DOCUMENT RESUME

ED 115 656 TM 004 926

AUTHOR	Anderson, Edwin R.
TITLE	Evaluation: The Justification of Practice.
INSTITUTION	Washington Univ., Seattle. Educational Assessment Center.
PUB DATE	May 75
NOTE	23p.; Not available in hard copy due to marginal legibility of original document
EDRS PRICE	MF-$0.76 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS	*Definitions; *Evaluation; *Evaluation Methods; *Evaluative Thinking; *Models; Professional Personnel; Training

ABSTRACT

Evaluation can be considered as a general activity not specifically confined to educational settings. This paper considers the assumptions leading up to a definition of value, defines value within the framework of the stated assumptions, and discusses evaluation as a conceptual, rationalizing activity. Terms defined in the paper's first section include infinite regress, distinction, potential set, variety, constraint, decision, conflict of interest, and value. An evaluation model is developed which regards the matching of concepts and world states as the key feature of evaluation. Section 2 of the paper draws implications from the general evaluation model for the specific role of the educational evaluator. Section 3 considers the training needed by an evaluator if the general model is accepted. (Author)

Educational Assessment Center

University of Washington

May 1975

Evaluation: The Justification of Practice

Edwin R. Anderson

Abstract

Evaluation can be considered as a general activity not specifically confined to educational settings. This paper considers the assumptions leading up to a definition of value, defines value within the framework of the stated assumptions, and discusses evaluation as a conceptual, rationalizing activity. Terms defined in the paper's first section include infinite regress, distinction, potential set, variety, constraint, decision, conflict of interest, and value. An evaluation model is developed which regards the matching of concepts and world states as the key feature of evaluation. Section two of the paper draws implications from the general evaluation model for the specific role of the educational evaluator. Section three considers the training needed by an evaluator if the general model is accepted.

Bureau of Testing Project: 297

Evaluation: The Justification of Practice

Edwin R. Anderson

University of Washington[1]

"However unreasonable or immoral an action may be, man has an insuperable urge to rationalize it--that is, to prove to himself and to others that his action is determined by reason, common sense, or at least conventional morality."

Erich Fromm, The Sane Society

Educational evaluators supply justifications for the practices and materials used in educational settings. Many of these justifications produced by evaluators are responses to questions from a question class having the form, why this and not that? For example, an evaluator might ask, should we use textbook A or textbook B in this program? A question, however, can be responded to in two ways: we can accept the question as stated and attempt to answer it, or we can look back from the question to the conceptual framework that produces it. Let us choose the latter recponse.

"Why?" is one of the most primitive and powerful questions we ask. Answering it leads directly into the nature of human thought, for in answering we connect the thought or action being justified with other thoughts and actions. The structure of our thought is dictated by the kinds of connections we are willing to accept. Since it is philosophy that has most rigorously searched for ways of discussing thought connections and implications, we must enter the domain of the philosopher.

I intend to accomplish three things by leading evaluators into the philosophical thicket.

(a) Although all evaluators can remember long hours of data dredging, evaluation is primarily a conceptual activity. Most of this paper is devoted to concepts appropriate to a formal theory of evaluation.

(b) Near the end of the paper I shall draw some conclusions from the theory about the role of the practicing evaluator.

(c) I'll draw some conclusions about the training evaluators need to carry out the roles assigned to them in section (b). Statements (b) and (c) seem oriented

[1] The author thanks Dr. Gerald Gillmore, Dr. John Smith, and Dr. Mildred Kersh for their thoughtful criticisms of an earlier draft of this paper.

toward practical evaluation, but the reader will find material contained in those sections to be general, just not as general as that in section (a). I'm leaving the connection of the theory to specific evaluation practice as the topic of a later paper.

A Formal Philosophy of Evaluation

Logical thought has direction

Philosophers use two general analytic methods. The first method starts from propositions and spreads outward in an effort to justify the premises leading to the propositions. A proposition is justified by finding an antecedent argument that cannot be denied and that is logically connected to the statement being justified. The "inside-out" philosophers, as I shall them, act as if they were at the center of an onion and were moving outward through the layers of the onion to find the outside layer. Each onion layer represents an antecedent concept for layers contained within it. The outside layer represents the absolute truth or the ultimate good. The inside-out method is losing favor in the philosophical community because practitioners of the method can't seem to find the required outside layer. In fact, Godel, a mathematician, has been able to show that at least one statement within any symbol system cannot be proven using the symbols of the system and therefore must be assumed (Nagel & Newman, 1958). Since propositions are symbolic, we shall remain disappointed if we seek the ultimate proposition. Thought is an open system if pursued from the inside-out; there are no fixed outer boundaries to conceptual structure.

The second method of analysis starts with a given, agreed upon, set of premises or axioms and then moves within the axioms to draw out their implications. This is the method used in symbolic logic and mathematics. The philosopher acts as if he were on the outside of the onion seeking to find its center. The "outside-in" philosopher knows by definition what the outside is, and so he has an advantage over "inside-out" philosophers. Thought can be a closed system, rationally constructed, if pursued from the outside-in.

The relevance of the direction of movement metaphor can be demonstrated in the following evaluation example. Suppose a two person team composed of one course developer and one inside evaluator. They develop a course, install it, and report on the effectiveness of the course; the team and their work form the inside of the onion. A third person, after looking at the work of the developer and evaluator, decides he would like an outside evaluator's opinion, a fourth person; he would like the evaluator's work evaluated. A fifth person could decide to evaluate the fourth person's evaluation of the first evaluator's report. In the solution to an analogous control problem, Stafford Beer (1972) makes a case for carrying this type of layered analysis back through five levels. He admitted that there was no logical reason for stopping at five layers; there are, of course, practical reasons for doing so. The point: a typical analysis of evaluation is an inside-out analysis and as such is an open conceptual structure We must think from the outside-in if we are to resolve the dilemma raised by asking, who will evaluate the evaluators?

Distinction

Nearly all introductory psychology textbooks have a chapter discussing human perceptual processes and in that chapter there is generally a drawing which illustrates the phenomenon of reversible figure-ground relationship. In figure 1 for example, if the white is seen as figure and the black as ground, then a challice is seen in the drawing. If the white is seen as ground and the black as figure, the drawing shows two silhouetted faces looking at each other. In figure 2 the same figure-ground switching can be observed. The white figure appears if the black area is seen as the ground and the black figure emerges if the white area is seen as the ground. These drawings illustrate three facts of perception:
(a) we perceive only when a distinction is made between the figure and ground,
(b) we can change what we perceive by switching the figure to ground, and
(c) we cannot perceive both figures at the same time. Distinction is the act of creating a perceptual or conceptual figure-ground relationship. We can think of the ground created by a distinction as a frame of reference which allows us to perceive the object framed. Now think of the frame of reference as conceptual rather than perceptual; we "see", i.e. understand, what our frame of reference allows us "to see."

Figure 1

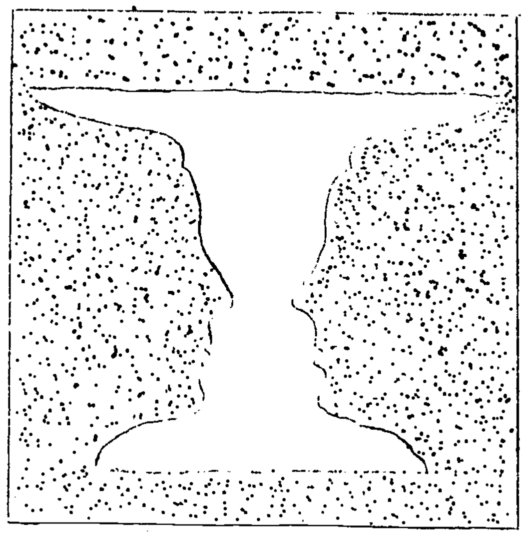
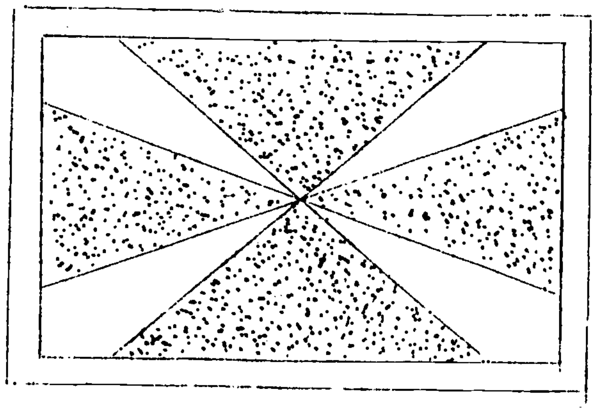


Figure 2



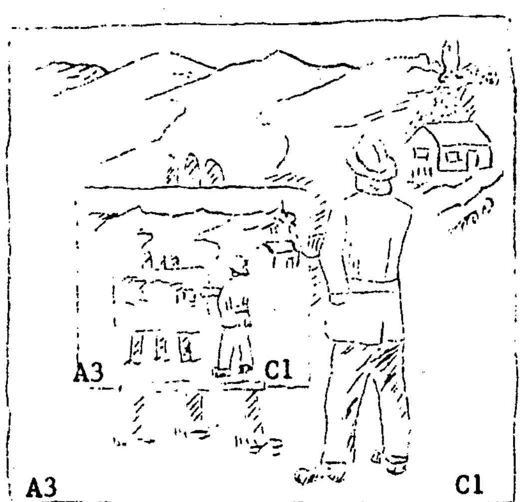
Infinite regress

A second visual metaphor is introduced in figure 3. An artist is happily painting a landscape when he realizes something is missing. He is not in his picture of the landscape but he is in the real landscape; therefore, the picture is incomplete. He then steps back and paints himself painting a landscape. And so on. Figure 3 provides a visual representation of this infinite regress. The regress is based on a distinction made by the painter, the "me/not me" distinction we commonly make in Western society. As soon as the artist paints himself painting the picture, he chops that off and calls it "not me" hence the picture is always incomplete since it never includes "me." A regress results from a distinction being repeatedly applied to itself. The field of evaluation is plagued by such distinctions. For example, who will evaluate the evaluators, and how will we decide that the good is good? If our outside-in approach to evaluation is successful, we shall avoid the problems of the infinite regress by not applying a distinction to itself.

Distinctions are arbitrary

Men are capable of making distinctions, but these distinctions are arbitrary. Consider some examples offered by Alan Watts (1966). Most of us assume our bodies to be a part of the "me" and the air around us to be a part of the "not me," but can we really imagine a functioning human apart from the air? Clearly the distinction is in some sense artificial for I might just as well imagine the

Figure 3



air to be a part of "me." The same is true of food; can we seriously imagine ourselves existing apart from food? The distinction of "me/not me" and the assignment of the world to these classes is clearly arbitrary, but often, because we all agree to make the distinction, we forget or are even unaware of its arbitrariness. The same conclusion can be drawn about other distinctions we commonly make. A cat's head and a cat's neck are normally distinguished from each other, but you can't show the exact place where the head stops and the neck begins. Here again we have an artificial split of an integrated system.

The arbitrariness of distinction, i.e. boundary drawing, doesn't mean that making distinctions can be avoided. Much human behavior has at its base the making of arbitrary distinctions and the agreement among people to use the same distinctions. Evaluation is one such behavior. Evaluators, for example, follow the convention of using "sound experimental design practices" where possible in their justifications, however, experimental design is itself a conceptual structure based on agreement. The roots of experimental design as a method of inquiry are laid in distinctions as arbitrary as those of any other conceptual system (Kaplan, 1964).

Time as an arbitrary distinction

Physiological studies show that men and other animals make temporary maps of the environment in their central nervous systems (Pribram, 1971). When an animal is in an environment and a new stimulus is introduced into that environment, e.g. a regular click, records of the animals brain activity show changes that correspond to the click. The amplitude of the change associated with each click decreases as the animal adapts to the click. Eventually the click is ignored. When the clicking terminates, the animals adapted to the click again show brain activity. The adapted animal has represented the environment in his nervous system and expects the click to occur, i.e. the mental representation is running slightly ahead of the environment. Adaptation to the click does not have to occur in humans. Zen buddhists, while meditating, seem to be able to suspend the activity of representing the environment (Kasamatsu & Hiroi, 1966). The Zen practitioners show brain activity of undiminished amplitude each time the click occurs. The verbal descriptions the Zen masters give when asked about the state of consciousness that correlated with the non-adaptation data offer clear reason to suppose they somehow manage to drop the time frame from their perception. Thus the time frame we all take for granted is related to our ability to run the mental representation of the environment in ways not dictated by environmental stimuli. When we run the representation faster than the environment, we create the ground for a concept of the future. The running activity is usually labeled as imagination. Our recollections of environmental events which are not now present, i.e. memory, leads us to a conception of the past. Note in both cases that the time frame develops from the occurance of mental events which indicate environmental events not immediately apprehended by the senses. The distinction drawn between mental events and environmental events is another artificial split of an integrated system therefore the use of these terms to develop the time frame means the time frame itself is arbitrary.

Potential, Variety, and Constraint

The question, why this and not that?, depends on the grounds of distinction and of time. We imagine some activity labeled "this" and subsequently imagine some activity labeled "that," "this" and "that" being distinct. We imagine further that both activities could occur in the future but that they are mutually exclusive in the same future. "This" and "that" are elements of a potential set.

> Given a set P with elements that represent distinct, imagined activities. P is a potential set if all the imagined activities of the set are seen as possibly existing but as mutually exclusive in a given future time frame. The number of distinct activities included in the potential set P is called the variety of the set. A constraint is introduced whenever the variety of the potential set is reduced.

As an example of a potential set, suppose an introductory psychology instructor had twelve textbooks that he was considering as candidates for use in next quarter's course. P then consists of (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12) where each number represents the selection of a particular text for use. The variety of the potential set is twelve. Suppose further that the instructor decides not to use a text published before 1970 and that texts 10, 11, and 12 were published in 1969. The potential set is now reduced to (1, 2, 3, 4, 5, 6, 7, 8, 9) therefore the elimination of texts published before 1970 is a constraint. By introducing the concept of potential early in the discussion of evaluation, we can later capitalize on the existence of powerful methods of logic already developed on the basis of this concept. For example, potential is a key concept in statistics, experimental design, decision theory, computing, control theory, measure theory, and symbolic logic.

<u>Elimination of the infinite regress</u>

The various constraints which could be applied to P to reduce its variety can be formed into a potential set of their own. Let C be the set of constraints which can be applied to P. Let a line between an element of C and an element of P indicate that the element of C where the line or lines originates selects the element or elements of P where the line or lines terminate. For example, the symbols

C(e, f, g, h, i, j, k)

P(a, b, c, d)

Indicate that constraint f selects element a from P and constraint h selects elements b and d from P. Note that the set of possible constraints may have greater variety than the potential set. This increased variety is another way of pointing to the openness of inside-out thinking. But C is also a potential set and has some set of constraints C' which select from it. For example,

C' (l, m, n, o, p, q, r)

C(e, f, g, h, i, j, k)

P(a, b, c, d).

This notion can be extended to C'', C''', and so on. We have introduced an infinite regress to the process of selecting an activity to be implemented from P.

The problem of justifying our selection of some activity from a set of activities can be viewed as a special case of the symbol system outlined in the preceding paragraph. When someone asks, "Why do you select 'this' and not 'that' "?, the answer takes the following form, "Because some constraint existed which could not be violated." For example, "Why did you select textbook A and not textbook B?" The answer comes, "Because textbook A covered an important topic more thoroughly than textbook B." The constraint is present in the concealed command, select that textbook which offers the best coverage of the important topic. However, note that we could continue asking, "Why?", thus seeking justification for the use of that particular constraint. We are now removed to the level of C'. Justification, i.e. the answering of the question, Why?, is logically an infinite regress. It is clear that "Why this and not that?" is an inside-out philosopher's question.

How do we stop an infinite regress? By giving up the activity which produces the regress, in this case by giving up justification. We choose an element from some potential set by agreement; we make it possible to become outside-in philosophers. A <u>decision</u> is the selection of an element from a potential set without justification for the selection. A decision is a constraint, it reduces potential, but there is no constraint applied to the decision; some constraint or constraints must be left unjustified. Those constraints are treated like axioms in mathematics, i.e. they are stated and agreed upon.

<u>The holders of the concepts in the potential sets</u>

Once we begin an infinite regress by adopting an inside-out direction through a conceptual structure, we have no logical reason for stopping the regress; however, concepts are held by people and people quickly tire of treading a conceptual regress. The logic of the regress gets terminated then when a person or a group of people are not required to justify their constraining concepts. We can see how this occurs by labeling the layers of our potential set diagram with the holders of the concepts represented by the diagram.

CONCEPTS	CONCEPT HOLDERS
C'(l, m, n, o, p, q, r)	Outside evaluator
C(e, f, g, h, i, j, k)	Program developer/ inside evaluator
P(a, b, c, d)	Program developer/ inside evaluator

CONGRUENCE

In this diagram l, n, and p are unjustified constraints because no one evaluates the outside evaluators. We can change from inside-out thinking to outside-in thinking by starting with the unjustified constraints the outside evaluator adopts. For him the system can be viewed as logically closed. The diagram shows that the outside evaluator and the inside people will agree on the selection of a, b, and d from the potential set.

Disagreement can be shown by the following diagram:

CONCEPTS	CONCEPT HOLDERS
C' (l, m, n, o, p, q, r)	Outside evaluator
C(e, f, g, h, i, j, k)	Program developer/ inside evaluator
P(a, b, c, d)	Program developer/ inside evaluator

TOTAL INCONGRUENCE

The outside evaluator holds l and n as unjustified constraints; a and b are his final selection from P. The inside people hold constraints i and j which select c and d from P. A <u>conflict of interest</u> exists when constraints held by different members of a group of people concerned with a system of concepts select different elements of the potential set. The degree of conflict is related to the number of mutually selected elements from P, i.e. partial overlap contains less conflict than the total incongruence diagram.

How are these conflicts to be resolved? Resolution comes through the use of two kinds of personal power, persuasion and authority. The authority resolution is followed by the course developer and inside evaluator recognizing that final approval must come from an outside source. If they want to produce congruence with that outside source, they must include the constraints the outside source will apply within their own set of constraints. This has the logical effect of putting the course developer and inside evaluator at the same conceptual starting point as the outside person. They can both follow outside-in logic from the same starting assumptions. For example, suppose a school principal who has final approval over an introductory science program doesn't believe multiple choice tests effectively measure scientific knowledge. The course developer might as well resign himself to accepting the same constraint. The persuasion resolution to conflict occurs when the course developer uses what constraints he feels are appropriate. He relies on his ability to persuade any person having an interest in his product that his constraints are reasonable. For example, the course developer whose principal doesn't like multiple choice questions can include multiple choice questions in his course if he feels he can later persuade the principal to accept them or if he is willing to test the principal's authority.

The activities leading to the development and constraint of potential sets are conceptual in nature. The holder or holders of the constraining concepts and the relationship those holders have to each other must be identified. We must know who selects unjustified constraints and what exact constraints have been chosen. When we have done this, we can design a course development plan and an evaluation plan which has stated boundaries and which moves "outside-in" conceptually. We escape from logical regress problems through identification of the people concerned with course or program development and implementation.

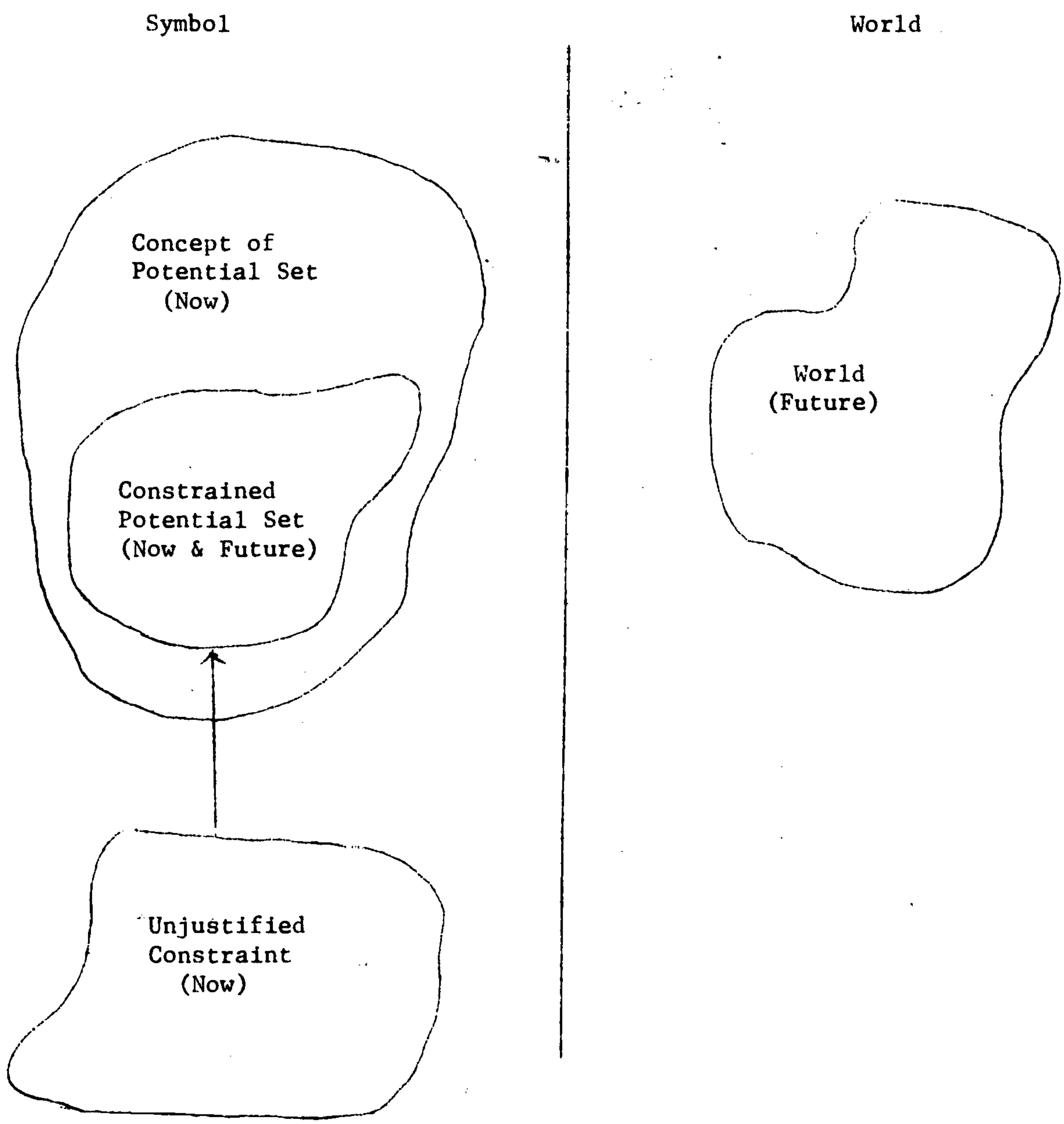
Definition of value.

We have not yet defined value and thus we have not yet determined the essence of evaluation as an activity. To do this, we need to draw a distinction between world events and symbolic events. On the symbol side of the distinction we have the potential set, the constrained set, and the unjustified constraint. On the world side we have a circle representing the world as it will actually occur in the future (See figure 4). We are concerned with the match between the constrained potential set and the future world. If the world matches the constrained potential set, then we say the world is good; if not, then the world is bad.[2] In the words of R. S. Hartman (1959), "A thing is good when it fulfills the definition of its concept." Hartman's statement is an axiom and as such is just one of many value axioms that are possible. We must agree on the fundamental value axiom, whatever it may be, if we are to agree on the nature of evaluation activity.

[2] The possible states of the world or the constrained potential set can be seen as ordered. "Good" and "bad" are not confined to being dichotomous. Gray "good" can be permitted.

Figure 4

Diagram of the Symbol-world Distinction



Some clarification of Hartman's value axiom is necessary. Examples are required.

Example 1. If an autombile is viewed as a transportation device for two people, as a device with four wheels, and as a self-propelled device which gets twenty-five miles per gallon of gasoline, then a cadillac is not a good automobile for a cadillac doesn't fulfil the definition of the concept. A Honda would be a good automobile. Note that a table is a bad automobile along with the Cadillac.

Example 2. If a table is viewed as a horizontal surface supported by four legs, then most tables are good tables. The same table that was a bad automobile may be a good table. To quote Hartman (1959) further, "Good is a property not of objects but of concepts. When a person understands that a thing 'is good', it is not necessary that he know anything of the thing in questions: but he must know something of the concept of which the thing is an instance." "One can value a thing only if one knows it, that is, if one knows its name and its properties."

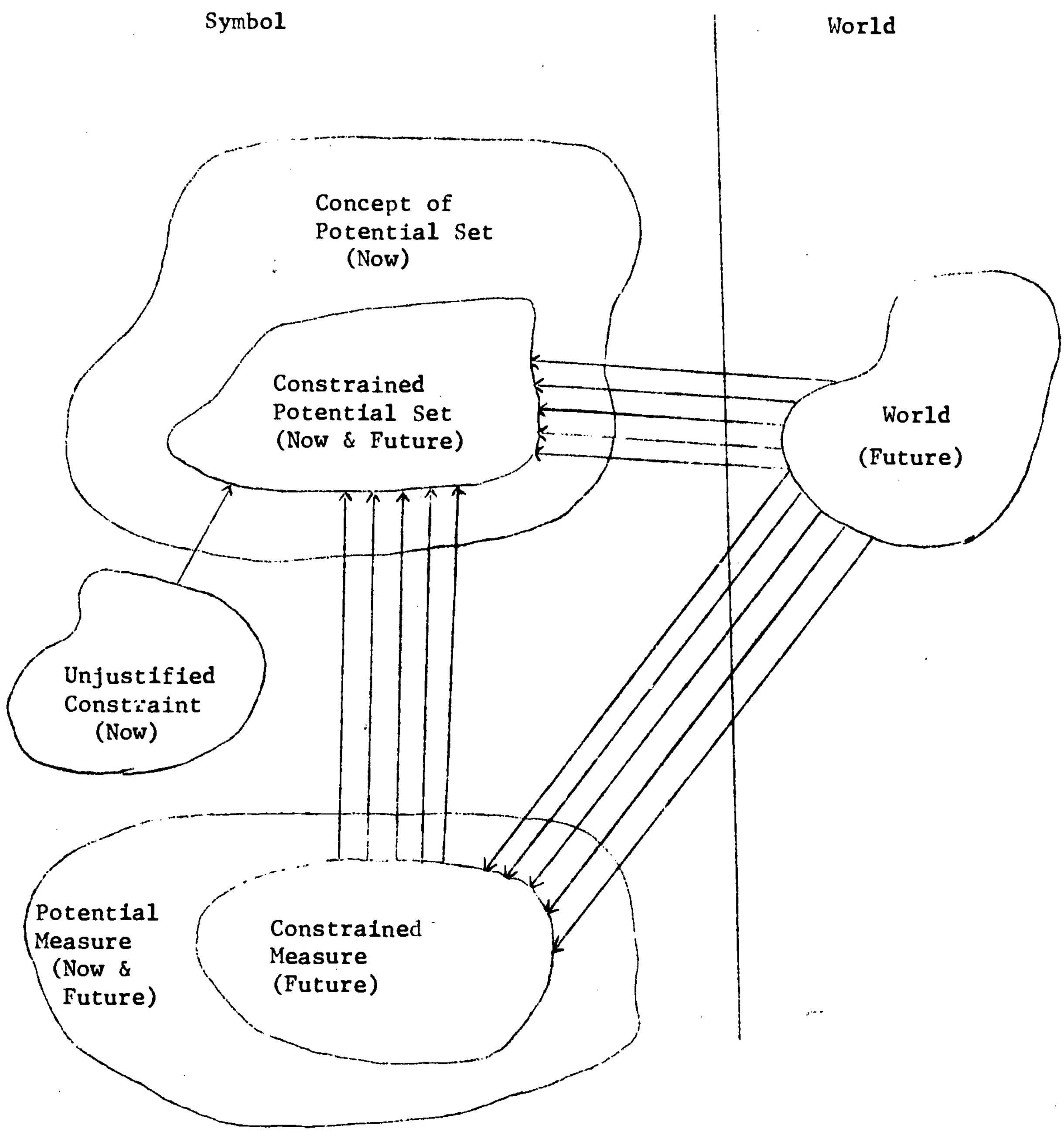
Value has a second meaning in common usage. If we have a variable with a set of elements determining the range of the variable, the element substituted for the variable in any particular situation is called the value of the variable. When we develop an objective measure M, we tie the elements of a set M to the world. Each element indicates a state of the world. The actual state of the world that occurs determines the value of the measure. No implication of goodness is intended in this use of value. Figure 4 can be expanded to include the measure. The lines connecting the circles in figure 5 shows the world reducing the potential measure to the constrained measure. For the measure the world acts as the unjustified constraint. In this case the constrained measure falls within the concept embodied by the constrained potential set of the interested social group. Note that their direct judgment of the world also falls within the constrained potential set. No action is required in this diagram; the diagram shows a successful evaluation.

Evaluation

Evaluation is at minimum concerned with making the three mappings indicated by the parallel lines in figure 5. Two of these mappings involve the concept of value as good, namely the matching of the world to the constrained potential set and the matching of the constrained measure to the constrained potential set. The third mapping uses the concept of value as a particular element from the range of a variable. The evaluator needs a constrained potential set, i.e. a concept of the activity being attempted which he can understand, in order to proceed.

Figure 5

Diagram of a Successful Evaluation



If the evaluator applies his knowledge of the constrained potential set directly to the world, he is using human judgment. Since there is widespread distrust of direct judgment in most concept holding groups, he will probably need to develop measures. If either the constrained potential set or the measures are missing, the evaluator cannot fulfil his role. The client, if he wants the evaluator's services, will surely ask the evaluator how to provide the constrained potential set and the required measures. Here the evaluator serves as an educational expert and as a measurement specialist.

Evaluation is at maximum concerned with the logical connections made in the conceptual structure of a developing or developed course and with making suggestions for the resolution of conceptual conflict produced by justification efforts. This role calls for the evaluator to act as a logician and arbitrater.

<u>The use of "outside-in" symbol systems</u>

What kind of symbol systems are we seeking? The evaluator needs abstract systems for economy, i.e. a small number of symbolic expressions should circumscribe much of the world's variety. The evaluator needs precision for clarity, i.e. the ability to state unambiguously what is intended with each symbol of the conceptual system. The symbolic expressions of physics are a good symbol system because they are general but still they are precise. On the other hand many expressions of everyday language are sufficiently general, but their lack of precision creates problems.

A more familiar example of an abstract but precise symbol system is offered in statistics and experimental design. They are based on mathematical expressions which are quite general; the formula for standard deviation, $s_d \equiv \Sigma(x-\bar{x})^2/n-1$, can be applied to a set of numbers generated in many different situations. Analysis of variance can be applied to research in psychology, education, agriculture, etc., with no change in the form of the analysis. The symbols involved are abstract, but the definitions of the symbols are precise. No problems of reference exist with the use of the symbol for standard deviation, s_d, once the symbol has been defined. In short the evaluator is a specialist in the application of "outside-in" symbol systems for abstractness and precision are obtained only in "outside-in" systems.

The role of the educational evaluator

<u>The evaluators role broadly defined.</u>

I turn now to the implications of the general evaluation model for the activity of the educational evaluator. The services of an evaluator are needed by clients from the teaching system when one or more of five problems must be solved. An evaluator is needed (a) when instructional systems must be justified to concerned others, (b) when the people developing the instructional system need assistance in identifying and agreeing upon the constraints they are going to meet, (c) when the match between the conceptual constraints adopted and the world is not readily determinable by the judgment of the concerned parties, (d) when a mismatch has occurred between conceptual constraints and the world, and action must be taken to eliminate the mismatch, and (e) when a measure is required and the parties concerned with the system being developed or tested do not know which measure to use.

The self-programmable evaluator.

The evaluator can treat himself like a self-programmable computer. A computer can carry out several operations but the order of those operations is left undetermined by the people controlling access to the computer. The user must specify with a program the sequence of operations needed to solve a particular problem. In the terminology developed in section one, the computer is all potential until the programmer supplies some constraints. The evaluator has a set of special skills which are appropriate in nearly all educational settings, but the use of those skills must be constrained by the particular setting to which they are applied. The evaluator must establish a contract with the people with whom he is working that clearly specifies the constraints under which he will operate. Three examples of this self-programmable approach follow.

Example 1. An instructor of an introductory computer programming course has developed a workbook which students can use during his course. He wants to know if the workbook is adding anything to the students performance on course examinations. In discussing the problem with the instructor, the evaluator determines that the following constraints are present: (a) the computer programming course is not in need of justification, (b) the instructor has clearly conceptualized and is satisfied with his definition of student learning, i.e. learning is satisfactorily measured by student performance on multiple choice questions and programming problems already prepared for the class, and finally (c) the instructor plans to publish the workbook and wants to know if it benefits his class and also if it will benefit introductory programming classes elsewhere. The evaluator programs himself to work within these constraints in accepting the contract with this professor. His activities will be limited to discovering the match between intended outcomes of the workbook and the actual outcomes of using the workbook, i.e. to determining whether or not the workbook increases examination performance. Should a mismatch occur between the conceptual constraints and the world, i.e. some parts of the workbook do not increase performance, he can offer advice concerning possible changes along with help in testing the effectiveness of those changes. The evaluator should not, in this instance, ask questions about whether teaching fortran programming to engineering students is of value. Doing so would violate the constraints of the contract established by agreement with the professor.

Example 2. A national agency wants to know if a widespread and expensive program designed to help children having reading problems overcome their reading deficiencies should be continued. A team of evaluators is hired and only one constraint is placed on their activities; they are to have the written report of their evaluation submitted by a particular date. In this case the evaluators will concern themselves with all five problem areas listed in the opening paragraph to this section.

Example 3. An evaluator seeking to develop a measure for use in student evaluations of instruction realizes that he needs a clear concept of what teaching should be before he can develop his measure. A sample of students, faculty, and alumni were asked for written descriptions of behaviors which contribute to effective teaching (Perry, 1969). Judges read the 13,643 behavior descriptions resulting and grouped them into 60 categories of behavior. A second sample of students, faculty, and alumni were presented with a list of statements in which each of the 60 category descriptors was included. The new sample indicated which of the following weights--critical, above average, average, below average, no importance--most characterized the importance of each statement. The value of each statement was determined from the weights. A rank order correlation of the ratings done by students and faculty showed a Spearman correlation of .91 for the 60 item questionnaire. This correlation indicates a high degree of agreement between students and faculty on the importance of the various items of the questionnaire to teaching. Perry recommends that the statements or a subset of the statements be used as a teacher rating form with student judgments of the teachers fulfillment of a statement being limited to always, most of the time, occasionally, very seldom, or never. He further argues that the students choices could be combined with the values of each item to produce a total score for the instructor. This is an example of a weighted utility decision model. In this example the evaluator established the contract with himself and constrained himself to sharpening the definition of the concept of effective instruction and to developing a measure for such a concept. He ignored the implications for action of the mismatch between concept and world that might come to light through the use of the student rating as a measure.

The evaluator must be self-programmable because the wide range of problems his skills are appropriate to cannot be encompassed by any one set of constraints. Conceptual constratints are held by clients and until we know who we are working with we cannot know what constraints are present. The constraints that are satisfactory to professor X in the evaluation of a single college course are not going to be acceptable in a full scale national program evaluation involving persons W, Y, and Z. **The nature** of evaluation as discussed in the general model remains the same but the specific constraints applied will differ. The determination of these constraints must be done by agreement between the evaluator and the specific people he's helping. If the evaluator does not want to accept the constraints of the people employing him or if they don't wish to accept the constraints the evaluator seeks to apply, then no contract is established between them.

The Evaluator's Training

Since the evaluator has been described as a person who programs himself to apply his skills to a particular evaluation problem, we must discuss the nature of those skills. The evaluator clearly needs knowledge of the educational ------ and the educational philosophy literature. He will often be questioned by his clients concerning effective methods of instruction and he should be able to give a research based answer to the question if such an answer exists. The client will often need help in clarifying his concept of what he is attempting to do with his instruction efforts and the knowledge needed to provide such help comes from the education literature. We would also expect the evaluator to be familiar with reports of evaluations done in other settings as methods from old settings may be appropriate to new ones. Finally the evaluator should be familiar with the philosophy of science.

Beyond the specific knowledges just mentioned, the evaluator should have introductory but firm knowledge of several abstract but precise symbol systems. For example, control theory is applicable whenever an attempt is made to in some way design and construct the world. This is particularly true when a test, change, test, change, test strategy of design is used. W. Ross Ashby (1956) has written a precise, highly readable introduction to control theory. He does not go into the advanced mathematics which engineering textbooks use because the engineering applications are mainly special cases of his general introduction. Solid, useable knowledge of this one book and familiarity with one more good text on control theory, e.g. Stafford Beer's Decision and Control (1966), would be all the evaluator would need. The same case can be made for statistics as an evaluator's tool. Thorough knowledge of Hays' Statistics for Psychologitst (1963) would provide the evaluator with the statistical tools he needs in his work. Winer's (1962) text on experimental design would serve as a source for that topic. In short, the evaluator should have introductory knowledge of several symbol systems gained from reading one or two outstanding introductions to each system, and he should know those introductions thoroughly.

The following list suggests some symbol systems the evaluator might wish to add to his repetoire:

1. Statistics
2. Experimental design
3. Measure theory
4. Educational and psychological test and measure theory
5. Decision theory
6. General systems theory
7. Control theory
8. Hartman's value calculus
9. Computer programming
10. Finite mathematics
11. Symbolic logic
12. Boolean algebra
13. Mathematical analysis thru calculus
14. Linear programming
15. Network analysis
16. PERT

The task of learning these systems is made easier by the interrelationships between them. The first seven systems listed, for example, are all special methods of handling the concept of potential. As such, they are direct extensions of the general evaluation model proposed in the philosophy section of this paper. Hartman's value calculus is a symbolic logic.

If we expected evaluators to be expert in all of these symbol systems, we would be asking for an impossible supercompetence. Instead we ask for a thorough, introductory knowledge of several of them. This knowledge is well within the capabilities of those people who should consider being educational evaluators. If we expected the evaluator to use the languages learned from these symbol systems in direct communication with the client, we would open ourselves to charges of deliberately confusing the client. Instead the evaluator is to use these symbol systems to clearly see what he is attempting in any evaluation he conducts. He would have to translate to everyday language in discussing the results of evaluation with his client. When the client comes to the evaluator for help with his problems, he should expect the evaluator to have the ability to clarify the instructional problems posed with his knowledge of logical systems, the ability to do the evaluation suggested by his tools, and the ability to translate the clarifications and the evaluation results into everyday language. If the evaluator can meet these expectations, at least partially, he will have a satisfied and expanding clientele.

References

Ashby, W. E. An Introduction to Cybernetics. London: Chapman Hall, 1956.

Beer, S. Decision and Control. New York: John Wiley, 1966.

Beer, S. Brain of the Firm. New York: Herder and Herder, 1972.

Fromm, E. The Sane Society. New York: Fawcett World Library, 1955.

Hartman, R. S. The science of value. In Abraham Maslow (Ed.), New Knowledge in Human Values, Chicago: Henry Regnery Co. 1959. pp. 13 - 37.

Hartman, R. S. The Structure of Value, Carbondale: Southern Illinois University Press, 1967.

Hays, W. L. Statistics for Psychologists. New York: Holt, Rinehart, & Winston, 1963.

Kaplan, A. The Conduct of Inquiry. Scranton, Pa.: Chandler Publishing, 1964.

Kasamatsu, A. & Hiroi, T. An electroencephalographic study of the Zen meditation (zazen). In Charles Tart (Ed.) Altered States of Consciousness. New York: John Wiley, 1969.

Nagel, E. & Newman, J. R. Godel's Proof. New York University Press, 1958.

Perry, R. R. Evaluation of teaching behavior seeks to measure effectiveness. College and University Business, 1969, 47 (4), 18 & 22.

Pribram, K. H. Languages of the Brain: Experimental Paradoxes and Principles in Neuropsychology. Englewood Cliffs, N. J.: Prentice-Hall, 1971.

Watts, A. The Book: On the Taboo Against Knowing Who You Are. New York: Collier Books, 1966.