Client-Centered Evaluation: Ethics for 21st Century Practitioners

Martin Bloom, PhD
Professor Emeritus
University of Connecticut

Abstract

Ultimately, evaluation belongs in the hands of the person most affected, the client, who should testify whether a conventionally evaluated intervention has attained his or her desired goal, as well as demonstrate that he or she can perform this intervention independently. Conventional evaluation is simultaneously performed on behalf of society to document the effective and efficient performance of public-financed service.

Key words: Client-centered evaluation, ethics, intervention phase, maintenance phase, single-system designs

1.0 Ethics in Evaluation and Research

Ethics, briefly defined with reference to the helping professions, involves a set of principles of right conduct that is supposed to govern practitioners' behaviors in clinical and social change situations (Reamer, 2006; Reamer & Shardlow, 2009; NASW Code of Ethics, 1999). There is no "ethics army" so that the enforcement of its principles is usually left to the consciences of mature individuals in the professions. As a constant reminder, these professions usually have concretized their ethical principles in written codes that are derived, more or less directly, from Hippocrates (circa 460 -377 B.C.E.). Veach (1981:22) quotes Hippocrates' The Epidemics as follows: "As to diseases, make a habit of two things – to help, or at least to do no harm." These 18 words or their equivalent have survived for nearly twenty-five hundred years to guide medical and other helping practices (Hartsell, 2006;

Loewenberg, Dolgoff, & Harrington, 2000; Meacham, 2007; Reamer, 2006). It is the task of this paper to connect 21st century evaluation with 5th century B.C.E. ethics.

I make the distinction between research and evaluation because the ethical requirements and implications of each differ in some important ways. This is true in spite of other applications of the terms, such as evaluative research, program evaluation, and the like, which constantly blend the two concepts as if they were the same thing. *Evaluation*, in the narrow sense in which I am using the term, involves producing practical and approximate knowledge for immediate use in client situations to provide guidance for achieving client goals (Bloom, Fischer, & Orme, 2009). *Research* involves producing abstract but usually mathematically specific knowledge for long-term understanding of conditions surrounding a group of persons with common problems or concerns. Research rarely provides any feedback to the subjects who composed the study group. Research studies supply the evidence for evidence-based general practice; my focus here is on evaluation-informed specific practice, that is, once a general intervention is selected based on the best available research, then practitioners have to individualize it or customize it in their intervention plans and then monitor its effectiveness with a specific client.

Evaluation involves the comparison of one client-system's current situation against some other reference point such as comparing back to that system's own initial baseline starting point (for example, what I weighed two months ago compared to what I weigh now); or comparing forward to some external standard or benchmark (like my trying to lose weight so as to reach the weight limits required

for joining the police force). Thus, evaluation uses the client as his or her own "control" group, a perfect equivalence with which to make these "before/after" comparisons, an equivalence that classical research designs and randomization seek to emulate – but never fully succeed in achieving. Evaluation, unlike research, is intentionally constructed to shape practice in field settings at the moment information is needed to make practical decisions (such as proceed as planned, change out of a deteriorating situation, or terminate as having achieved a stable goal). Classical research, often called the "gold standard" for empirical information, usually requires much more time for funding, institutional reviews, completion, and analysis, let alone the rare translation into concrete practice suggestions for clients/subjects in general, so that this gold is seldom transmuted into everyday currency to shape practice in real time.

Evaluation, in contrast to research, is very much influenced by its participants, which includes on-the-dime changes of direction when local evidence supports such change; additions of new targets as needed; and consideration of client's specific positives (to be maintained) as well as that client's problems (to be resolved) – any of which would not fit well into classical research designs, which more likely deal with problems in general. Evaluation is sensitive to the nuances of the particular client situation, in distinction to classical research, which has to consider the same outcome measures for all participants. The approach evaluators take has enormous advantages in the immediate client situation. For example, on-going measurement and monitoring makes it possible to detect signs of deterioration early and thus be able to make suitable adaptations. Of course, the same monitoring can inform practitioners about early signs of probable success, which allows planning for additional interventions if needed. The point is that evaluation is time-focused on the here-and-now, while research holds a more futuristic view of outcomes, namely, those that are the basis for evidence-based general practice.

These and like considerations have led my colleagues and me to consider a form of evaluation we call single-system design (Bloom, Fischer, & Orme, 2009) because of its applicability to a wide array of immediate practice situations involving persons, groups, or collectives. The chief characteristics of this approach may be briefly described as follows: single-system designs involve a set of empirical procedures to observe changes in an identified target (the dependent variable, a problem or objective of the client) that is measured repeatedly over time using the same procedures. A baseline reference pattern is used to compare the same

targeted pattern during intervention with reference to client goals or other benchmarks for desired outcomes. Some of the more elaborate designs used in single-system evaluation permit inferences of causality, such as multiple baseline designs or experimental repetition designs (A-B-A-B). The basic (or minimal scientific) A-B design compares baseline against intervention, which logically permits objective assessment of change, without specifying causal factors.

I call attention to the fact that at times in most practitioners' lives, they need to know the causal situation, especially in the maintenance phase when they teach the client to use a successful intervention on his/her own, as I will discuss shortly. These reasons are very different from why a researcher wants to have causal information, to build the evidence basis of social science. In conventional single-system evaluations, data are analyzed visually and/or by statistical methods, and practical decisions are made using the results, in conjunction with other considerations (such as agency policy, values, and professional standards). It is this immediate practicality of evaluation results that is critical, compared to the longer term research results that eventually reach a published form, which becomes the basis of evidence-based general practice.

There are threats to the validity of the findings in A-B designs that more elaborate evaluation designs can minimize, but all evaluations are tentative, approximate measures of a changing state of affairs that are to be used within the entire context of information, something like watching a moving picture of the client's targeted life. Thus, the practitioner cannot say with absolute certainty that evaluation results are the final word on outcomes; rather, these immediate outcomes themselves have to be assessed by the client as exhibiting a positive and meaningful change in that client's life. Such personalized results are in contrast to the usual statistical results that summarize or express changes in research for the population studied.

2.0 Client-Centered Evaluation in Practice

I use the term client-centered evaluation to represent a new configuration of choices and decisions practitioners and clients have to make together. Let me enumerate them:

First, a client-centered evaluation involves the client as much as possible in identifying the *goals* of intervention. The phrase, client-desired outcomes, names the process by which goals of the case are first enunciated. The practitioner usually has to re-state these desired outcomes to ensure that they are

publicly clear and agreed on, so that the rest of the intervention can focus on attaining them.

Second, goals are broken down into feasible intermediary steps, called *objectives.* Targets are the operational proxies for objectives, in which the specific client concern is selected for intervention by means of a specific set of actions, as clarified by the practitioner. The measured degree of a given target is plotted on a graph, and over time, these data represent a moving image of the target. Often clients are involved in collecting data as well as in discussions about changes in targets.

Third, as usual, *baselines and intervention phases* are conducted as needed to provide a logical basis for observed changes in the target as in the basic AB design. Advanced designs permit the logical inference of causality, such as ABAB and multiple baseline designs.

Fourth, regardless of any improvement in client functioning that occurs in the B phase of an AB design or the second B of the ABAB design, I would assert that it is unethical to stop here. Rather, I propose that the basic evaluation design be described as an ABM design, where M refers to the *maintenance phase* in which the client is instructed on how to take over the entire machinery of his or her improvement (to the extent possible) and demonstrate that he or she can continue obtaining positive results without practitioner assistance for a sustained time period. This time period will vary depending on the seriousness of the concern, but stability in desired outcomes is a basic condition for client-centered practice. Then, and only then, will the client's goals of attaining desired and sustained outcomes be a reality, so far as the practitioner can make this happen. Moreover, this dual analysis – a logical outcome in a B phase and the psychological outcome of the M phase – emphasize the dual ethical concerns for societal welfare and personal well-being reflected in the current NASW Mission Statement. The same would be true of any ABABM design, and in a multiple baseline design in which the client has to demonstrate independent control over any client-confirmed positive outcome.

3.0 Hippocrates in the 21st Century

The great possibility of twenty-first century client-centered evaluation's contribution to ethics is to fulfill its goals. Say you want to *help* if you can? Good, practitioners now possess the tools for each client in almost every clinical or social setting to evaluate whether change has occurred, and possibly whether they have contributed to that change. But is it no longer acceptable to say merely that "we helped." Ethical accountability demands that

practitioners specify what was achieved with the help that was given. More than that, they must specify to what degree they achieved the goals that were sought by clients and society. And further, they must indicate at what costs. These issues will lead us into an ancient ethical discussion, in a few moments.

Practitioners also need to deal with *do no harm* by returning to some ancient ethical issues that play out dramatically in our own day. Ultimately, I would argue that we can never say that "no harm has been done" because helping an individual client is like every surgical procedure that "harms" or cuts into the integral body surface in order to bring about a greater healing. Ethics is a comparative enterprise. Practitioners have to weigh what it costs to achieve what desired effects, and with what costs to the parties involved. There is no free lunch.

Likewise, evaluators, unlike researchers, must be aware at all times that it is never easy to be a client seeking help, effectively admitting failure at self-resolution or local solutions by family and friends. That cuts deeply into the integral personality, but it is a necessary "harm" to get the process of healing started. By engaging both the strengths of clients, as well as their presenting concerns and problems, we actively minimize harms. Research does not concern itself with such nuances, especially with the control group during the research period. This is demonstrated in the mathematics of research where all subjects are effectively considered equivalent, which practitioners know is never the case.

4.0 Evaluation in Ethical Theories, Ancient and Modern

Ethicists tell us that there are two broad classes of ethical theory that can be described as 1) the absolutist or deontological approach, and 2) the consequentialist or teleological approach (Meacham, 2007). There are many positions in between, taking elements of each model and combining them, often in Rube Goldberg fashion, to generate a new theory of ethics. The absolutist approach blends with the all-or-nothing position, which asserts that either we helped achieved client goals, or we didn't. Our duty is to reach this level of perfection and to keep matters there. Then the client and we live happily ever after.

I find the absolutist approach to be fairy-tale-like, and unsuited to the real world. The consequentialist approach tells us that by the fruits will you know whether the objectives have been achieved, especially if you can ask the client at risk. Most importantly, this position recognizes that everything under the social/cultural sun comes in degrees, more of this, less of that. This is true of

evaluation as well, which leads to the obvious question: how much of a desired outcome has to be achieved before we accept the intervention as producing "help (if you can), or at least do no harm"?

The issue, namely, what do we set as standards of successful outcome, is complex. Let me raise a few questions: First, recognize that practitioners can obtain degrees of improvement (or deterioration) compared to a baseline pattern. But how different must these two patterns (baseline and intervention phase data) be before victory is declared and we all go home? Let's assume we have a clear operationally defined target, such as minutes Junior practices each day on the piano, or pounds pudgy Sister loses after a period of vigorous exercise and controlled eating, or the degree to which aged Mr. Smith feels less angry for having been effectively forced to move into a nursing home after he lost practically every social support he had. What is measurably good/desirable/moral in these cases?

Perhaps the music teacher, based on years of experience, sets the goal of 30 minutes of practice a day (no procrastination, real playing). Maybe the school nurse has suggested to Sister's parents that she exercise X minutes and eat Y calories a day, which has been shown to take off pounds. Maybe there is no goal for Mr. Smith, other than survival at some acceptable level of contentment.

Evaluation of the first of these examples (piano playing) is simple, and both child and parent can verify the number of real practice time to some acceptable level of reliability. The second example is more complex because both exercise and eating have to be monitored, which can be difficult when Sister is on her own (and hungry) at school. The third example is very difficult, since there are no norms for contentment under these conditions, and Mr. Smith himself might doubt that nursing home living is really living at all.

This leads to a second consideration on ethical judgments when no clear goals or norms are available. In these instances, we are likely to fall back on science as being neutral and disinterested (compared to practitioners who are not neutral and are definitely interested in positive outcomes). Are the patterns at intervention significantly different than those at baseline?
But notice that this question, frequently asked in the literature on single-system design, will not tell us if Mr. Smith is contented. It only tells us that his contentment level (however we choose to measure it) is significantly different between the two time periods. And even this is relative: if Mr. Smith was at the 5% level of contentment at first, and then zoomed up to 10% after this intervention, even 10% may not be of any practical significance in Mr. Smith's life.

Say we have norms on contentment from a variety of studies of older people who had been moved into nursing homes from various earlier living sites. Let's say 40% of these elders were contented with their new surroundings. Or, let's say 4% were contented. The point is that for any pattern of change comparing Mr. Smith's scores with these general findings, we can make statistical interpretations, such as that Mr. Smith's 10% level of contentment is far below the 40% norm, or slightly above the 4% norm, which is an important first step in evaluation.

However, this is only the first step in an ethical analysis. I believe that ultimately, the client has to interpret any change within that client's own perspective. Mr. Smith could say that his level of contentment (at 10%) is not satisfactory, no matter whether it is compared to the 40% or the 4% norm. Unless we build this client determination directly into our evaluation process, we are avoiding the ultimate ethical decision regarding this client's situation.. This is not to say that I have any less enthusiasm for repeated systematic observations or the rest of the machinery of single-system designs, but rather I have come to appreciate more fully the operationalized fulfillment of client ethical concerns.

5.0. Research Evidence Versus Evaluation Evidence

Let's take a second tack in looking at research and evaluation in the ethical context. The rising tide of science in support of practice has been labeled evidence-based practice, and I do not intend to surf against the tide. However, with single-system designs, the practitioner has a marvelous surf board to ride with the wave.

Let's begin by defining terms, since evidence-based practice comes with many associated concepts (empirically-based practice, scientific practice, among them). "Evidence-based practice represents the practitioner's commitment to use all means possible to locate the best (most effective) evidence for any given problem at all points of planning and contact with clients" (Bloom, Fischer, & Orme, 2009, p.13). However, the underlying meaning of this definition includes one more term: evidence-based *general* practice. This means that the analysis and combination of many studies leads to general conclusions for practice, such as the probability that such-and-such a procedure, if properly executed, will likely lead to this pattern of results. This is vital information, and propels the tide in favor of universal evidence-based practice. And this is why we use the enlarged phrase, evidence-based general practice (Bloom, Fischer, & Orme, 2009).

However, what practitioners need, after they have some evidence-based general information on what works under what conditions for what kinds of people and problems, is much more specific: what will work with my client sitting in front of me? This is where evaluation enters the professional scene. Having selected one (or more) methods from the evidence provided by analysis of studies on this clinical or social change topic, the practitioner has to tailor the general methods, questions, time table, to fit the conditions of the immediate client. Recall that NASW Code of Ethics requires social workers to base practice on recognized knowledge, including empirically based knowledge (4.01.c). How will they know if this general evidence is working, especially having made some modifications to suit the conditions for their specific clients? By evaluating their own practice.

Using the evidence-based general practice ideas, the practitioner now fine tunes his or her evaluation to become informed about the specific effects with a given client. Hence, we call this evaluation-informed specific practice (Bloom, Fischer, & Orme, 2009). I now want to connect this to ethical considerations: By combining both evidence-based general practice and evaluation-informed specific practice, and involving clients (as far as possible) in choice of goals, demonstrating control over a successful intervention, and determining when an accomplished outcome in fact attains those goals, we have come one step closer to ethical practice for the 21st century. To paraphrase Kant, evidence-based general practice without specific evaluation is clumsy; evaluation-informed specific practice without general evidence is blind. Both are necessary; both must be taught to the new generations of ethical helping professionals.

6.0 Specific Ethical Considerations in Using Single-System Evaluation

I now turn to some of the specifics in using ethical methods of evaluation within the single-system design paradigm. Let's begin with the fundamental assumption, that any and every choice and action a practitioner makes in the field of practice with clients is an ethical decision (Reamer, 2006). Such choices and actions presumably introduce changes to the client's life -- that is to say, in his or her interactions with others, which is the moral ground of all human behavior. Here are the ethical particulars in single-system evaluation:

1. We must seek to provide *demonstrable help* for the individual client in his or her social setting. For this, a single-system design provides the closest form of hard evidence that we can attain easily, quickly, at low costs, without the use of elaborate methods. Evaluation is an approximation of the best research; sometimes the approximation is very close (with advanced single-system designs); other times, evaluation provides only hints (with the basic A-B design).

2. We must demonstrate that *no harm* was done in the process of intervening and collecting ongoing data, for both the individual client and for his or her social context. Thus, even if help could be provided for the client while at some harm to others in the social context, the principle of doing no significant harm within the larger social context takes precedence over doing good for the client (Reamer, 2006; but also see Hartsell, 2006). The single-system form of evaluation can make these distinctions, which would likely be hidden by group data in research -- and likewise in using only evidence-based general research for our unique client. Single-system design has a significant contribution to the dual focus of combining evidence-based general practice with evaluation-informed specific practice.

3. Since ethical considerations involve the client directly, so in evaluation, we often involve the client as directly as possible in selecting targets, choosing methods of data collecting, participating in those data collections, and interpreting the outcomes as part of real world changes the client is seeking. (We may not involve some clients directly, if they have limitations of age or functioning, in which case their adult caretakers are directly involved.)

4. Evaluation has the luxury of stopping itself, should the intervention or data collection prove to be painful or harmful to the client or client-system, physically, psychologically, or socially. The difficult part is performing this action without prejudicing the services being offered. The clinical and lay literature is littered with terms such as "noncompliant patient" (Bloom, 2008) and the like -- and I have no doubt that clients can be difficult -- but in fact clients are telling us something by "acting out" or "violating agreements" that we don't like to hear, but must hear if we are to act appropriately. Evaluation offers many ways to achieve its results that may not involve actions clients find offensive, like having to report on their own behavior, by means of indirect methods including unobtrusive observer ratings made in private. In research, we are generally stuck with the methods and designs we started with.

5. Evaluation shares with research the obligation of confidentiality with regard to data, records, and the like. However, it may be easier for the one-on-one practitioner/client situation to explain the limits of confidentiality (see Tarasoff v. Regents of California [17 Cal. 3rd 425, 1976; Kopels & Kagle, 1993). The difference between confidentiality

in evaluation as contrasted with research is that a bond of trust is likely to exist between the practitioner and client in the former that does not exist in the latter, that all information obtained in confidence must benefit all parties and harm none. It should be obtained through informed consent (Bloom, Fischer, & Orme, 2009: 572).

6. Evaluation takes into consideration directly the unique factors of the individual client, including ethnicity, income level, educational level, sexual orientation, and gender. These become factors to be controlled in research, and yet in evaluation, they are the specific guiding factors on which decisions have to be made. It was for this reason that we constructed a client bill of rights (Bloom, Fischer, & Orme, 2009: 571) that explicitly states what are client rights regardless of individual differences (Wilson, 1983). With these kinds of sensitivities in evaluation, we can increasingly know where this

specific client is (regarding targets); we can reduce the drop out rate (which often occurs when the client is dissatisfied with the intervention); we can have fewer treatment sessions with no worsening of outcomes for the clients making progress -- this will be shown in increased cost-effectiveness (Harmon, et al., 2007; Lambert, 2007).

7.0 Summary: Client-Centered Evaluation in Practice
In this summary, I offer this paradigm of the nature of client-centered evaluation as it is combined with practice. On the left side of this summary are short-hand expressions for what is clarified on the right side. This represents the blending of practice and evaluation processes, reflecting the central place clients should have in both. It is constructed as a kind of dialogue with the client, although I use technical terms to communicate with this professional audience:

Practice and Evaluation: General	Practice and Evaluation: Particulars
[Thorough assessment]	1." After careful consideration of the scope and nature of your (the client's) presenting concerns,
[Evidence-based general practice]	2. " and a thorough review of the relevant empirical literature that discusses how people in your situation with your kinds of concerns fared when a given intervention was used to address these concerns,
[Intervention and Evaluation Design chosen, with client-chosen goals and informed consent]	3. "I (the practitioner) have what I think are the best plans to address your concerns, as well as how I'll check with you on how well we are doing. I want to give you this information so that you can make an informed decision on whether to proceed or not. This will involve your discussing your desired outcomes regarding these concerns.
[Probable positive outcomes and possible risks]	4. "I do not know whether this intervention will work with you as it has for others with similar concerns, so I will be evaluating the positives and negatives, if any, of the outcomes of this intervention, and sharing the information with you.
[Individualized plan: Evaluation-informed specific practice]	5." I have adapted this plan of action to fit your unique circumstances, emphasizing your strengths and available social supports, and taking into consideration your problems and limitations.
[Start now, and do on-going monitoring]	6." So I propose we start now, with each of us taking parts in this process. I will carefully monitor your progress toward your desired goals. I will use this information that we collect as we go along in our contacts to fine tune the intervention. .
[Criteria for Success]	7. "We will know if these have been good choices of interventions when the concerns that brought you here are resolved, and you feel that you have attained your desired goals as much as possible. The point of this service is to give you the tools to succeed on your own."

References

Bloom, L. Z. (2008). (Im)Patient. *The seven deadly virtues and other lively essays* (pp.184-195). Columbia, SC: University of South Carolina Press.

Bloom, M., Fischer, J., & Orme, J. (2009) *Evaluating practice: Guidelines for the accountable professional* (6th ed.). Boston: Allyn & Bacon.

Hartsell, B. D. (2006). A model for ethical decision-making: The context of ethics. *Journal of Social Work Values and Ethics*, 3:1.

Harmon, S. C., Lambert, M. J., Smart, D. M., Hawkins, E., Nielsen, S. L., Slade, K., & Lutz, W. (2007). Enhancing outcomes for potential treatment failure: Therapist-client feedback and clinical support tools. *Psychotherapy Research*, 17, 379-392.

Kopels , S., & Kagle, J. D. (1993). Do social workers have a duty to warn? *Social Services Review, 67*, 101-126.

Lambert, M. J. (2007). Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome

.

in routine care? *Psychological Research, 17,* 1-14.

Loewenberg, F. M., Dolgoff, R., & Harrington, D. (2000). *Ethical decisions for social work practice* (6th edition). Itasca, IL: F.E. Peacock.

Meacham, M. G. (2007). Ethics and decision making for social workers. *Journal of Social Work Values and Ethics, 4*:3.

NASW . (1999). *Code of ethics of the National Association of Social Workers*: Washington, DC: Author.

Reamer, F. G. (2006). *Social work values and ethics*. NY: Columbia University Press.

Reamer, F. G. & Shardlow, S. M. (2009). Ethical codes of practice in the US and UK: One profession, two standards. *Journal of Social Work Ethics and Values, 6*: 2.

Tarasoff v. Regents of University of California [17 Cal.3rd. 425, 1976]

Veach, R. M. (1981)*. Theory of medical ethics*. New York: Basic Books

Wilson, S. J. (1983). Confidentiality. In A. Rosenblatt & D. Waldfogel (Eds.) *Handbook of clinical social work.* San Francisco: Jossey-Bass.