those examinations.

People say, our schools are failing.
Teachers are flailing and parents are wailing.

Kids can’t read, kids can’t add
Kids skip school, kids act bad.

Kids who are poor, whose skin is not white
They do the worst, and this is not right.

So how do we help? What do we do?
My kid is in school, what about you?

Should I volunteer?
Try not to fear
Try instead to trust?
No, says Russ (Whitehurst, director of IES)

What you need to do is ….

Randomization
Experimentation
You can forget about … imagination
Show the door to … deliberation
Shove in your locker … socialization
But give a big hand for … memorization
Bring 'em on … those examinations.

What do these snapshots suggest about the character and position of evaluation today? Where is evaluation located amidst global turmoil, national divisiveness, and local disharmony -- amidst private spaces of mistrust and suspicion, public spheres of irresponsibility and blaming, and relationships frayed by fear and strangeness -- amidst these contemporary challenges of accountability, responsibility, and diversity? What are some problems with evaluation's current location? And then, how can we re-locate evaluation and thereby *reclaim* its more engaged and committed potential? My argument is one of reclamation. I will turn now to it.

Where is evaluation located in today's troubled world?

I will start with accountability.

<u>To whom is the evaluator accountable? And for what?</u>

We are in the midst of another cycle of public demands for accountability – captured in such concepts as the new public management, performance management, results-based management, as well as outcomes measurement and performance measurement, and perhaps, stretching just a bit, also evidence-based practice and evidence-based decision making. Evaluation is intrinsic to these ideas and systems, but evaluation is not in its best form in the accountability systems that currently exist. I would like to *reclaim* from the

narrowness of these systems evaluation's grander potential for generating meaningful, respectful, and useful knowledge, understanding, and insight.

Each of the scenarios I just described contains strands of the contemporary infatuation with accountability, especially the first two scenarios. Both Even Start case worker Mary and the Union Township Middle School teachers were discouraged from engaging in what they perceived to be worthy and important activities due to demands from accountability systems in their work environments. Mary found it hard to balance the raw and immediate needs of Sally Jo with her responsibilities for insuring measurable progress on periodic literacy assessments. And the Union teachers were not able to engage in an exciting professional development activity, lesson study, because of acute pressures to raise kids' test scores under *NCLB*.

These kinds of accountability systems arose in the 1980s and 1990s in multiple sectors of westernized democracies for good reasons, I think. In the US, decades of social spending during the latter half of the 20th century yielded few visible results. Social problems appeared intractable. It was time to concentrate not just on providing services to those in need but also on understanding what difference those services made in the lives of program participants. And if a given program did not make a meaningful difference, it was time to find a new one. It was time indeed to concentrate on measurement of outcomes and results.

In previous work, I and others have noted our complete agreement with this underlying rationale for new systems of performance measurement and outcomes accountability, alongside our considerable disagreement with the ways in which these systems have been fashioned and implemented (Greene, 1999, 2001). These are now

familiar discussions to many of us. They include concerns about the narrow and limited definitions of outcomes used in these systems, their disregard of context and process, their privileging of a managerial perspective on program quality, their adoption of market-oriented business models and values for the public sector, and their coupling of performance results with employment security especially for front-line staff and managers, yet much less so for those higher up in an organization.

In brief, most of the accountability systems of today emphasize the "directing and controlling" function of accountability, to the serious neglect of its "understanding and learning" function and potential (Benjamin, 2004). The former requires laws, rules, and contracts along with monitoring, sanctions, and other external control mechanisms, while the latter rests on mutual deliberation and collective action along with trust, respect, and shared responsibility. The former rests on rationality; the latter on learning. The former risks being punitive; the latter can be generative. The former is usually a top-down activity, the latter usually reciprocal.

We can do better than this. We can ask important accountability questions without giving up or sacrificing what we know are multiple and complex facets of program contexts and therefore of program quality and effectiveness. We can find out about important program outcomes while also learning about important program processes and practices, thereby preserving evaluation's vital educative roles and responsibilities (following the legacies of Lee Cronbach and Carol Weiss). We can develop accountability systems that are not punitive or authoritarian, but rather respectful and affirming. We can, in other words, "summon the best countenance of accountability" (Greene, 2001), rather than only its controlling aspects.

We in the evaluation community have the understandings and the strategies needed to engage meaningfully with, rather than simply satisfy, contemporary demands for accountability. We can contribute significantly to imagining, developing and implementing accountability systems that offer meaningful learning and that therefore resist simplification, memorization, and sole reliance on examination. Our understandings and strategies include:

- Insisting that evaluation fulfill its important educative mission, with particular emphasis on understanding and improving practice
- Ensuring that our work values and affirms the importance of context and complexity in human action and interaction
- Getting serious about evaluating our own practices, making meta-evaluation a more routine and more transparent part of our own processes – routine in that we all do it regularly and transparent in that we explicate and make public not only the technical and methodological criteria that matter for good evaluation work, but also the political and values-based principles that underlie our practice decisions and actions

I will return to these ideas after discussing the other two intersections of contemporary issues with our field of evaluation (responsibility and diversity).

<u>What are the citizenship responsibilities of evaluators</u>? (partly from Greene, in press)

Many in our evaluation community position our work on the sidelines of decision making and deliberation about important public issues. I get as upset about this as I do when my now-18 year old daughter would choose to be a spectator of some athletic event

peopled by her male peers. I would say, “You are a player, not just a spectator. You belong on the field, not the sidelines!” To be sure, the dynamics are different with evaluation, but the point about being a player is quite a bit the same. The issues relate to action and engagement, as well as to responsibility, citizenship, and democracy. I would like to *reclaim* evaluation’s potential *and* responsibility to be an active player in important deliberations about important public issues, to envision ours as a public craft with civic responsibilities and commitments.

Many social and political critics continue to lament the absence of moral and ethical considerations from contemporary public decision making. We remain centered in a modernist era, anchored in the Enlightenment -- an era where technical rationality rules, where scientism reigns, where we have endless new electronic gadgets and toys, but we have lost our moral compass. Our pathways and decisions are scrupulously reasoned but empty of the moral courage required to include in these decisions considerations of what is “right to do and good to be,” in the words of my colleague Tom Schwandt.

Evaluation is squarely in the center of this contemporary controversy over what counts in public decision making and who is responsible, for our social scientific legacies have long honored technical soundness, objectivity, and value-neutrality. Our job traditionally has been to contribute “truth and facts” to public debates about public issues, and thereby make them more rational and grounded in empirical reasoning. Our traditional location has been one apart from the political fray of democratic decision making – a sheltered, protected position on the sidelines.

Yet, as we know from our own immersion in the philosophical challenges to this way of thinking about and framing science – best encapsulated as the “quantitative–qualitative

debate" – social science is not and cannot be value free. Evaluative knowledge claims are also evaluative value claims. Our claims to know about a given social program or policy are made from within a particular value stance – one that privileges participatory democracy, liberal pluralism, technical progress, utilitarianism, or social betterment. In addition, as we have learned from Ernie House, the practice of evaluation is not separable from the socio-political practices and institutions to which it is designed to contribute or in which it is embedded. There are no viable sidelines in democratic political discourse. The very activity of generating evaluative knowledge about social programs and policies helps to constitute the form and function of this discourse. An evaluative practice that divorces technical rationality from moral responsibility helps to shape democratic institutions that do the same. And I don't think we want to do this.

Moreover, evaluators should not be absolved from the moral and ethical responsibility for the practice choices we make and the knowledge claims we generate. Rather, we need to think of our work as a *public* craft. Following William Sullivan, we need to aspire to an ethic of civic-mindedness in our profession, of using our scientific expertise and the power and privilege of our status for the common good, not just for self-interested gain. Sullivan bemoans the technocratic takeover of many professions in the US during the past century, whereby utilitarian and individualistic ideals stripped professionals of their public responsibilities, leaving only de-contextualized, self-interested technical expertise. He presents an eloquent and persuasive argument that by reclaiming their civic and moral agenda, professionals can importantly help to reverse the current unraveling of the social fabric, the decline of social trust, and the fraying of civic bonds. Professionals need to be a part of the democratic civic order, argues Sullivan.

They need to recover their commitment to the public good, a commitment that extends beyond technical competence to moral and social issues of trust, equity, and civic cooperation. Genuine professionalism requires engagement with moral and civic aims, as well as technical means to reach those aims.

How can we do this? What are some important understandings and strategies that we have as evaluators in this regard?

- Most importantly, we can use our experience and understanding of participation and collaboration in evaluation to re-position our work as conducted in *partnership* with evaluation sponsors, particularly in government, rather than in service to such sponsors. This, I believe, would heighten our own civic responsibilities for our evaluative work and judgments, as well as make more public the decisions and judgments of government.
- More serious and thoughtful meta-evaluation will also contribute to the idea of re-positioning evaluation as a public craft and thereby reaffirming our citizenship responsibilities therein.

Again, I will return to these ideas after discussing the third domain of contemporary challenge to our work and our position in the world (diversity).

(See also McKie, *Evaluation, 9*, 307-324)

<u>What is evaluation's moral commitment to engaging with difference and diversity</u>?

(drawn from Greene, in press)

The challenges of diversity and difference are central to the contemporary landscape of American community life. There is no more urgent public issue than healing the racial and class fractures of our nation, seriously committing to overcoming our ugly legacies of

prejudice, and striving to realize our wonderfully ambitious national dream of a huge and wealthy country that does indeed welcome into each community and make a place within each neighborhood for the tired, the hungry, and the poor from around the globe. It is inarguably in the public interest in contemporary American society, and indeed in the broader global society, to engage with our diversity, to learn how to live with, appreciate, and accept our differences. I would like to *reclaim* evaluation from the wastelands of value neutrality and recommit our work to engaging with these vital national issues.

This part of my reclamation agenda has two main points – first, the meanings of difference and diversity, and second, the importance and meanings of context.[1]

Let's start with meanings of difference and diversity. What I mean by "difference" in this argument includes *both* the historical markers of "disadvantage" in our contemporary society – notably, race, ethnicity, class, gender, sexual orientation, able-bodiedness, and cultural traditions – *and* the infinitely astonishing other ways in which human beings are different from one another. That is, difference is defined in ways that both acknowledge and honor, as well as trouble and disrupt existing social categories that mark historical disadvantage. Let me elaborate on these two key meanings of difference.

Clearly, the process of social categorization – sorting and labeling people by race, ethnicity, class, gender, age, able-bodiedness, and so forth – distorts the meanings of diversity and often converts good intentions into prejudicial actions. Yet, to simply erase or ignore such meanings of diversity is also politically naïve. Moreover, it devalues the profound injustices suffered by those who rights have been seized in the past. So, some of the meanings of difference to be engaged are these very social categories – but not as

[1] The good work on contextually and culturally responsive evaluation being done by people like Stafford Hood, Rodney Hopson, Veronica Thomas and colleagues at Howard, Lois-ellin Datta, Fiona Cram and other Maori New Zealanders is directly relevant to my remarks.

fixed or essentialized ways of being in the world, rather as multifaceted, situated, dynamic, and socially constructed dimensions of experience and identity.

To illustrate, "the" African American experience in the US today is anchored in the shared ancestry of slavery and the continuing racism of our society, but is otherwise differentially mediated by all aspects of context and by the interaction of context with the unique characteristics of each black American today. So, there are as many African American experiences as there are African Americans. To meaningfully engage with difference then is to listen well to the story of each African American and to listen for its own contextualized particularities.

Beyond these standard social categories marking difference, there are many other important dimensions of difference in human experience, dimensions that extend well into the infinite constellations of characteristics, abilities, and aspirations that make each human unique. To meaningfully engage with difference, therefore, is to see beyond, within, and through social categories to that person's shyness, wit, spirituality, creativity, kindness, and the next person's and the next and the next.

And now let me turn to the second point about how to engage with contemporary challenges of diversity in evaluation – and that is the importance of seriously attending to context in our work. It is fully recognized in the US today that most inequities in our society today arise from historical tragedies and exploitations. And it is fully recognized that people with dark skin, a native language other than English, an income below a livable wage, a culture that is not Anglo or western European, a religion that is not Judeo-Christian live in communities and neighborhoods that are "different" from mainstream (white, middle class) American culture. To attend to difference and diversity in

evaluation is thereby to attend to context in all of its varied cultural richness, to be respectful of and responsive to the unique features of the particular contexts in which our work is located. Further, to attend to difference and diversity in evaluation is to acknowledge and understand that the very meaning and quality of a given program is embedded in and constituted by its context. Decontextualized questions like, "how good is this health care program?" make no sense. Instead, evaluators need to ask questions like, "how good is this program for these people with this disease in this community at this time?" Moreover – and this is the critical point here – definitions of goodness themselves become contextualized; that is, critical features of the context become embedded in or intertwined with meanings of quality and thus with standards or criteria for judgment. And so the program gets judged by how well it responds to these sick people in this place at this time, rather than asking how well these people succeed in the program. This is a profound reversal of figure and ground (Kushner, 2000).

Boyd Johnson in the earlier vignette was troubled because the evaluation report he was reading was not even attentive to, much less engaged with some of the program's and context's challenges with difference and diversity. How could these be missed, he wondered? What are these evaluators attending to, if not these critically important challenges of difference and diversity?

In contrast, evaluation that is engaged with challenges of difference and diversity is evaluation that privileges attention to what differences are manifest in a given context and how those differences matter to effective pedagogy, to quality health care, to meaningful employment, to fairness in court, and so on. Again, this engagement can be enacted in two main ways:

- By asking how well the program being evaluated serves the people participating in it, rather than asking how well the participants succeed in the program. Contextually responsive and engaged evaluation, that is, insists that educational and social programs be judged, at least in part, by how well they respect and advance the well being of children, youth, and families who are traditionally ill-served by our educational and social systems. Another way of saying this is, is this the right program for these people? Are these the right goals, objectives, materials, activities? Who decided this? Who should participate in such decisions?
- And by asking the meta-evaluative question again, this time in terms of judging evaluation in ways that centrally highlight and legitimize diversity in context, which again refers to both historical conditions of advantage and disadvantage in our society, as well as the abundantly magnificent other ways in which humans are different from one another. That is, evaluation becomes accountable to democratic ideals like social justice and equity, and evaluations get judged by how well they advance these democratic ideals (Greene, Millett, and Hopson, in press).

Interlude … Can evaluation save the world? (Stake, 2004)

It's about here in my talks, or actually it's usually about 15 minutes before this, that someone, or likely some many, wonder something like this. What is Jennifer talking about anyway? Evaluation is a modest activity (see Weiss, 199x), aimed primarily to help inform our understandings of educational and social interventions and the policies that

shape them. Evaluators are not radical social activists, nor should they be. Rather, evaluators are mere social scientists with good technical training and an interest in, even commitment to social betterment or incremental social change or making the world a better place. What is all this talk about value engagement, the public good, contextual responsiveness, social justice anyway?

My highly esteemed colleague, Bob Stake, recently asked the same question in a paper he titled, "How far dare an evaluator go toward saving the world?" (presented at AERA in San Diego, April 2004). Bob, like many others, has long been troubled by the presence and character of advocacy in evaluation. In this brief paper, Bob revisits his concerns about advocacy in evaluation. He begins his reflections by acknowledging that all evaluators are advocates for some things, commonly for:

- Evaluation – in his words, "we care about evaluation. We want to see others care about it. We want to encourage them to do it."
- Evaluators are also commonly advocates for rationality.
- And "we care to be heard. We are troubled if our studies are not used."
- "We are [also] distressed by under-privilege" and "we are advocates of a democratic society."

Bob then speaks briefly of the ethical dilemmas of full disclosure of purpose and practice in evaluation. He notes that evaluators should disclose all their advocacies and biases but "there is no way for the evaluator to identify all predispositions, nor even to *know* them," and he turns to our Joint Standards and Guiding Principles for ethical guidance and finds some ideas there.

But he still finds a gap between thinking of evaluation as a professional activity, uncontaminated by advocacies, biases, personal predispositions and the like, and acknowledging evaluation's inherent advocacies. Moreover, even if we try to control all of these potential advocacies in our work, says Bob, "much will differ from evaluator to evaluator … In the complex determination of program quality and accomplishment, there is no single reality we can capture. The only reality is that constructed by people, and people differ."

In summary Bob says:

> We have an evaluation practice that is [inevitably] influenced by the value commitments of the evaluator [juxtaposed against] a set of operating standards that imply we can attain a widely-agreed upon picture of merit and worth…. Clearly we should develop our standards and principles so that they deal better with the uncertainty and individuality of evaluating. [Furthermore] evaluators should be encouraged to "have a life" and to "have a dream" so their interpretations are enriched by their experience. Comprehensive idiosyncratic interpretations are small steps toward saving the world.

In short, in his characteristically oblique way, Bob Stake is suggesting that advocacies in evaluation are inevitable, and thus there is indeed room for evaluators to position and locate their work in service of saving the world.

Return to "so what?"

Now to summarize these thoughts and ideas. This morning, I have suggested that urgent public issues today are rooted in private mistrust, public shirking of responsibility,

and relational fear and divisiveness. I have presented some thoughts on how these moral and ethical concerns are manifest in our world as evaluators and discussed three issues: evaluation as accountability, which I characterized as a manifestation of mistrust, evaluation's responsibilities and location in public decision making, and evaluation's engagement with diversity and difference. In each discussion, I critiqued what is currently happening (in many places, though certainly not in the state of Oregon), and I suggested ways to reclaim evaluation's potentialities and promise. I would like to return now to my agenda for reclamation and underscore its key tenets.

In my remarks this morning, I have suggested that evaluators can respond to urgent public issues by the following. We can:

1. Insist that evaluation fulfill its important *educative* mission, with particular emphasis on understanding and improving *practice*

   Some decades ago, Lee Cronbach urged the evaluation community to aspire to be educators, to help society learn about its enduring challenges and how best to address them.

   Today, Tom Schwandt quietly and persistently urges the evaluation community to re-center our work around the improvement of social practices – teaching, healing, organizing, consoling – wherein practice is conceived as involving both craft knowledge and wise moral judgment.

2. Ensure that our work values and affirms the importance of context and complexity in human action and interaction

   Bob Stake, and the many others who have followed his vision of an evaluation practice that is responsive to situational concerns and issues, offer inspiration and

countless fine exemplars of evaluation that fully engages the messiness, contingencies, and complexities of particular human lives in particular places and times. Particularity is a personal advocacy for Bob. We can continue to learn a lot from him and others in the responsive tradition.

3. Re-position our work as conducted in *partnership* with evaluation sponsors, particularly with government, rather than in service to such sponsors, thereby emphasizing our own civic responsibilities in policy and program decision making

   Ernie House, Helen Simons, and others remind us that evaluation is intertwined with government and the ideal of an informed and engaged citizenship. It is time, perhaps even an urgency, for evaluation to become a public craft in service of the public good.

4. Ask how well the program being evaluated serves the people participating in it, rather than asking how well the participants succeed in the program

   Saville Kushner encourages evaluators to ask this question in his "personalizing evaluation" approach (2000). In the context of public policy in this country, this question also rejects deficit-model thinking, in which the problem is located in the people needing the services, and instead emphasizes the responsibilities that government has to its citizens.

5. Turn some critical attention and energies to meta-evaluation, specifically:
   a. Make public not only the technical and methodological criteria that matter for good evaluation work, but also the political and values-based

principles that underlie our practice decisions and actions, or our own personal advocacies (from Stake)

b. Include as core meta-evaluative criteria some meaningful engagement with diversity in context, thereby making evaluation accountable to democratic ideals like social justice and equity (Greene, Millett, and Hopson, in press), my own personal advocacy

And finally, back to the US Department of Education's current promotion of narrow and simplistic ideas about how to "fix" our schools with technical knowledge about "what works" derived from randomized experiments -- As you recall, Russ Whitehurst, the director of the Department of Education's Institute for Education Sciences, is promoting through his policies about research and evaluation:

Not … imagination or deliberation or thoughtful concentration
But rather, in inquiry, simple randomization and experimentation
And in classrooms, memorization as preparation for those examinations.

In response, may I recommend an evaluation practice that promotes:

True education … about our enduring social problems and how best to address them
Meaningful reclamation … of evaluators' civic responsibilities in service of the public good
And generative restoration .. of the frayed and tattered fabric of our collective social and political health and well being

Thank you.

References available upon request (jcgreene@uiuc.edu)