EVALUATION IS TO HUMAN RESOURCE DEVELOPMENT
WHAT LOSING WEIGHT IS TO THE AMERICAN MIDDLE CLASS.
BUT LIKE LOSING WEIGHT, THE RESULTS
OF EVALUATION ARE RARELY WHAT ONE HAD ORIGINALLY HOPED FOR.

PRAGMATIC EVALUATION

BY ANTHONY O. PUTMAN

Evaluation is to human resource development what losing weight is to the American middle class. Nobody denies its importance (who would dare?), almost everybody has plans to do it (or vague guilt about not having already done it), and the mere act of trying brings automatic approval. Heads nod thoughtfully, smiles and words of encouragement are offered; one is, one feels, doing the right thing.

But like losing weight, the results of evaluation are rarely what one had originally hoped for. Having talked with literally hundreds of HRD professionals about evaluation, I feel compelled, like the little boy who saw the naked Emperor, to blurt out certain self-evident truths:

(1) Evaluation is widely held to be a "good thing" and equally widely done only under duress;

(2) Actually doing evaluation is typically a costly, time-consuming, messy pain-in-the-neck;

(3) (Most glaringly obvious of all) The results of even the most methodologically sound evaluations almost never tell you what, as it turns out, you really needed to know.

In short, evaluation is seldom done, hard to do, and the results are rarely really useful.

Hard to believe? Check it out: how many people do you know (outside of academe) who have actually done evaluations? Of those, how many did it because they were required (by law, regulators, or management) to do so? And how many got results that were both convincing and made a real difference to the decision-makers to whom the results were presented?

But just as there was nothing inevitable about the naked Emperor — It is not that he could not wear clothes, just that he observably was not wearing them — I submit that these "truths" about evaluation need not be the case. That they observably are the case, I suggest, is the result of our having borrowed the classic but inappropriate "truth-seeking" paradigm of evaluation from academic research. There is an alternative to this "truth-seeking" paradigm, namely the equally classic but less widely known "pragmatic" paradigm of research.[1] This article intends to contrast these two paradigms, show why the pragmatic paradigm is by far the better choice as a framework for doing evaluation in HRD, and present a step-by-step sequence of eight "guide-questions" that illuminate how to do pragmatic evaluations.

The Truth-Seeking Paradigm

Consider the following familiar scenario:

VP Personnel: "We've been running that supervisory skills program for over a year now. How much good is it doing us?"

Training Director: "We've got really great feedback on the sessions — the participants are enthusiastic, they say they learned a lot, we've got a waiting list. . . ."

VPP: Fine, but we need some harder data than that — there's a recession on, you know, got to justify all our expenditures to corporate."

TD: "Right. When do you need the report?" (*Fade, cut to . . .*)

Training Director: "Corporate wants an evaluation of the supervisory skills program, by yesterday."

Super Trainer: "(Numerous expletives deleted) . . . but I guess we have no choice, right?"

TD: "Right. So where do we start?"

ST: "Well, let's look at the program objectives. Number One: 'Participants will understand the role of supervisor at Amalgamated Assets.' Hmmm . . . how are we going to measure that?"

This scenario, or some close cousin, is played out in thousands of organizations each year. On the face of it, it looks straightforward and sensible. But a closer look will reveal that the players, with the best of intentions, have already made a key mistake that will almost certainly result in an evaluation that does not accomplish what it was originally intended to accomplish.

The mistake is easy to see. Notice that when the VP personnel began this evaluation process, she had some issues and concerns that needed addressing, and saw "evaluation" as a means of addressing them. As is typical, her concerns were not sharply focused to start with; also typically, they had to do both with making intelligent allocation decisions and fulfilling the CYA imperative. In short, she was asking for information that would be useful in certain specific ways. But by the time Super Trainer started designing, the emphasis had shifted away from her needs for information, onto the training program itself. The name of the game for Super Trainer is "evaluate the supervisory skills program", and that's what will happen, whether it addresses the original concerns or not. Typically at best it will do a haphazard job of addressing them.

Super Trainer fell into the trap of assuming that the foundational purpose of evaluation is the same as that of academic research, namely, to find out what is the truth of this matter. What really happened as a result of that supervisory skills program? Has it really increased production? How much good has this program actually done us? This "truth-seeking" approach keeps the primary focus on the issue of truth — and, although this may be an appropriate model for academic research (even that is questionable, e.g. Ossorio, 1980), it is demonstrably inappropriate for a model of HRD evaluation.

To see this, it is useful to notice a few of the assumptions and commitments of the "truth-seeking" paradigm:

(1) The truth-seeking approach is fundamentally past-anchored: an event (e.g., a training session) took place in the past; it has already had certain effects; and our task is to ascertain what these effects were. The time-span with which this approach concerns itself is from the past to the present, with only incidental and informal concern for the future (e.g. "projecting future trends" from past data).

(2) The truth-seeking approach concerns itself primarily with sorting out the one actual version of reality from among the many possible versions. Thus a researcher takes great pains to eliminate potential sources of "error" or confounding, by focusing on issues of design, methodology, and analysis: How you evaluate at times seems almost more important than what you are evaluating.

(3) The "truth-seeking" task is essentially accomplished when the "data" are collected, analyzed and reported: "What actually happened was. . . ." Questions like, "OK, now what do we do with all this information?" are seen as somewhat ill-mannered or yokelish. The truth is the truth, after all. Do with it what you like.

By contrast, consider the following assertions, which I take to characterize what is actually the case in HRD evaluation:

(1) Organizations that survive are fundamentally future-anchored: they concern themselves with the past only insofar as it helps them in the future. Thus, evaluations are always requested because somebody believes that information will be helpful at some,

usually specific, time in the future (if only to justify future budget requests!).

(2) Organizations rarely have much stake in having exact, unchallengeable accounts of "the truth." Managers and administrators are accustomed to dealing with that irreducible kernel of uncertainty that characterizes real-world functioning; what they require is reasonable evidence, not irrefutible proof. Highly "rigorous" methodology and design are typically not appropriate in HRD evaluation; using them is like using a microtome to slice cheese. Great tool — wrong job.

(3) Managers and administrators are keenly aware of the absurdity of thinking that data represent a unified truth. They are awash in reports, statistics, "facts" of all sorts, many of which contradict each other and some of which seem to mean nothing at all. What organizations require are sharply-focused data which serve as information that enables them to make important decisions.

It seems apparent, on the face of it, that the assumptions of the "truth-seeking" enterprise are a poor match indeed to those of organizations. I believe that HRD evaluations have used this paradigm primarily because they are not aware of an alternative: as a gambler once put it, "Sure, I know the game's rigged — but it's the only game in town!" But there in fact is an alternative:

The Pragmatic Paradigm

Just as the question "What is the truth of this matter?" characterizes the essence of the "truth-seeking" paradigm, the "pragmatic" paradigm can also be seen as addressing itself to one essential question, namely:

"What should we do to accomplish what we want to accomplish?"

In other words, the pragmatic approach takes as its fundamental task the gathering of information that will enable us to choose intelligently among possible courses of action. Shall we continue the sales training program as is, expand it, or put the money into a new incentive program? Should we regulators certify this plant for compliance? How much should training's budget be next year — and what kind of review should the training director get? What should be the emphasis of Super Trainer's own professional development program? These are all decisions for which an evaluation study could provide useful information — if the study were designed with that end clearly and constantly in view. Designs that maximize the utility of the data are characteristic of the pragmatic paradigm.

A primary focus on utility is not, of course, the same as "rigging the data." Truth is part of the pragmatic paradigm, not as its focus, but as a check or "boundary condition" on the data: false information is rarely useful. But merely being "true" is not enough; from the pragmatic perspective data must be both true and useful.

A considerable body of literature exists on how to do evaluation research from the "truth-seeking" perspective: methodology, design, etc. By contrast, comparatively little has been written about the "pragmatic" paradigm.[2] To assist evaluators who want to maximize the actual usefulness of their results, I offer the following sequence of eight "guide-questions." Experience has shown that answering these questions thoroughly and sequentially leads easily to designs that are straightforwardly doable and which yield useful results.

1. *What are the results of the evaluation intended to be used for?*

This is the primary question of pragmatic evaluation, which anchors the research in utility rather than in truth-seeking. It should be answered thoroughly and in thoughtful detail before going further; the other questions refer to and build on the answer to this one.

Experience shows that good answers to this question always take the form: "To enable Smith to decide about Y."

In other words, the results of your evaluation study are intended to enable some person or persons (your boss, the president or chief administrator, board of directors, customers, regulators, yourself, etc.) to decide about some course of action (give you a raise or promotion, increase funding for training, purchase the package, revise a module, buy your product in the future, certify the plant, etc.). Vague words like "management" are rarely adequate here — which specific persons in management will be the decision-makers? The more specific your answer to this question, the more focused and impactful your results can be.

Typically, of course, you may want your data to address more than one issue. Nonetheless, since trying to do all things for all people may result in doing nothing for anybody, it is wise to prioritize and focus your design to ensure achieving your top priority.

2. *What kind of information counts with "Smith"?*

The "truth," like Baskin-Robbins ice cream, comes in a marvelous variety of flavors. Just as some people love chocolate but loathe butter pecan, every decision-maker is impressed with "truth" as found in some kinds of information, and totally unimpressed with other kinds. The pragmatic paradigm suggests a simple rule: If you want your evaluation to count with a chocolate-lover, feed him chocolate — even if you are morally convinced that butter pecan is really the best.

Case in point (sadly, an actual one): A young industrial engineer was called upon to evaluate a new, controversial quality-control procedure in a labor-intensive manufacturing environment. He did a classic study, generating solid data and curves that demonstrated clearly the superiority of the new procedure. He took the report into the president, expecting congratulations on a job well-done — and emerged a half-hour later having experienced the most thorough chewing-out of his career. His boss, afraid of losing a quality employee, asked me as the plant's organization development consultant to find out what went wrong. In a nutshell, what the president told me was:

"These damn college kids, think they can come in here and snow me

with their charts and numbers, hell, I was using figures to lie to the board before that punk kid was born! Tell you how I find out if a new procedure is worth a damn, I let it run for a while and then I go talk to (he named six or seven long-time employees), they can tell me if it's any good. These people have been around!"

On the other side of the coin, I once heard the chief executive officer of a large bank remark to his training director about a management-development program: "Forget what the 'old hands' think of it — you'll never change them. What I want to know is: what impact is this having on our bright young people, who we count on to take over in 10 years?"

Technical supervisors have told me: "If you want to get through to an engineer, put it on a curve." Many HRD professionals profoundly distrust statistics but are swayed by personal anecdotes. And so on. The point is simple: Once you know by whom your data will be used, find out what kind of data they actually find useful. Then give them (at last) that kind of data.

3. *What shall I assess to get that kind of information?*

Now that you know what kind of information you need, it is relatively easy to figure out *what* to assess: knowledge or skills gained, on-the-job performance, supervisory report, peer relations, customer complaints, reject percentages, etc., etc. Notice that, by taking this approach, what you actually assess may bear great, little or no resemblance to the "objectives" of the program. One could argue that this simply means that the objectives were poorly thought out in the first place. I would agree that program objectives tend to be written with a too narrow and parochial focus; however, even objectives that are written in terms of the contribution made to the overall mission of the organization will inevitably not address some legitimate "outcomes" that one could and perhaps should evaluate. While the truth may be one, human activities are many and various — so there's actually no reason to expect that simply "assessing the objectives" will yield useful information.

4. *What constraints on the evaluation exist?*

There is little point in designing a study if you don't have time to carry it out ("The board meets next month!"), regulations or labor contracts prevent you from doing part of it (think of all the tests that you *can't* use with your employees), or if the methodology is too expensive (I once heard a consultant recommend a $75,000 computer-based "simulation" study to evaluate a $47,000 management development program!). Knowing in advance what you can't do can help to clarify what you can do.

5. *Whose cooperation, sanction, approval, etc. are needed?*

Obviously, we need approval of an evaluation study from our management. Equally obviously, we need at least sanction from other managers if we are going to come around asking questions of their

"When approached from the pragmatic paradigm, evaluation turns out to be, in most cases, considerably less messy, time-consuming, and painful."

employees or rating performance in their shop. And if we need access to reports, or computer support to analyze our data, we had best get cooperation from the people responsible for those functions.

But most crucial of all — and most frequently, in my experience, taken for granted — is the cooperation and involvement required from the people who will be giving us our data: to use the jargon term, our "subjects." It is easy to forget that just because this study is important to us, does not automatically make it important or even sensible to them. Example: as a graduate student I helped evaluate a federally-funded seminar internship program. We developed a beautiful questionnaire utilizing Likert scales and semantic differentials; the items tied to concepts that were theory-based, our design was impeccably orthodox, and we had worked out a sophisticated multiple-regression analysis technique. Unfortunately (this was in the late '60s, when doing your own thing had the force of a moral imperative) our interns were bored with questionnaires; the return rate was less than 20 percent. The lesson was clear in retrospect: Always include, as part of the design, a plan for motivating respondents to respond.

6. *How shall I collect the data?*

Now we get down to the "nuts and bolts" of the evaluation process: methodology, design, measurement techniques, etc. Common sense is your best guide here. Nonetheless, this is the "truth-seeking" paradigm's strongest suit, so it seems sensible to borrow from its considerable literature whatever is useful (always remembering not to borrow a microtome to slice cheese). A particularly good book on design is Campbell and Stanley's *Experimental and Quasi-Experimental Designs for Research;* as it turns out, virtually all HRD evaluations are appropriately "quasi-experimental."

Note, by the way, that this sixth question of the pragmatic paradigm is usually the first question of the truth-seeking approach. That seems to be taking far too much for granted.

7. *How shall I analyze my data?*

In thinking through how to analyze the data, it is useful to keep in mind the GIGO principle of computing: Garbage In, Garbage Out. Unfortunately, from the point of view of most sophisticated statistical methods, HRD evaluation data is crude, unrefined garbage — so we are wasting our time using such methods. (Fortunately, for *practical* purposes we rarely need true "ratio" scales anymore than we need a Lamborghini to transport the little league team.)

Stick to the minimum analysis needed to do the job; pick the simplest statistic (usually nonparametric) that will tell you what you need to know; and be certain that your data *in fact* meet the requirements for using this method. And then, of course, make sure you have access to computing facilities that support your chosen analysis: it's hard to do multiple-regression in COBOL, for example.

8. *How shall I use my data?*

We began by asking *what* the data are to be used for; we conclude by asking how the data are to be used. The "truth-seeking" approach is to document the study with a well-written report. But we're after more than that; we want people actually to use the results in making decisions, so we need to present the results in a way that facilitates that use.

I recommend adopting an "action research" strategy here: convene a utilization meeting with the decision-makers, at which the results are presented, processed, and tentative action conclusions reached. The key is to ensure that the decision-makers engage with the data, thinking through implications and reservations, instead of merely accepting or rejecting our conclusions. The considerable body of literature on "action research" is a good support in thinking this through.

A Final Note

When approached from the pragmatic paradigm, evaluation turns out to be, in most cases, considerably less messy, time-consuming, and painful. Even so, I expect that the first "self-evident truth" will remain true: people will almost certainly continue to do evaluation only under duress. But when they *do* evaluate, a pragmatic evaluation will help to make their results *actually* useful — and that's clothing enough for any emperor.

REFERENCES

1. The distinction between the "truth-seeking" and "pragmatic" paradigms was taught to the author by Dr. Peter G. Ossorio of the University of Colorado, on whose work in Descriptive Psychology this article is based (Ossorio, 1964). It is a distinction that dates back at least to Aristotle's "pure" and "practical" reasoning.
2. While preparing this article, the author encountered Michael Patten's book, *Utilization-Focused Research.* Although it is written from a value-perspective-agency program evaluation — that may need a bit of translation for most HRD evaluators, it nonetheless is replete with tips and wisdom about doing pragmatic research.
3. Ossorio, Peter G., *Explanation, Falsifiability, and Rule-following.* Los Angeles: Linguistic Research Institute, 1964.
4. Ossorio, Peter G., "Evaluation Research." In *Advances in Descriptive Psychology.* Keith E. Davis, Ed., New Haven, JAI Press, 1980.
5. Patten, Michael. *Utilization-focused Research.* Beverly Hills: SAGE Publications, 1979.

Anthony O. Putman is president of Descriptive Systems, a human systems consulting and design firm based in Ann Arbor, Mich. He has been an active consultant to a wide variety of private and public organizations for over 10 years. A speaker at the two most recent ASTD National Conferences, he is the author of numerous articles, and edits the Bulletin of the Society for Descriptive Psychology.