Articles should deal with topics applicable to the broad field of program evaluation. Articles may deal with evaluation practice or theory, but if the latter, implications for practicing evaluators should be clearly identified. Examples of contributions include, but are not limited to, reviews of new developments in evaluation and descriptions of a current evaluation effort, problem, or technique. Results of empirical evaluation studies will be published only if the methods and/or findings have clear utility to other evaluation practitioners. Manuscripts should include appropriate references and normally should not exceed 10 double-spaced typewritten pages in length; longer articles will occasionally be published, but only where their importance to AJE readers is judged to be quite high.

Professional Reasons for Declining an Evaluation Contract

NICK L. SMITH

ABSTRACT

Professional evaluation is a market-based activity in which economic considerations play a dominant role. But are there ever professional reasons for turning down an evaluation contract? Considering how evaluations can promote either guild maintenance interests or societal improvement interests illuminates several dilemmas in the proper conduct of evaluation and suggests the need for greater explicit attention to what it means for the evaluation profession to contribute to the public interest.

Nick L. Smith

INTRODUCTION

> Ethical discussion aims at making us more critical of what we are doing. It brings us back to thinking about what it is to be a good evaluator, and to asking, In whose interests should we be acting, and to what purpose (Schwandt, 1997, p. 36)?

Nick L. Smith • 330 Huntington Hall, School of Education, Syracuse University, Syracuse, New York, 13244-2340; E-Mail: nlsmith@syr.edu.

American Journal of Evaluation, Vol. 19, No. 2, 1998, pp. 177–190.
ISSN: 1098-2140

Although professional evaluation is practiced under a wide variety of forms, it can be viewed as promoting primarily either its own guild interests or societal interests. Evaluation is commonly held to be both a business seeking to profit from services provided to individual clients, and a learned enterprise based on scholarly activity and dedicated to providing service to society.[1] Although individual evaluators may prefer one characterization over the other, both are embedded in current notions of the "evaluation profession." It is also important to note that, in practice, both characterizations are fundamentally economic: the practice of evaluation is essentially a market-based activity (see House, 1993). If evaluations are done principally for economic gain, are there ever reasons for not accepting a handsome evaluation contract? More specifically, are there *professional* grounds for turning down an evaluation contract?[2]

The following analysis is based on these two characterizations of the evaluation profession which suggest concerns with guild-maintenance issues and societal-improvement issues. Although few actual evaluations may fit solely under one or the other characterization, the dichotomy illuminates a variety of reasons for turning down evaluation work. I will first examine the issue of rejecting evaluation work from both perspectives, and then highlight how, in some instances, these two perspectives produce competing or contradictory rules for accepting or declining evaluation contracts. Also, for clarity of argument, the criteria for rejecting evaluation contracts are stated unilaterally, with the necessary compromises and trade-offs of actual practice discussed later in the article.

Guild-Maintenance Issues: When Evaluating Injures Evaluation

A guild is "an organization of persons with related interests, goals, etc., esp. one formed for mutual aid or protection"; today, as with the original medieval guilds, such organizations seek "...to maintain standards and to protect the interests of its members..." (Random House, 1967, p. 629). Clearly, the evaluation profession, as a collection of individuals seeking to maintain their livelihoods through the provision of client services, is in many respects a guild. If useful, quality services are not provided, clients will cease to pay for evaluators' services; if individuals cannot secure a livelihood through evaluation work, they will leave the profession.

A guild typically seeks to control the entry, training, methods, and practice of its members. Such attempts have been highly visible in evaluation over the years, from the development of the Joint Committee Standards for program and personnel evaluation begun in 1975 (Joint Committee on Standards for Educational Evaluation, 1988, 1994), to the publication of the *Guiding principles for evaluators* published in 1995 (Shadish, Newman, Scheirer, & Wye, 1995), to the American Evaluation Association's current consideration of certification procedures for evaluators. The preface of the *Guiding principles for evaluators*, for example, states that "These principles are broadly intended to cover all kinds of evaluation." (p. 20); "...should proactively guide the behaviors of professionals in everyday practice." (p. 21); and, are intended "...to foster continuing development of the profession of evaluation, and the socialization of its members." (p. 20). Clearly, guild issues are important to professional evaluators.

The evaluation profession serves, in part, to protect and promote the livelihood of its members. From the point of view of the profession as a guild, then, one could argue that evaluators should not accept evaluation contracts that are *bad for business*.

There are a number of ways in which an evaluation contract may be bad for business; each suggests a criterion for declining evaluation work.

1. *A contract should be turned down if the work is not going to be profitable to the evaluator.* In order to provide a livelihood, an evaluation must minimally be lucrative enough to reimburse expenses and to pay salary and benefits. Occasionally, an evaluator might work for a client at below cost as an investment in securing more profitable work in the future. In the long run, however, the work must pay the bills. We should note in passing, of course, that "profitable" work need not refer only to financial gain. A handsome contract might be one that promotes the evaluator's reputation or career, or simply provides an opportunity to work with valued individuals or in a desirable geographic setting.
2. *A contract should be declined if the desired work is not possible at an acceptable level of quality.* Rog (1997), for example, has discussed the use of evaluability assessment in determining when to turn down an evaluation project because good quality work is not technically feasible. She has suggested four indicators of lack of readiness for an outcome study:

   - The logic of the program is not clear: linkages between program activities and outcomes are not clear or not logical; goals are not explicit; etc.
   - There are insufficient resources to carry out the program (financial; time; expertise).
   - There is no real agreement on what the program is to accomplish.
   - The evaluation resources are insufficient to carry out the evaluation (financial; time; expertise). (Rog, 1997, pp. 3-6)

   The reasoning in using these indicators to decline evaluation work is illustrated below:

   > In a recent multisite evaluation, for example, we found that the resources provided to the individual evaluations were insufficient to conduct the outcome assessment needed. It was a recipe for failure, without sufficient resources to design comparison groups, measure the outcomes of interest, etc. The federal officials and others were cognizant of the problems but hoped that we could work out something. Though they said that they understood that we could not talk about outcomes, others involved believed it could be done and must be done. There was a sense that despite the limits, somehow it could heroically be accomplished. Given these discrepancies and the fear that we could not pull it off with the resources available, we walked away (Rog, 1997, p. 6).

   A corollary of this criterion is that *an evaluator should not accept a contract for evaluation if the work is not evaluation*, such as when "evaluation" contracts are let for work which essentially concerns management, or planning, or development problems. It is not unprofessional for evaluators with skills in these areas to accept such work, it is unprofessional to accept it under the guise of evaluation. As Stake (1997) has put it:

   > There is need for assistance, need for more action research, need for advocacy, but it is unethical, deceptive, a disservice to the profession if we provide these services when we have promised to evaluate (p. 2).

3. *Neither should an evaluation contract be accepted if the work is likely to harm the client or violate ethical principles.* Mabry (1997) has discussed such conditions, which, like the issues of technical feasibility, are related to the evaluator's ability to provide useful and ethical services to a client.

> Evaluation is the most ethically challenging of the approaches to research inquiry because it is the most likely to involve hidden agendae, vendettas, and serious professional and personal consequences to individuals. Because of this feature, evaluators need to exercise extraordinary circumspection before engaging in an evaluation study (Mabry, 1997, p. 1).

Mabry has identified a dozen issues that ideally should be deliberated in a pre-contract program review to assess possible ethical implications before accepting evaluation work:

- Access to crucial or sensitive data—(Is it ethical to accept a contract if it appears likely that evaluation could provide no more than a seriously incomplete or distorted representation of the evaluand?)
- Client openness to critique and use of evaluation findings—(Is it ethical to take a contract if a client is likely to be demoralized and incapacitated by findings?)
- Sufficiency of evaluation resources—(Is it ethical to accept a contract one might not be able to complete?)
- Support of management—(Most evaluation clients are managers; should we support management? Should we give the appearance of supporting management?)
- Conflict between program values and an evaluator's unsuppressible values—(Is it fair to evaluate a program with whose stated aims we strongly disagree?)
- Noncompliance with legal and regulatory requirements—(If I find noncompliance, will I feel it should be reported? Will reporting it be a responsibility I am willing to accept?)
- Some kinds of internal and external political conflicts—(Will the evaluation increase partisanship in a way that is counterproductive to program management?)
- Personnel divisiveness, especially where evaluation clients are in a position to intimidate other stakeholders—(Will the evaluation empower the bad guys?)
- Likelihood that an evaluation report would be used to distort impressions of program quality—(Will the evaluation be used to exaggerate claims of the program's success, to compete successfully for funds and support at the expense of better or more worthy programs?)
- Preexisting relationships among evaluators and clients or stakeholders—(Is there a conflict of interest?)
- Partisan support—(Whether an evaluation might or should contribute to the status quo or advance a particular agenda.)
- Resource reallocation—(Whether the evaluation will divert funds from stakeholders in unjustifiable amounts.) (adapted from Mabry, 1997, pp. 2-4)

Mabry admits that it is difficult to foresee and deal with all these issues prior to signing an evaluation contract, especially since many such issues may surface only as the study proceeds or nears completion. In spite of such difficulty, an evaluator who

becomes known as unethical, or incompetent in dealing with ethical issues, is not likely to remain in business long.

4. A related guild-level consideration is that *a contract for an evaluation that may harm fellow evaluators should not be accepted.* Accepting such a contract would not only violate ethical principles, but, from a guild perspective, would betray the evaluator's responsibility to protect the interests of fellow guild members. Evaluations that might contribute to intimidation of other evaluators, restriction of others' evaluative activities, or undermining of fellow evaluators' professional status and source of livelihood should not be conducted.
5. A broader but related issue is that *no evaluation contract should be accepted if it undermines the professional interests of evaluation.* Thus, evaluations should not be conducted if they violate the profession's own ethical or quality of work standards (i.e., criteria 2 and 3 above), not just because these actions are unacceptable in themselves and thus legitimate independent criteria, but because they threaten the credibility and effectiveness of the entire profession within society. Evaluations should contribute to the preservation of the profession, or, at least not contribute to its detriment. Contracts should not be accepted whenever the existence, process, or effects of the evaluation may threaten the profession by preventing or curtailing the proper conduct of future evaluations, restricting the right or opportunity to practice the profession, or in any manner discrediting the profession. Consider the following types of evaluation activities that might undermine the professional interests of evaluation.

   (a) *Evaluation contracts that lead to the de-professionalization of evaluation should not be accepted.* Critical to the profession's position in society is its ability to control who conducts the activities under its purview. To the extent that formal evaluation is conducted by non-professionals, the profession loses quality control, income, and social influence. If such approaches as empowerment evaluation genuinely lead to an increase in evaluation by non-professionals and a resultant lowering of evaluation quality, loss of business for professional evaluators, and diminishing regard for professional evaluation, then they should not be promoted.

   One might argue, of course, that society would be better off if everyone (managers, program developers, policy makers, stakeholders, etc.) did better evaluation, and that evaluation staff-development is an important professional service. This tension might be seen as a conflict between evaluation guild interests (restricting who can practice evaluation) and societal interests (broadly improving information use and decision making); such tensions are discussed further below. Alternately, this tension might reflect the need to define more explicitly what level of expertise constitutes "professional" practice. In medicine, for example, all citizens should follow better health care practices, but only licensed physicians should make complex diagnoses and prescribe regulated drugs.

   (b) *Contracts for evaluations likely to endanger the reputation of evaluation should be declined.* In situations where current evaluation methodology is inadequate, the issues of concern are highly sensitive, and any evaluation results are likely to anger all audiences, the risks may be too high to conduct a particular evaluation. It is important to "pick your battles," especially if successful evaluation is seen as providing technical services in a long term relationship dependent on client

satisfaction. As internal evaluators quickly learn, if evaluation is to make important contributions tomorrow, it must carefully select what it evaluates today. Under this criterion, the maintenance of the profession and practice of evaluation is judged more important than the conduct of a particular study.

(c) *Evaluations that antagonize major market clients should be avoided.* A large share of the evaluation profession's financial and political support comes from governmental agencies, the health sector, corporate industry, etc. Evaluations that alienate large portions of these sectors jeopardize the future of the profession itself. Brickel (1978) once noted that in giving negative feedback to clients, evaluators had to learn "how to bite the hand that feeds them while only appearing to lick it." Not even guard dogs are likely to be kept if they repeatedly bite their owner's hand.

(d) *Evaluation contracts should not be accepted if they could ultimately lead to a loss in evaluation's market share.* In describing how modern professions protect their guild interests, Rice (1997) emphasizes that

> ...a profession lays claim to a set of tasks or a domain of work over which it attempts to gain as much control as possible, a region of society that ... [Abbott, (1988)] calls a professional jurisdiction. One major task faced by a profession is legitimizing its jurisdictional hegemony, an objective often aided by legal and governmental authority (Rice, 1997, p. 1174).

Consider, for example, the guild-level disputes between the American Medical Association (AMA) and the American Psychological Association (APA) over third-party insurance payments and prescription privileges for psychological clinical care. These disputes concern federal legislation restricting who will be reimbursed by insurance companies for patient care; reimbursement significantly influences who patients can afford to see for treatment. The resolution of these disputes has clear implications for the economic survival of licensed psychologists. Similarly, evaluation competes for market share with management consultants in the evaluation of program management, auditors and accountants in the evaluation of program cost, and information technologists in the development and evaluation of information management systems, to name just a few groups. Evaluation studies that might result in the comparative advantage of other professions over evaluation should not be conducted.

Some evaluators may find these criteria for turning down a lucrative evaluation contract to be too self-serving, but they are the logical social and economic consequences of evaluation as a market-based professional guild. Such criteria serve to maximize benefit to the guild of evaluation, and are similar to the criteria used by such professions as law and medicine to insure their survival and prosperity, as when physicians categorically refuse to refer patients for care by chiropractors or acupuncturists. Licensing requirements, accreditation standards, and professional associations are employed to insure these criteria are met: only members of the profession may do the work, quality is a matter of professional purview, harm to clients is prohibited, members are protected, and the reputation and influence of the profession is enhanced. Before turning to the second view of evaluation as a profession, then, let us acknowledge that no matter how high-minded and altruistic we might wish to think of the pro-

fession of evaluation, for many evaluators it provides a livelihood they consider worthy of guild protection.

Societal Improvement Issues: When Evaluating Injures Society

The guild protection of economic gain is not the entire evaluation profession story. Although without the marketplace there would be no profession, the professional authority of evaluation consists of more than that.

> Profit is not a sufficient motive for professionals ... Although professionals exist in markets, the market cannot be the primary consideration ... Legitimation of professional authority involves three distinct claims: that the knowledge and competence of the professional have been validated by a community of peers, that this consensually validated knowledge rests on scientific or scholarly grounds, and that professional judgment and advice are oriented toward important social values ... (House, 1993, p. 19).

Under the guild perspective, evaluation contracts should not be accepted if they are *bad for business*. Under this latter perspective, contracts should not be accepted if they constitute *bad business*.

House's (1993) three claims of legitimacy each suggest a way evaluation contracts could represent bad business.

1. *Evaluation contracts should not be accepted if the work would violate professionally accepted standards of practice.* Just as physicians collectively judge what is to be considered the "treatment of choice" for a particular malady, or peer review panels of physical scientists determine what is reliably known about a given problem, evaluators are expected to follow what the evaluation profession generally accepts as best practice. Best practice is reflected in training materials, professional standards, and awards for exemplary work. At present, there is such diversity in the focus and methods of evaluation practice that a universal consensus of what constitutes best practice does not exist. Pragmatically, this criterion for accepting work may only be feasible within sub-areas of professional practice; for example, best practice for summative, internal, private sector, impact assessment in product evaluation, or best practice for formative, public, process analysis in social program evaluation.
2. Related to best practice is the criterion that *contracts should not be accepted unless the work can be based on the latest scientific or scholarly knowledge about the evaluation process and the evaluand.* As with the above principle, implementation of this criterion is problematic. Although science is seen as providing the fundamental intellectual authority for some professions (e.g., medicine), there are those in evaluation who argue that science is not the proper basis for practice in the evaluation profession. Eisner (1991), writing in support of qualitative approaches to educational evaluation, has argued that "...qualitative inquiry has no obligation—moral, epistemological, or otherwise—to be scientific in character." (Eisner, 1991, p. 180). There is fundamental disagreement about what the proper knowledge base for evaluation ought to be. Debates rage in the journals and at professional meetings about what should be the proper epistemological, ideological, and methodological basis of evaluation practice, with each position supporting a possibly different resolution as to what work should be accepted or declined. Evaluation practice should be intellectually sound, but on what basis should we judge that soundness?

3. Finally, *an evaluation contract should not be accepted if the evaluation work does not support the values of society and contribute to the social good.* Since evaluation has a professional obligation to society to promote the public welfare, no evaluations should be conducted which fail to do so. Contracts should not be accepted for evaluations whose existence, process, or effects may violate basic human rights (such as health, safety, or freedom), or undermine democratic values (such as free speech, equal opportunity, or right to privacy), or be injurious to individuals or society.

   While evaluators would all agree to "do no harm," there is undoubtedly less agreement on how to "promote the public good." For example, the professions of journalism and medicine both seek to contribute to the public welfare, but the different social roles they play lead them to deal with issues of intellectual rigor, disclosure of information, and public versus individual rights, in very different ways. One of the professional purposes of evaluation may be to promote the social welfare, but what is the proper role of evaluation in society—to foster participative, democratic decision making?—to advocate for equality of under-represented groups?—to provide unbiased, bi-partisan, information in support of public dialog? All three roles (and others) can be said to promote the public welfare; which is the proper role for evaluation as a profession? Those evaluators who would argue that the *only* proper role of evaluation is, for example, as an advocate for under-represented groups, would have to argue against accepting any contract that did not support such groups as not being in the best public interest.

> At issue here is the question of the moral end in view or the human or social good that evaluation practice is intended to serve...most answers to the question, To what purpose or in whose interest should evaluation practice be conducted?, are given in epistemological or political terms. But there are ethical claims at stake in answers to this question as well—claims about the nature of the moral life, claims about agency, identity and human interaction, claims about the way society ought to be and the ways professionals ought to serve that society, and so forth (Schwandt, 1997, pp. 30-31).

Schwandt (1997) provides a concise summary of three ethical frameworks (an analytical, value-neutral framework; an emancipatory, value-committed framework, and a value-critical framework) which he suggests are "...something like ideal types that various kinds of practices of evaluation aspire to" (p. 31). Although this criterion suggests that an evaluation contract should not be accepted if it contributes to an unacceptable social role for the profession, Schwandt's analysis strikingly illustrates the difficulty evaluators have in agreeing on how evaluation can best promote the social welfare.

Balancing Guild and Societal Interests

It quickly becomes clear that general, professional criteria for turning down a handsome evaluation contract are open to varied application depending on what one considers to be in the best interest of the evaluation guild and of society. Promoting guild or public interests can lead to competing decision rules, and the interests of society and of the profession are not always likely to coincide. For example, in order to protect the evaluation profession, one might avoid contracts that "antagonize major market clients," but such evaluations may be in

the best interests of society. Recent examples in the public media include evaluations of automobile manufacturing, the tobacco industry, and the medical establishment.

This tension between guild interests and societal interests is evident in the *Guiding principles for evaluators* (Shadish et al., 1995), whose authors opted to favor the evaluation guild more than society. The authors struggled with the problem of determining

> ... the *kind of profession* that evaluation is. For instance, is it most like medicine with its concern for helping the patient, most like auditing with its concern for accountability, or most like public administration with its concern for public service and the public good? [emphasis in the original] (Shadish et al., 1995, p. 6).

Their resolution strongly favored the guild:

> The final draft of the Guiding Principles includes a principle entitled "Responsibilities for General and Public Welfare." reflecting a task force consensus that no evaluator can *ever totally ignore* the public interest [emphasis added] (Shadish et al., 1995, p. 12).

From the perspective of evaluation as a profession, then, the *Guiding principles for evaluators* (Shadish et al., 1995) suggests that evaluators should be more concerned about refusing contracts where evaluating hurts evaluation, than where evaluating hurts society.

Hennessy (1997) has cautioned that a narrow guild orientation can ultimately harm both society and the guild itself. In commenting on the prescription privileges dispute between the AMA and the APA, he argued for attention to

> ... the degree to which prescription privileges and other 'guild issues' undermine efforts to promote and sustain the necessary collaborations among all mental health professionals that are essential to counter the powerful economic and political forces that seek to minimize insurance coverage for services, regardless of which discipline provides them...it is often in the best self-interest of psychologists to align with other mental health professionals to ensure that the services they all provide are appropriately recognized and valued by the general public (Hennessy, 1997, pp. 272, 273).

Raising a similar caution, Lagemann (1997) argues that preoccupation with guild issues has inevitably led to the increasing isolation of educational research, and has "...been a barrier to the effective linking of knowledge and action in education" (p. 15). "...the development of wide support for education research will depend on surmounting the jurisdictional challenges that have heretofore precluded sustained agreement about the methods and focus of the field" (p. 5).

Although guild and societal interests may at times be in conflict, the problem is more complex than deciding which should prevail. Facione, Scherer, and Attig (1978) discuss two types of tension found in conflicts between values: divergence and essential tension. Divergence occurs when two sets of reasoning lead to two different conclusions which, although different, may not be in conflict. In such cases, it may be possible to take the actions suggested by both conclusions. For example, guild criterion #1 argues that evaluators should not accept work which is not profitable, while social criterion #3 suggests that evaluators should do pro bono work. If it can be shown, however, that pro bono work demonstrates the utility of evaluation and hence leads to increased future paid contracts, then doing pro bono work satisfies both criteria and benefits both the guild and society.

Essential tension arises when there is no way to simultaneously satisfy both criteria—satisfying one automatically precludes the other. For example, an essential tension would seem to arise if we were to argue both that non-professional evaluators should be trained to do better evaluation work (social criterion #3), and that only professional evaluators should be allowed

to do such work so the guild does not lose income, quality control, and social influence (guild criterion #5a). Facione et al. (1978) suggest there are three ways of dealing with essential tensions: dissolve the conflict, establish a hierarchy of values, or compromise. In this case, we might dissolve the tension by charging clients for training rather than evaluation services, assisting clients with quality control methods in their own work, and clarifying what levels of expertise require work by a professional evaluator.

Essential tensions can also be resolved by establishing a hierarchy of values. For example, it may be socially and politically very important to evaluate a national educational program (social criterion #3) even though it is not possible to do so at an acceptable level of quality (guild criterion #2). In such a case, the evaluator might proceed if the social benefit is deemed to exceed to guild cost: "There may be the rare situation in which, despite its lack of evaluability or perhaps because of it, an evaluator feels it is justified to rigorously evaluate a program to demonstrate its problems" (Rog, 1997, p. 7).

Finally, essential tensions may be resolved through compromise in which both criteria are observed, but neither fully. Meta-evaluations of large social programs might be warranted because evaluation work should not violate professional standards (social criterion #1), but the results could harm fellow evaluators (guild criterion #4) and damage the reputation of professional evaluation (guild criterion #5b). Having professional evaluators conduct these studies both provides a benefit to society (social criterion #3) and enables the profession to insure its work is of acceptable quality (guild criterion #2), but is a compromise since professional evaluators are likely to be more sensitive to issues of harming fellow evaluators and the profession than would be non-professional reviewers. (See Smith, 1982, for a more detailed discussion of how value conflicts may be resolved in evaluation.)

A primary impediment to recognizing and resolving value conflicts such as these is that evaluators are often vague or silent on how, specifically, evaluation is to improve society. Because the evaluation profession as a whole lacks a clear understanding and widespread agreement of what it means to act in the public interest, it is easier to promote guild interests—sometimes accompanied by apologetic rhetoric about promoting the social good. Rossi (1995) has provided an example of how this lack of clarity about the public interest weakens the *Guiding principles for evaluators*. He argued that the principle, "Responsibilities for General and Public Welfare," is vague and ineffective:

> For example, this principle would be more meaningful if it flatly stated that evaluators would be serving the public good if they refused to undertake classified evaluation research, or that serving the public good means making evaluation findings available to all stakeholders, professional journals, and the mass media. Analogous specific rules concerning other ways of serving the public interest might also be stated (Rossi, 1995, p. 57).

At present there is insufficient discussion in evaluation of how the profession might better serve the public interest. For example, evaluators might be expected to provide pro bono services so that those least able to afford evaluation could still obtain qualified help. Or, the profession could encourage collegial training of other professionals (for example, teachers, managers, service providers, administrators, policy makers), so that they might better perform the evaluative aspects of their own responsibilities. Of course, as discussed above, this could threaten guild control of evaluation practice and quality.

The American Evaluation Association (AEA) currently makes no provision for supporting the evaluation of important social problems; for example, through an AEA foundation. Indeed, in a recent call for proposals "...that show promise of furthering the [member] service

or fund-raising aims of the organization" ("Proposal Guidelines for Financial Ventures," sponsored by the American Evaluation Association, 1997, p. 5), a number of issues are raised of potential costs and benefits to AEA members, but no mention of possible contributions or harm to society. Neither is there any indication that part of the revenue to be accumulated might be used for anything other than guild interests.

Evaluators *are* concerned about ethical practice and contributing to the social good. Concern with such issues has been raised in the literature, and, in part, underlies the current interest in incorporating client participation in evaluation studies (for example, Cousins & Earl, 1995), empowering clients through the evaluation process (for example, Fetterman, 1995), and evaluators becoming client advocates (for example, Greene, 1997). These issues are of concern in daily evaluation practice, as evidenced by an extended discussion about reasons not to conduct evaluations that reverberated on the AEA Internet listserv, EVALTALK, over several weeks in February of 1997 (EVALTALK, 1997). The sensitivities exist, but evaluation policy and associated practical procedures that might serve to promote the public interest have not been worked out in sufficient detail.

CONCLUDING THOUGHTS

Considering how the evaluation profession may promote guild or societal interests gives rise to other professional dilemmas, as the following two examples illustrate.

If the principles of promoting the evaluation profession and contributing to the social good determine when an evaluation contract should be turned down, do they also suggest when an evaluation contract should be accepted? Does an evaluator have a professional obligation *not* to turn down a contract for an evaluation that would contribute to the profession or society? The question posed initially in this article was, "When is it right to say 'no'?"; but we should also consider, "When is it wrong to say 'no'?" Under what conditions should a priest or physician or evaluator turn someone away? As Mabry (1997) points out:

> A useful complication is whether refusing a contract might, in some cases, be less ethical than accepting one. It is not difficult to imagine situations in which there is relatively little possibility that evaluation might help but even less assurance that lack of evaluation—doing nothing—would help (Mabry, 1997, p. 3).

Further, does the principle of doing no harm to the profession rule out the possibility of an honest summative evaluation of the evaluation enterprise itself? Although a formative evaluation of the profession could be seen to be in the best interest of both the profession and society, a negative summative evaluation might be professional suicide. Evaluators may tend to equate the evaluation profession, and, more narrowly, professional evaluation associations such as the AEA, with the societal function of evaluating, but they are clearly not the same. Consider, for example, the claims that the medical profession, specifically the American Medical Association, has harmed the health of society by creating a monopoly which focuses on only "disease" models of health care and has excluded numerous alternative treatment traditions (homeopathic, acupuncture, and the like). Are summative evaluations of the medical profession possibly in society's best interest? Similarly, are such evaluations of the evaluation profession also likely to be in society's best interest? Some evaluators have so argued, such as Stake (1986) in his book, *Quieting Reform*. If so, professional evaluators would be precluded from conducting such evaluations, unless evaluation as a profession accepts that societal

improvement interests should take precedence over guild maintenance interests. House (1997), resolving the apparent essential tension here, has suggested that violating guild principles, such as daring to antagonize clients or to criticize fellow evaluators through meta-evaluation, may strengthen the integrity and independence of the profession, thereby furthering both guild and societal interests.

Greater understanding of what it means for the evaluation profession to contribute to the social good is needed if we are to resolve dilemmas such as these. The economic self interest of members of the evaluation profession is immediate, clearly felt, and easily used as a basis for deciding whether to turn down a handsome evaluation contract. Issues of contribution to the public welfare are more complex and subject to interpretation based on one's values about the proper roles and methods of evaluation practice. The diversity of philosophical and political positions within the evaluation profession promises no easy resolution to the search for widely accepted principles for determining whether to accept evaluation work, or the trade-offs of guild versus societal interests. Under the guise of pluralistic acceptance, evaluators may continue to focus primarily on their own financial self interest, being cautious, of course, not to "totally ignore" the public interest. Alternately, evaluators might risk the disagreements of sorting out their differences in an attempt to clarify the evaluation profession's responsibility to society. As Rossi (1995) says, "In my view, evaluation is not getting along with everybody but getting it right. It is not negotiation for approval but the search for the best approximation to the truth." (Rossi, 1995, p. 58).

Or, should we say, "best approximation to the truths"? The evaluation profession may already be too diversified to have a unified view of the public good, as suggested by the struggle of the authors of the *Guiding principles for evaluators* in determining, "...the *kind of profession* that evaluation is." [emphasis in the original] (Shadish et al., 1995, p. 6). Perhaps the proper contribution to the public interest is different for evaluators conducting product evaluation in private sector corporations, than it is for evaluators conducting advocacy evaluations of public service programs. Perhaps, but we will not know until the issues of guild interests versus societal interests are treated more explicitly by the evaluation profession.

ACKNOWLEDGMENTS

An earlier version of this paper was presented at the American Educational Research Association annual meeting, March, 1997, in Chicago, Illinois. My thanks to Ernest R. House and Karen E. Kirkhart who offered helpful suggestions, and to three anonymous reviewers and this journal's editor, Blaine R. Worthen, who provided constructive reviews and advice.

NOTES

1. The English language itself suggests a guild versus societal distinction. For example, the unabridged edition of *The Random House Dictionary of the English Language* (Random House, 1967) provides two distinct connotations of the word "profession." The first is that a profession is "any vocation or business," and a professional is someone "following an occupation as a means of livelihood or for gain" (p. 1148). The second meaning is profession as "a vocation requiring knowledge of some department of learning or science," as in a "learned profession." Under this meaning, a professional is someone "engaged in one of the learned professions" (p. 1148), which historically were "one of the three

vocations of theology, law, and medicine, commonly held to require highly advanced learning, high principles, etc." (p. 816).

2. It has been pointed out that most evaluations done in the United States are not done by persons who consider themselves "professional" evaluators, but by substantive scholars, management consultants, program directors, and many others with differing types of expertise. While this observation is certainly accurate, the present discussion focuses on the responsibilities and obligations of those who seek to engage in evaluation as a profession. This enterprise has matured considerably in the twenty years since Morell and Flaherty (1978) noted that "...evaluation shows unmistakable signs of going down the path of professionalization." (p. 11)

One reviewer noted that, "In sharp contrast to medicine and the law, no one can be criminally charged with 'practicing evaluation without a license.'" This is true at present, but one implication of the full professionalization of evaluation is that one day a license may be needed to practice evaluation. Indeed, licensing, credentialing, and ensuring that only qualified individuals practice evaluation, are perennial topics among professional evaluators. My purpose in writing this article is to stress the need to be as concerned with society's interests as with evaluation guild interests in developing a profession of evaluation.

REFERENCES

Abbott, A. (1988). *The system of professions: An essay on the division of expert labor.* Chicago: University of Chicago Press.

Brickell, H. M. (1978). The influence of external political factors on the role and methodology of evaluation. In T. D. Cook, M. L. Del Rosario, K. M. Hennigan, M. M. Mark, & W. M. K. Trochim (Eds.), *Evaluation studies review annual* (Vol. 3). Beverly Hills, CA: Sage.

Cousins, J. B. & Earl, L. M. (Eds.). (1995). *Participatory evaluation in education: Studies in evaluation use and organizational learning.* London, Falmer.

Eisner, E. W. (1991). Taking a second look: Educational connoisseurship revisited. In M. W. McLaughlin & D. C. Phillips (Eds.), *Evaluation and education: At quarter century.* (pp. 169-187). Chicago, The University of Chicago Press.

EVALTALK (February 14-26, 1997). American Evaluation Association Discussion List. [On-Line]. Available: LISTSERV@UA1VM.UA.EDU

Facione, P. A., Scherer, D., & Attig, T. (1978). *Values and society: An introduction to ethics and social philosophy.* Englewood Cliffs, NJ: Prentice-Hall.

Fetterman, D. M. (1995). *Empowerment evaluation.* Newbury Park, CA: Sage.

Greene, J. C. (1997). Evaluation as advocacy. *Evaluation Practice, 18*(1), 25-35.

Hennessy, K. D. (1997). Neglecting our common—and the public—interest. *American Psychologist, 52*(3), 272-273.

House, E. R. (1993). *Professional evaluation: Social impact and political consequences.* Newbury Park, CA: Sage.

House, E. R. (1997, March). [Discussant Comments]. In R. E. Stake (Chair), *Grounds for turning down a handsome evaluation contract.* Symposium conducted at the meeting of the American Educational Research Association, Chicago, IL. (Cassette Recording No. RA7-40.34). Chicago, IL: Teach'em.

Joint Committee on Standards for Educational Evaluation (1988). *The personnel evaluation standards.* Newbury Park, CA: Sage.

Joint Committee on Standards for Educational Evaluation (1994). *The program evaluation standards* (2nd ed.). Newbury Park, CA: Sage.

Lagemann, E. C. (1997). Contested terrain: A history of education research in the United States, 1890-1990. *Educational Researcher, 26,*(9), 5-17.

Mabry, L. (1997, March). Ethical landmines in program evaluation. In R. E. Stake (Chair), *Grounds for turning down a handsome evaluation contract.* Symposium conducted at the meeting of the American Educational Research Association, Chicago, IL.

Morell, J. A. & Flaherty, E. W. (1978). The development of evaluation as a profession: Current status and some predictions. *Evaluation and Program Planning, 1,* 11-17.

Proposal guidelines for financial ventures sponsored by the American Evaluation Association. (1997, Winter). *Evaluation Practice News,* 5-6.

Random House (1967). *The Random House dictionary of the English language.* New York, Random House.

Rice, C. E. (1997). The scientist—practitioner split and the future of psychology. *American Psychologist, 52*(11), 1173-1181.

Rog, D. J. (1997, March). When NOT to do an outcome evaluation: Assessing the evaluability of a program. In R. E. Stake (Chair), *Grounds for turning down a handsome evaluation contract.* Symposium conducted at the meeting of the American Educational Research Association, Chicago, IL.

Rossi, P. H. (1995). Doing good and getting it right. In W. R. Shadish et al. (Eds.), *Guiding principles for evaluators* (pp. 55-59). San Francisco, Jossey-Bass.

Schwandt, T. A. (1997). The landscape of values in evaluation: Charted terrain and unexplored territory. In D. J. Rog & D. Fournier (Eds.), *Progress and future directions in evaluation: Perspectives on theory, practice, and methods* (pp. 25-39). San Francisco, Jossey- Bass.

Shadish, W. R., Newman, D. L., Scheirer, M. A., & Wye, C. (Eds.). (1995). *Guiding principles for evaluators.* San Francisco, Jossey-Bass.

Smith, N. L. (1982). Evaluation design as preserving valued qualities in evaluation studies. *Studies in Educational Evaluation, 7,* 229-237.

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

Stake, R. E. (1997, March). Response to 6 papers on when to turn down a handsome evaluation prospect. In R. E. Stake (Chair), *Grounds for turning down a handsome evaluation contract.* Symposium conducted at the meeting of the American Educational Research Association, Chicago, IL.