DOCUMENT RESUME

ED 160 652 TM 007 884

AUTHOR Alkin, Marvin C.
TITLE The Title of This Speech is... CSE Report No. 96.
INSTITUTION California Univ., Los Angeles. Center for the Study of Evaluation.
PUB DATE Jan 78
NOTE 25p.; Paper presented at the Annual Meeting of the California Educational Research Association (Burlingame, California, November 18, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Context Clues; *Decision Making; *Educational Assessment; Educational Testing; *Evaluation; Evaluation Criteria; *Evaluation Methods; *Evaluation Needs; Evaluators; Models; Problems; Research Methodology
IDENTIFIERS *Context Oriented Evaluation

ABSTRACT

This speech concerns the dimensions of context-oriented evaluation and begins with a discussion of the differences between measurement, research, statistical methodology, and evaluation. Decision-oriented evaluation is described as a process in which the evaluator has, as a major concern, the responsibility to determine likely decisions prior to data collection or analysis. The assumption that the link between evaluation and decision-making is direct and immediate, and that this approach is appropriate for all situations, is seen as a drawback which limits the decision-oriented evaluation approach. On describing the context-oriented evaluation approach, the author provides a three dimensional method for classifying the evaluation context. The first dimension concerns evaluation intent and entails process, results with a predisposition, and openminded results. The second dimension of evaluation context involves whether evaluations are commissioned for individual or organizational purposes. Several organizational and personal motivation factors in this dimension are discussed. The third dimension, the evaluation structure, entails the extent to which those commissioning the evaluation have a specific set of expectations as to the areas to be investigated, the goals to be examined, and the role of the evaluator. The three dimensions comprise a method for providing an evaluation consistent with the contextual needs of a given situation. (Author/JAC)

Center for the Study of Evaluation

UCLA Graduate School of Education
Los Angeles, California 90024

CSE

CSE CSE CSE

CSE

Center for the Study of Evaluation

The Center for the Study of Evaluation (CSE) is a university-based research and development center established in 1966 by the Department of Health, Education, and Welfare. The mission of the Center has been to conduct research, development, training, and dissemination activities for the purpose of improving the planning and evaluation of education and other social action programs. More specifically, CSE activities include: guidance for educators and researchers in the areas of tests and measures and in educational decision making; development of easy to use evaluation procedures and materials in the form of manuals, workshops and kits; and evaluation of projects for federal, state, local, and university agencies requesting technical assistance in evaluation. The work of CSE during its first decade has clarified the ways in which systematic evaluation can be helpful in the improvement of many kinds of programs.

Information about CSE and its publications may be obtained by writing to

Director, Public Information
Center for the Study of Evaluation
UCLA Graduate School of Education
Los Angeles, California 90024

THE TITLE OF THIS SPEECH IS...

Marvin C. Alkin

CSE Report No. 96
January 1978

Center for the Study of Evaluation
UCLA Graduate School of Education
Los Angeles, California 90024

THE TITLE OF THIS SPEECH IS...[1]

Marvin C. Alkin

The title of this speech is: (choose one)

a) Why Evaluations Don't Work

b) Confessions of an Evaluation Sinner

c) Testimonial of a Twice-born Evaluator

d) A Description of the 3 Dimensions of a Context-oriented Evaluation Matrix

e) Musing and Mutterings of a Middle-aged (Moderately Middle-aged) Evaluator

I used to think many things that I no longer believe. I believed things about my physical abilities. I believed things about the extent to which I could drive myself to further achievements. I believed things about the extent to which certain kinds of accomplishments were important to me. Perhaps even I believed things about my sexual magnetism. And, there were other beliefs. Many of them also have been dispelled, partially as a function (I expect) of the aging process.

Some of the things that I believed about evaluation have also been modified. Perhaps ten years ago, when I left my other disciplinary endeavors in Education and became "born" as an evaluator, I surveyed the field, noted

---

[1] Keynote speech presented at the annual meeting of the California Educational Research Association, Burlingame CA, November 18, 1976.

the extent to which there were so many misguided practitioners, and decided that knowledge by me of the "true" way of evaluation and personal practice of the "right" way of evaluation were not in and of themselves sufficient. Indeed, what was necessary was a massive missionary effort at converting the heathens (and assorted other misguided individuals) to what I called the true religion, "decision-oriented evaluation."

I felt then that there was clearly a need to do something about misconceptions of evaluation. I looked around, and saw the host of misguided evaluators.

Look at them, look at them, thought I, glancing at the host of researchers pretending to be evaluators. Don't they know, don't they know that research is not evaluation?

Webster says (dictionary Webster) that research is "studious inquiry, usually a critical and exhaustive investigation or experimentation, having for its aim the revision of accepted conclusions in the light of newly discovered facts." And, oh how I hammered at that. Experimentation? Was evaluation experimentation? Was the aim of evaluation simply the revision of accepted conclusions? Certainly, evaluation had as its goal something other than adding to the body of scientific knowledge. Surely evaluation had as its aim something other than discovering additional insights into the nature of an entity based upon incontrovertible evidence. Clearly evaluation is different from research. Clearly evaluation meant more than (or less than--depending upon your bias) the accumulation of research findings.

But I felt that this was understood. The distinction between evaluation and research was fairly broadly accepted. The message was that research conditions are not always possible and that research conditions may not even

most appropriate in the conduct of evaluations in real world program settings. Schools are not laboratories; children are not easily experimented with; and the demand for rigorous experimental control frequently is not possible in educational administrative settings. This, however, did not mean that the evaluator should not attempt to approximate research conditions in the field to be the best extent possible. But, research conditions could usually not be met within school contexts. Moreover, the concern of the evaluator as opposed to that of the researcher was the provision of information, the best possible information under the conditions existing within the field setting.

And so the gospel got spread. But there remained some disbelievers. There remained some who still felt that good evaluation was good research and that bad research was, of necessity, bad evaluation--an evaluation could only be classified as "well done" if it was good research and anything thought of as a poor research study must quid pro quo be poor evaluation.

Also, there were the measurement specialists and there were the statisticians.

Look at them, look at them, look at them, thought I, glancing at the host of measurement specialists pretending to be evaluators. Don't they know, don't they know, that evaluation is more than simply measurement? And so, the traditional measurement masters, the measurement establishment if you will, wisely shook their heads, recognized a rising trend, saw this new word "evaluation" coming over the horizon, and with perspicacity and sagaciousness succumbed to the new movement. And how did these traditional masters of surement--the norm-referenced test makers succumb? Simple, they continued writing their measurement texts in precisely the same way, with the chapter

headings almost exactly the same, with the content almost identical, but changed the titles of the books to "Measurement and Evaluation in the Schools", "Measurement and Evaluation in Education and Psychology", "Measurement and Evaluation in Teaching". (And these are, as you know, honest to Gosh titles.)

But, the emphasis was the same. The claimed synonymity of the words "measurement" and "evaluation" was implied in the very title of the books. And, when they infrequently ventured into the real world to do an evaluation, it was simply a question of selecting or devising a measure and reporting the results. That was it. Evaluation was measurement of results. PERIOD.

But there was another failing of the work of the measurement masters--the results. And what results! Results based almost exclusively on the use of norm referenced tests whose match with program objectives was dubious at best.

But there were some, there were some among the measurement specialists who saw the folly of some of this reasoning. There were some among the modern measurement sages, who recognized the error of basing evaluative judgments in school situations upon the results of norm referenced tests that might or might not (usually not) have a relationship to the content taught in educational programs. This cult developed holy words that all throughout the land came to know -- "Behavioral Objective," "Measureable Objective," "Criterion-referenced testing," "Objectives based Evaluation." The dogma was learned by all, (no, perhaps "experienced by all" is a better way of putting it). "Thou shalt write behavioral objectives," "Thou shalt write them and write them and write them." "Thou shalt write them in measureable terms," "Thou shalt write them upon the doorposts of Thy house and upon Thy A127's," moreover, "Thou shalt place them in column F on page 6.1," and so on. The experience was

religious. Small doses of dogma, with similar small doses of practice led to feelings of self-righteousness. Large doses of both led to feelings of indignation toward non-believers.

Yet, the approach remained simplistic. The key to evaluation was measurement. True, a new kind of measurement, but measurement and evaluation were still taken as synonymous terms. To measure (properly measure in criterion reference terms that is), to measure was in essence to perform evaluation.

But again I implored, don't you see that evaluation is far more than just measurement, don't you see that the decisions to be made as a consequence of the evaluation perhaps have as much to do with the way in which the evaluation should be conducted as do the objectives to be measured. Measurement is a part (I grant you, an important part) but nevertheless only a small part of the science (perhaps in reality it is an art) of evaluation.

To no avail. Many heard and many heeded the call to decision-oriented evaluation, but the seductive entrappings of this neatly presented, entertaining, and simple road to evaluation salvation gained many followers. Didn't they know, didn't they know, that measurement is just one part of evaluation?

But the mistaken measurement specialists were not alone in their innocence, there were also the sophisticated statisticians.

Look at them, look at them, thought I, glancing at the host of statisticians acting as though statistical methodology was synonymous with evaluation. At that, most of the stuff that they were peddling was psychological statistics, the statistics of psychological experimentation. What about social research methodology which in many respects is more appropriate for

the kinds of situations faced by most evaluators? What about the statistical procedures of the economists, as they cast light upon cost effectiveness kinds of questions in evaluation? What about the statistical procedures employed in operations research models?

Surely, statistical analysis is an important part of the evaluation operation. But to hear them talk, one might think that the major failing of the Coleman Evaluation was the way in which the analysis was done. Article after article has been written criticizing the analytic procedure and suggesting even more esoteric models for "fine tuning" the data. From this point of view, the failings of the Coleman Evaluation are not related to the appropriateness of the data collected, to the way it was collected, to the attention (or lack of attention) by the evaluators to the political context, or to the attention (or lack of attention) by the evaluators to other external data sources that might be viewed and considered by those in major policy positions. All are incidental to the majesty of the analysis. The sophistication of the statistical sophistry reigns supreme.

My words, and the words of other similarly inclined evaluation prophets, fall on deaf ears. For many, evaluation was synonymous with statistical analysis. Sure there were a few other things that one thought about, but it was primarily statistical analysis. Wouldn't they accept, couldn't they accept, that statistics is just one part of evaluation?

In juxtaposition to these feelings of what evaluation was not, I and others had been pulling together thoughts on what evaluation is. I had been at U.C.L.A. developing, refining, and aging my brand of evaluation stuff. Others had been tilling a similar, but distant, evaluation vine-yard.

Dan Stufflebeam, and the late Mal Provus, to name two, had each been approaching evaluation in a similar way. (Indeed there were some who scoffingly dubbed us fellow "handmaidens of decision makers"--referring of course to our concern for recognizing the specific decision issues involved. ["Handmaiden" thought of Dan Stufflebeam wearing a scanty swimsuit, rose in teeth, laurel on head, carrying bowl of fruit, tripping lightly (?) through woods, following a decision maker]).

We all said similar things about what evaluation is not. Evaluation is not synonymous with research. Measurement alone is not the essence of evaluation. Statistical analysis alone is not the essence of evaluation.

We all believed that the purpose of evaluation was the collection, analysis, and reporting of information relevant for specific decision concerns. Evaluation must be decision-oriented and the activities that the evaluator engages in must continually use the decision concerns as a guiding framework.

We all defined various decision levels or decision types that we felt occurred as a part of the educational decision process. The distinctions between the decision types that each of us identified were based primarily upon differences in point of view and differences in the kinds of school systems that we had been exposed to, or were operating in. Mal Provus, for example, was at the time serving as director of research and evaluation for the Pittsburgh Public Schools and had to be particularly tuned to the needs and organizational functioning of a district such as that. My own experience and orientation, perhaps, was towards the middle-sized kind of school district.

We all tried to develop specific evaluation procedures that could be employed by evaluators in acquiring the kinds of information necessary for each

of the decision types that had been stipulated. In our own ways, each of us attempted to clarify and define precisely the steps that a decision-oriented evaluator must follow in order to be maximally effective. In our own ways, each of us attempted to disseminate these views on evaluation procedures to practitioners. I, for my part, instructed efforts at dissemination of evaluation information, through the Center for the Study of Evaluation, which I directed. I began the process by creating a variety of evaluation training materials -- Evaluation Workshops/Needs Assessment Kit/Other Evaluation Kits for Elementary School Practitioners.

I believe that the decision-oriented view of evaluation enjoyed (and enjoys) some modest success. Many have adopted the viewpoint that evaluation has as its prime function the provision of information for decision making -- (that the evaluator has, as one of his major concerns, to determine the likely decisions prior to commencing evaluation data collection or analysis).

And now, for the confession that I had earlier promised. "Precepts that I have preached are not without their failings." (Let's see, it says here [looking down at notes] "Leave plenty of time for the gasp of astonishment from the audience to subside before continuing"). Yes, many people have seen the decision-oriented evaluation point of view, many people have seen the light. Perhaps the problem is that some have seen too much light.

The naivete and the simplicity that I abhorred in other approaches is equally present in the practice of decision-oriented evaluation as conducted by many. These deficiencies stem primarily from the incorrect assumption by some that the evaluation - decision-making link is clean, firm, direct and immediate. This assumption invariably leads to the understanding that the

evaluator is to determine what decisions are likely, and having made those determinations, he/she conducts his/her evaluation and presents the findings, then sits back and enjoys the fruits of his/her efforts. The assumption is that evaluation reports will be received, believed, and acted upon and when this doesn't happen the evaluator generally feels that he's been deceived and [illegible] upon [illegible] the assumption and [illegible] expectation that evaluation [illegible] and [illegible] impact upon decision making [illegible] Those who [illegible] lack of impact of evaluation on decision [illegible] while failing to recognize that there are other inputs into the decision process, are naive.

[illegible] oriented evaluation

[illegible] It is my feeling that the notion of decision oriented evaluation is [illegible] limiting. Perhaps what [illegible] the assumption that decision oriented evaluation [illegible] What's the decision? And, if there isn't one by golly there ought to be one!

I became acutely aware of [illegible] the deficiencies a number of months ago when I was asked to develop a speech on evaluation which was to become a part of the AERA cassette tape series on evaluation. [illegible] that one of the most crucial activities that the [illegible] place prior to the collection [illegible] prior to the analysis and reporting of the evaluation information to the decision maker. I chose to refer to that activity as "locating the decision context." This was a period in which the [illegible] relevant decision makers, talked about the kinds of concerns they had, helped them to phrase those concerns into decision questions, and then discussed the types of data presentation that would b[illegible]

considered by the decision maker as an adequate response to the decision question that he had in mind. Thus, in essence, I as the evaluator simulated some evaluation results ("suppose I told you that X percent of the students had accomplished such and such an objective "; "suppose I told you that 75% of the parents with children in the program averaged a score of 4.5 on an attitude-toward-program questionnaire where 4 is favorable and 5 signifies highly favorable"), and in the process of this refining the conditions of adequate evaluation information, decision maker and evaluator gained further refinement of the decision question. Typically this refinement would occur as a consequence of a decision maker indicating that the data presented would not really be adequate for him to make the kind of decision that he wanted and in fact that maybe the way in which the decision question had been stated was inaccurate. And so, the evaluator and decision maker would jointly redefine decision questions into a form that more properly reflected what the decision maker (or decision makers) really meant.

I completed that AERA cassette-tape, titled "Framing the Decision Context," but felt somewhat uneasy about the tape. I felt uneasy not because of the quality of the tape, (which of course I was convinced was "great") but rather the uneasiness stemmed from the fact that part of the discussion of the tape dealt with situations that might logically be called framing the non-decision context. That is, the evaluator in the course of the "framing" might determine that indeed in many instances there was no decision that was to be made as a result of the evaluation information, or that the likelihood of decisions was remote. Perhaps another reason for uneasiness was that I was aware of a great deal of non-decision related evaluation activity that perhaps was not discussed within the AERA cassette-tape.

Evaluation does not always lead to decisions that will be made. Occasionally, possibly even frequently, evaluations were never intended to lead to decisions. Thus, what I was driving at (what I was concerned about at the time that I developed that tape speech), was perhaps not the decision context but the necessity for understanding the total context in which the evaluation would be occurring.

University professors are often stereotyped as ready to develop a model, a theory, a formulation, a framework, a conceptualization, a tautologogy, a theoretical formulation, or any of dozens of other esoteric conceptual entities, at the drop of a hat or at the merest indication of the potential possibility that one might be helpful. I would find it helpful to develop a categorization (you'll note that that was not one of the terms I previously ridiculed) showing three dimensions or ways of classifying the evaluation context; there are undoubtedly others that might come to your attention or which might become immediately apparent. For the moment I find this a helpful way to think about the evaluation context. (I can't really say how long it will remain unchanged).

The first dimension of the evaluation context is something that I will refer to as evaluation intent. As I look at the research that I had done on evaluation utilization, at the case studies that I am currently conducting related to this particular topic and to the evaluation projects that I have personally conducted in school districts and in other kinds of organizations, I become acutely aware of the necessity of making the distinction between evaluations which are commissioned for the purpose of obtaining some results (or learning about some outcomes of the program) as opposed to those evaluations in which the major intent is the commissioning of the evaluation itself.

In the latter instance, it is the process of having the evaluation conducted that is important rather than the results that might be obtained from the evaluation. In some respects, the commissioners of the evaluation see it as almost immaterial whether or not the evaluation ultimately produces one set of findings or another (or none at all).

As I've indicated, one potential evaluation intent then is engaging in the process. Obviously another possible dimension of evaluation intent is having a concern for the results or an understanding of the outcomes of the program. But I would like to break this into two categories, the first, a concern for the results with a particular predisposition as to what the nature of those results should be, and the second, having a concern for the results where there is an openmindedness on the part of those commissioning the evaluation.

I have become aware of instances in which those commissioning an evaluation really have done so in order to "show that the program is doing a good job." The intent is not to "determine if the program is doing a good job" but to show. Hopefully, instances of this type are few and lessening. Hopefully, more and more evaluators are refusing to participate in evaluations which have this as the intent.

Thus, I see three categories of evaluation intent: 1) process, 2) results with a predisposition and 3) results (or results openminded, if you wish).

Now let me consider a second dimension of the evaluation context. (Help me decide on a name for this dimension of the evaluation context.) It seems to me that evaluations are commissioned either for individual or for organizational purposes, and moreover, that the variety of organizational purposes

differ rather substantially. For example, I can conceive of an instance where an evaluation might be commissioned to fulfill personal needs of the individual requesting the evaluation. That is, the function is not so much the demonstration of a program but rather it is an action that is taking place and which has been commissioned by a decision maker, primarily because of a view that it might add to his own personal glory, or to his status within the district, or to his feelings about himself, or to the way his colleagues in other districts view him.

I suppose there is a certain amount of the personal element in every situation. When I describe the alternatives within this dimension, I suspect that they are not mutually exclusive, but that rather we're concerned with the extent to which one or another of them is the dominating motivation.

As a second alternative within this dimension, I suppose that one might think that the dominating motivation could be found within the school program itself (perhaps there is an insistence by teachers or by a program director that an evaluation be conducted to find out how "we" are doing). Possibly, the dominating motivation is at the school or district level. Possibly, the prime mover is the principal of the school or the Superintendent of Schools -- although my experience tells me that this is not usually the case because they typically do not want to shake up the system unless it is required from elsewhere.

A further possible motivating force is frequently the school board. I recall performing an evaluation for a school district of an alternative high school. In that case, it was the school board that insisted that an evaluation take place. (Something novel was being introduced into a relatively

conservative school [illegible] and they wanted the a[illegible] of knowing whether it worked -- and if it didn't, they wanted the political insulation of having called for the evaluation that demonstrated it.) Another possibility is that the dominating force, the motivating force, might come from the community itself. Perhaps an example of this might be on highly volatile issues which invoke a great deal of community fervor (e.g. busing) and an evaluation, (generically, some kind of data acquisition and summary activities) -- an evaluation is called for in response to an evaluation motivating force from the community.

A final example of this dimension of the context is the external agency. And I feel that this is a particularly appropriate example because many evaluations, (possibly most evaluations), are commissioned primarily because of an external agency (a state or federal government) and the requirements that agency has placed upon organizations that receive program funds. Let me take this one example and explore it somewhat further since it forms such a major portion of what is done under the label "evaluation."

Government agencies impose reporting requirements on school districts. These reporting requirements demand the presentation of certain kinds of data -- some of which is outcome data. These reporting requirements are generally called by the term "evaluation" by the agencies [illegible]d "evaluation" is in.) The procedures when conceived were felt t[illegible]portant evaluation activities that would be meaningful. Typically, procedures get modified, refined, enlarged, re-refined, re-modified, re-enlarged. And while the intent (usually to provide an information system and an evaluative procedure useful for decision making at local levels and at state levels) may have been clear

to those who initially conceived of the system, each subsequent refinement to the procedure makes the rationale less clear. Soon, those who become enmeshed in the procedural contortions become a captive of the procedures and in a very real sense, the agency is the dominating force. Domination by the agency usually precludes whatever valid decision utility intent that might have initially been present as the cornerstone of the "evaluation procedure." In practice, agency dominated or motivated evaluation frequently has little if any impact on decision making. It would be surprising if this kind of evaluation did have impact. The smallest amount of knowledge of how federal agencies and other large governmental bureaucracies operate, quickly convinces one that this is not an arena in which empirical rationality is always the dominant mode. Furthermore, a little common sense, understanding of the time schedule with which funding decisions get made in agencies at the state and federal level (for example), in juxtaposition with the time of the year in which governmentally required "evaluation reports" are received, quickly convince one that there is no way that such reports can have impact upon funding decisions. Agency-required evaluation reports typically are received from three to six months a the funding decision has already been made.

Typically, the only way that the evaluator can have impact in such programs, in terms of the provision of evaluation information leading to program modifications, is for the evaluator to develop a relationship with project personnel, where project personnel are attentive to the evaluation findings, respect the evaluator and his integrity, and are aware of major evaluation findings, prior to the publication of the final evaluation report, and are

able and willing to incorporate these recommendations into program changes for subsequent years. In many cases, the routinization of the procedures convinces the evaluator that compliance with agency requirements is synonymous with good evaluation and doing more is redundant.

(This belief is reinforced by many practicing school administrators, who believe that evaluation is simply an event that leads to compliance with various agency requirements. There is no real expectation that major basic decisions will be made. The name of the ballgame is simply not to get "dinged" by the governmental agency -- in short to show the agency that you've played the game their way.)

Let me summarize one point with respect to this second dimension, and that is that we are not concerned with who actually commissions the evaluation, but rather the question is who or what is the motivating force for the evaluation being commissioned. In the past, when I have placed emphasis and concern upon designation of the evaluation report commissioner [illegible] son who actually set in motion the selection of an eval[illegible] me that [illegible]he right question in terms of unde[illegible]ng the relationship between the evaluator and the person he was to report to but it missed a very important dimension, and that dimension dealt with who or what was the moving force that necessitated an evaluation. Now what name would you give to this second dimension of the evaluation context? (Shall we call it the evaluation motivating force dimension? Kind of clumsy -- perhaps you'll help me think of a better name for this dimension.)

There is a third dimension to the context of evaluation which I would like to identify. This dimension deals with the extent to which those

commissioning the evaluation have a specific set of expectations as to the areas to be investigated, the goals to be examined, or what it is the evaluator is to do. Perhaps one could think of the ends of a continuum as restricted agenda evaluation or open agenda evaluation. It seems to me that this description goes far beyond the goal-based/goal-free distinction that Michael Scriven has made in his writings, because we are dealing with far more than the goals of the organization. Of concern might also be program characteristics and the extent to which they're implemented. Also included within this category is a general expression of the autonomy granted to the evaluator, pointing at the program to be evaluated and asking him to present a report. The typical program evaluation situation of which most of us are aware has a tightly prescribed agenda within which the evaluator is expected to live. Typically, this tight prescription is related to the [illegible] for completing a report to an external agency. However, even without such a report to an external agency, I have performed evaluations in some instances where school districts have let out requests for proposals that specifically state the activities that the evaluator is to engage in and the kinds of data that he is to collect. (And I am sure that many of you have performed such evaluations as well.)

At the other end of the continuum, I can recall some instances of evaluation situations with "open agendas." In one case, I had been requested by a State Legislative Committee to "evaluate" the educational program of that State's juvenile detention facilities. This evaluation was to take place within a three day period. And there were little if any additional prescriptions presented to me on the way in which this evaluation would take

place, the data would be collected, the way in which my observations would take place, or the way in which my final report would ultimately be framed. I was the evaluator. I had performed evaluations for this State Legislature on a number of other occasions. The Legislature had been impressed and pleased with my work. Thus, in this instance, there seemed to be a belief in the evaluator as an omnicient presence who needed no guidance or framework, who was neither to be bounded by the goals of the program, by reporting requirements and formats, or by specific decision concerns of those responsible for programs. In this instance, what was wanted was a sage, a seer, perhaps an oracle of Delphi. (I hope that when such oracles are commissioned, that those who do so get not the oracle but the true priests. [Story of DELPHI -- stone, smoke rising, etc., a mad illiterate woman, priests who interpreted her ramblings into meaningful statements in line with the context.])

At any rate, what I have identified here is a third dimension of the context of evaluation. Perhaps I might call this third dimension of the evaluation context something like evaluation structure -- implying the structure provided to the evaluation by those who commission it rather than the structure provided to the evaluation by the evaluator in the process of his conducting it.

What it appears to me that I have said thus far is that the notion of decision-oriented evaluation (or of any other single kind of approach to evaluation) is extremely limiting and that instead we ought to be thinking of the context in which the evaluation takes place -- context-oriented evaluation if you will. And that one way of framing that context, one way of describing the context in which evaluation takes place is in terms of three dimensions that I have referred to as evaluation intent, where the primary difference

is between wanting the evaluation because you want the results, as opposed to wanting the evaluation simply to demonstrate that the process has taken place. The second dimension of this context of evaluation is the identification of the evaluation motivating force -- the motivating individual or force for the evaluation being commissioned. The third dimension of the context of evaluation is the evaluation's structure which I defined as the restrictiveness or openness with which the evaluation could be conducted -- that is the initial structure imposed upon the evaluation prior to the evaluator commencing his work.

Now it appears to me that the first and major job of the evaluator is to play to this dominant context, to provide an evaluation consistent with the major contextual needs of a given situation. Or, alternatively if that context is disagreeable, repugnant, ethically unpure or what have you, to either refuse to do the evaluation or (with greater risk) perform the evaluation in a manner not intended to serve the dominant evaluation context. But if the latter course is chosen, if the evaluator chooses to perform an evaluation inconsistent with the dominant context, he ought to do it with a full understanding of the situation and the potential consequences instead of with the naivete that currently exists in evaluators who sometimes perform an evaluation, have high expectations for it and suffer the disappointment of apathy on the part of those who receive the report. If the evaluator is aware of the dominant context and finds it appropriate to work within that context there is no reason why he should not be able to attempt to extend the implications and findings of his work to other appropriate contexts or to even attempt to change the dominant context. The important point, however, is that if evaluation is to in any way be useful we need to be aware of what is the contextual starting point.

Against these various contextual situations there are perhaps a variety of evaluation forms and styles that become imposed. Different evaluators have different ways of approaching the evaluation situation. Sometimes the match with the context is good, sometimes it is poor. When the match of the evaluator with the context is good, he will typically praise himself for an evaluation well done. But when the match is poor, he will typically bemoan the lack of responsiveness of the various program participants to his evaluation.

Now, what about these different forms and styles? Evaluators approach their task in different ways. Some evaluators have been labeled or have labeled themselves as decision-oriented evaluators (they must use different models to describe the steps in the sequence as they see them) and in each situation they must struggle to identify the potential decisions even in situations where none exist. Some evaluators prefer the role of seer and try to impose that upon all situations. Some evaluators readily accept the role of "federal form filler-outers" and "state statute satisfiers" and willingly do what is necessary to satisfy those requirements. That too is evaluation.

Context oriented evaluation involves the necessity of recognizing contextual needs and of matching evaluative orientation and style to contextual needs. To the extent that an evaluator has the insight and the abilities to modify his style in his response to the contextual situation, he will be more successful. Generally, the evaluator cannot do very much more than the context will allow. Generally, evaluation will not be more than the context will allow.

If someone should ask you (I can't imagine who would) what that speech by Alkin was all about, it seems to me that you might tell them that it dealt with why evaluations don't work, that there was a certain amount of confession on his part as to how he felt he had erred in his evaluation prescriptions of the past, there were a lot of musings and mutterings, and substantively he focused on context-oriented evaluation -- the notion that we ought to be aware of the context in which evaluation takes place and be responsive to that context.