DOCUMENT RESUME

ED 259 011 TM 850 380

AUTHOR Dillon, J. T.
TITLE Finding the Question for Evaluation Research. Paper and Report Series N., 102.
INSTITUTION Northwest Regional Educational Lab., Portland, OR. Research on Evaluation Program.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Oct 84
CONTRACT NIE-400-80-0105
NOTE 36p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Classification; *Evaluation Methods; Evaluative Thinking; Program Evaluation; *Research Design; Research Methodology; Research Needs
IDENTIFIERS *Evaluation Research; Northwest Regional Educational Laboratory; *Questions; Question Types

ABSTRACT

Evaluation research is conceived as a question answering process. Four strategies are suggested which emerged from four approaches to the study of questions. Evaluation questions must be identified and classified before research can effectively begin. This paper discusses the four approaches to studying evaluation research questions and their implications for practice. The presentation is organized under the headings: Identifying the Question; Posing the Question; Answering the Question; and Reporting the Answer. The importance of selecting the right question for evaluation research is emphasized. Systematic methods for locating appropriate questions for evaluation research are identified. These include: (1) classifying the questions that can be asked before identifying the question; (2) analyzing the question before posing it; (3) making a dummy of the answer before addressing the question; and (4) stating the question again before stating the answer. (DWH)

paper and report series

No. 102 FINDING THE QUESTION FOR EVALUATION RESEARCH

J. T. Dillon

Research on Evaluation Program

Northwest Regional Educational Laboratory
300 S.W. Sixth Avenue
Portland, Oregon 97204
Telephone (503) 248-6800

Interim Draft

Do not cite or quote without author's permission. Comments and suggestions are welcomed.

No. 102 FINDING THE QUESTION FOR EVALUATION RESEARCH

J. T. Dillon

University of California, Riverside

October 1984

Nick L. Smith, Director
Research on Evaluation Program
Northwest Regional Educational Laboratory
300 S.W. Sixth Avenue, Portland, Oregon 97204

Published by the Northwest Regional Educational Laboratory, a private nonprofit corporation. The work upon which this publication is based was performed pursuant to Contract No. 400-80-0105 of the National Institute of Education. It does not, however, necessarily reflect the views of that agency.

PREFACE

The Research on Evaluation Program is a Northwest Regional Educational Laboratory project of research, development, testing, and training designed to create new evaluation methodologies for use in education. This document is one of a series of papers and reports produced by program staff, visiting scholars, adjunct scholars, and project collaborators--all members of a cooperative network of colleagues working on the development of new methodologies.

Evaluation research can be conceived as a question-answering process, and its resulting knowledge as a question-answer proposition. The theoretical study of questions suggests four practical strategies for undertaking an evaluation study. (1) Before identifying the question to investigate, classify the questions that can be asked. (2) Before posing the question, analyze it. (3) Before addressing the question, construct a dummy answer. (4) Before stating the answer, state the questions. In general, it makes pragmatic sense to expend at least as much effort on finding the question as on finding the answer.

Nick L. Smith, Editor
Paper and Report Series

CONTENTS

	Page
Identifying the Question	1
Posing the Question	11
Answering the Question	17
Reporting the Answer	25
Conclusion	27
References	28

FINDING THE QUESTION FOR EVALUATION RESEARCH

Evaluation research can be conceived as a question-answering process, and the knowledge that results as a question-answer proposition. This conception will be left in abrupt summary form in order to pass rapidly on to its implications for practice. What can be done with it? How might it enhance the practice of evaluation research, in respect of asking and answering evaluation questions?

Four broad strategies can be suggested, emerging from four approaches to the study of questions: categorial theory, erotetic logic, pragmatics, and epistemology. Phrased as didactic suggestions to a docile evaluation researcher--as pieces of advice imploringly requested--they run as follows.

1. Before identifying the question to investigate, classify the questions that can be asked.

2. Before posing the question, analyze it.

3. Before addressing the question, construct a dummy answer.

4. Before stating the answer, state the question.

These will be set forth under headings of identifying the question, posing the question, answering the question, and reporting the answer.

Identifying the Question

Before identifying the question to investigate, classify the questions that can be asked.

Classification is one useful way to set about finding the question for research. There are other ways to find a question--curiosity, luck, inspiration, genius--but these are not ways to set about finding it. There are still other ways that questions are identified but neither are they ways of identifying one--copying a question from the literature, being assigned one, answering someone else's question as posed.

Two relevant concerns that classification satisfies are: What might be some novel and original (possibly fruitful and significant) questions to investigate? What is the range of questions that ought to be considered for study (so that important, obvious, or necessary ones will not have been overlooked)? These concerns might be said to relate to the field. A third concern relates to the investigator: What personal conceptual involvement does the researcher have in the enterprise?

Classification is a conceptual act. It consists in conceiving of the kinds of question that can be asked about the domain or phenomenon being investigated (evaluated). The result is a set of concepts, or classes, each constituted by a generic question--of which the

question/s formulated for evaluation will represent particular, individual variants. The generic questions thus establish the categories of thought about the phenomenon to be investigated. To classify the questions is to construct a conceptual framework that defines thought and action in that domain. (On the classification of questions see Dillon, 1984a; on the relation between questions and categories see Kahn, 1978 and Rescher, 1982.)

In a categorial approach to classifying questions--the approach I recommend (Dillon, 1984a)--all of the questions will be the simplest ones conceivable, each addressing one of the most fundamental aspects of the phenomenon imaginable. No more than one question per aspect will be identified, and no aspect left unidentified. The result is a set of all possible (kinds of) questions askable about all possible (kinds of) existential aspects of the phenomenon. The classes and subclasses--distinguished to whatever degree of detail is thought useful--are arranged to display the relations that are conceived to hold among the categories of the phenomenon.

One such a scheme appears in Table 1 (from Dillon, 1984a). We should be mindful that it is not an example of what to do, but an instance of the kind of thing that can be done. Each classification is different in the principle that it uses, in the classes it identifies, in the domain that it describes, in the detail and scope of its description, and in the purposes for which it is constructed. Moreover, there are in fact very few classification schemes available for questions. It is for these reasons that Table 1, although a good classification, is not what is ordinarily thought of as an example; and it is also for these reasons that I recommend that each evaluator construct a classification for the particular domain and purpose of evaluation.

With such a scheme as in Table 1 it becomes obvious what possible questions can be asked, about which possible aspects, of the phenomenon under investigation; and, further, which questions are being emphasized and neglected in previous investigations (e.g., review of research) or in the complex of current knowledge (e.g., the evaluation sponsor's or client's) about the phenomenon.

Moreover, it becomes easy to locate within the scheme the question/s of interest to investigate, and to determine (a) whether these questions can be asked, or others must be asked first; and (b) in what form the questions are to be asked.

What is more, the classification scheme reveals in what relation any question of interest stands to other questions--both questions of that kind and questions of other kinds.

Finally, the scheme defines the set of possible answers to that question, anticipates the form of the eventual answer, and establishes a way to construe it--(a) as part of a proposition formed by the question-answer pair, and (b) in relation to answers of that kind and

other kinds. The scheme thereby provides a way to assess the significance of the evaluation undertaking.

Illustrations

It would give relief to these points and their readers if some illustrations would now be given. Unfortunately, little relief can be forthcoming. Remarkably few schemes are available, all of them deficient.

As for categorial schemes, the two best ones known to me are those constructed by Rescher (1982) and myself (Dillon, 1984a and Table I infra); the latter has also been applied to analyzing the questions posed in the educational research literature (Dillon, 1983; 1984b). These schemes can be used most fruitfully in conjunction with a topical scheme for a given domain, such as evaluation, so as to identify various kinds of questions that can be asked about various kinds of thing constituting the subject under investigation (e.g., the program being evaluated).

To illustrate, suppose the program to be evaluated is the program called Program Evaluation in a university. What are the kinds of thing that constitute this subject for evaluation?

An abbreviated classification might alliteratively run: Purposes (Why?), Persons (Who?), Places (Where?), and Processes (What? How?). Each of these categories can be elaborated to a degree thought useful. First one asks about omitted categories. One does not ask if anything has been left out, but if any kind of thing has been left out. Thousands of things are left out; can we conceive of things that do not fit in? and, do these things require us to conceive of another category, or of four-five new categories to replace these four so far? Ordinarily (for reasons we needn't concern ourselves with here), one cannot just add another category to a list once conceived; one reconceives and identifies new categories.

"Cost," for example, is a popular item in evaluation, but need not constitute a new category: it can be conceived either as an element of all four categories (the cost of attaining this purpose, the cost of these faculty, etc.), or as part of the fourth category, Processes. Of course, depending on purpose, an entire classification of this program could be based upon the characteristic of cost alone.

Another classification might re-arrange topical categories in order to reflect curricular aspects of the evaluation program: Purposes, Teachers (& Teaching), Students (& Learning), Subject-matter (& Materials), Milieux, Implementation, Evaluation, and Profession. "Outcomes" is a popular topic in evaluation research, but need not be a category: it can be conceived as an element attaching to all eight main categories, or part of the one category, Students (& Learning). This scheme, here given in outline only, has been used to classify all

Table 1

A Classification of Research Questions

Category of Question	Knowledge in Question-Answer
ZERO ORDER	NONE
0. Rhetorical	no knowledge or no answer.
FIRST ORDER: PROPERTIES	INDIVIDUAL ATTRIBUTES OF P, OF Q
1. Existence/Affirmation-Negation	whether P is.
2. Instance/Identification	whether this is a/the P.
3. Substance/Definition	what P is.
a. Nature	- what makes P be P.
b. Label	- whether "P" names P.
c. Meaning	- what P or "P" means.
4. Character/Description	what P has.
5. Function/Application	what P does.
a. Modes	- how P acts.
b. Uses	- what P can do.
c. Means	- how P does it or is done.
6. Rationale/Explication	why or how P has a certain attribute.

Table 1 (Cont.)

SECOND ORDER: COMPARISONS	COMPARATIVE ATTRIBUTES OF P-AND-Q
7. Concomitance	whether P goes with Q.
a. Conjunction	- whether P and Q are associates.
b. Disjunction	- whether $\overline{P}$ and $\overline{Q}$ are alternatives.
8. Equivalence	whether P is like Q, and wherein.
9. Difference	wherein P and Q differ.
a. Disproportion	- whether P is more/less than Q.
b. Subordination	- whether $\overline{P}$ is part/whole of Q.
THIRD ORDER: CONTINGENCIES	CONTINGENT ATTRIBUTES OF P-AND-Q
10. Relation	whether P relates to Q.
11. Correlation	whether P and Q covary.
12. Conditionality	whether or how if P then Q, or if Q then P
a. Consequence	- whether if P then Q, or what X if P.
b. Antecedence	- whether if Q then P, or what $\overline{X}$ then P.
13. Biconditionality	whether or how if P then Q and if Q then P.
EXTRA ORDER: OTHER	OTHER ATTRIBUTES OR WAYS OF KNOWING P.
(14.) Deliberation	whether to do and think P.
(15.) Unspecified	to know P in other ways.
(16.) Unclear	not known.

the topics of all articles found in curriculum journals during one year (Dillon, 1984b).

Now one conjoins the topical scheme with the categorial scheme in Table 1, articulating them to classify the kinds of questions for evaluation: Which aspect of which topic to evaluate?

We can appreciate that any and all of the Table 1 categories can be applied to any one or to all of the topical categories, both singly and in combination. But I do not recommend making a grid and spinning out all the hundreds of possible combinations. Rather, I have recommended classifying the kinds of questions that can be asked. Then, with this conception of the domain--the department called Program Evaluation--one can identify the kind of question that is of most interest, and lastly formulate the particular questions of that kind.

As an alternative, where the question is already given by a client or sponsor--or Faculty Senate or State Legislature--I recommend construing that question by situating it within the appropriate categories of topic plus knowledge about that topic. (In the end, the department will be cut anyway, but at least one will learn the kind of question that had the cutting edge, and the kind of answer that enjoyed the decisive edge.)

Approximations

The evaluation literature contains some approximate examples of classification. Three are very instructive.

Making and collating lists of questions--about everything, from everybody--seems to be an emergent approach to "generating" questions for evaluation. The multiple questions thus generated are then "prioritized" for study. One praiseworthy example are the lists offered by Smith (1982) as part of his general effort (e.g., Smith 1981) to broaden the range and variety of questions (and methods) used in evaluation research, "The heart of an evaluation design is the questions asked" (1982, p. 112).

As ways of "generating" questions about a teacher education program, Smith (1982) proposes that the evaluator consider five things: the purpose, audience, and focus of the study, the criteria of the program, and general sources of information about the program. He lists 10 possible purposes, 23 audiences, 16 foci, 9 criteria, and 10 sources of information. For each he gives as "examples" some particular questions that are "suggested" by the item. The many questions thus generated are then prioritized according to pressing program needs, external mandates for information, and audience information needs; final priority is assigned by the purpose and the focus of the evaluation study.

A listing-generating approach can be useful but far less so than

classification, which yields greater conceptual range and greater practical control of the questions for evaluation.

First, lists specify multiple individuals (almost certainly of one or two kinds), whereas classification identifies multiple kinds, for each of which multiple individuals may be formulated. What looks like an impressive variety of questions on the list can actually be a monotonous multiplicity of individual variants, all of a kind. For range and variety, one does better to stop thinking up examples, individuals, particulars and start conceiving of kinds, classes, and categories.

Second, items on lists cannot of themselves generate questions. Consider "program goals": What is the question thus generated? "Efficiency": What question does that suggest? Any question whatsoever--an infinite number--can be formulated to include the word listed; the questions that are listed as examples must derive from some other principle or operation. They are appended to the list, not generated by it. By contrast, in the classification approach recommended here, each category is itself constituted of a generic question, of which the particular questions are individual and interrelated variants.

If questions must nonetheless be generated by listings, one estimates that a computer would do so better than a poor human evaluator. And, indeed, computers are being used to do this very thing. Mechanizing Hypothesis Formation (Hajek & Havranek, 1978) reports how a program called GUHA (General Unary Hypotheses Automaton) automatically lists all observational statements possible from a set of data. Applied to industry, for instance, data on simultaneous overflashing of the generator and motor of diesel-electric locomotives permitted GUHA to list 197 "important" statements that then "serve as a source of hypotheses for further investigations" (p. 380). But what does the evaluator do with the 197 questions thus generated, save return them to a computer? From start to finish, classification would give greater conceptual range and practical control of the questions.

A second approximate example also comes from Smith (1980). Here he conceives that evaluators involve themselves in four kinds of events, each entailing a kind of question which in turn yields a corresponding kind of answer (via a kind of method).

1. Policy. What is the desireable state of affairs? The answer is a value claim: X's are good.

2. Research. What is the state of affairs, and why is it that way? The answer is a fact claim: X is true; X because Y.

3. Management. What needs to be done to achieve the desireable state? The answer is an action claim: Do X here.

4. Evaluation. Is this specific event desireable? The answer is a value attribution: This X is good.

Whatever the merits of this scheme as a classification, it illustrates two advantages of proceeding by categories rather than by individuals. (Interesting to note, this scheme follows upon Smith's effort to classify the evaluation literature.)

First, it makes immediately apparent that there are distinct classes of questions that are different--in kind--from most questions used for evaluation. And it gives the generic questions of these other kinds, so as to provoke any number of individual questions of that kind. It conceives of four kinds of question instead of listing 400 individuals.

Second, the scheme shows how the answer depends on the question: both are of one and the same kind. Thus, if you formulate even 100 questions of a kind, you will get answers all of a kind, e.g., X is true. If you wish to know something of another kind, you must formulate not just another question, and not just a different question, but a question of a different kind. That is another way of saying that questions and answers proceed within categories of thought. Only when you first conceive of the categories do you know the kinds of knowledge that your questions will yield; and you know further the other kinds of knowledge that could be yielded by asking questions of another kind.

A third approximate example is David Nevo's (1983) list of ten dimensions of the domain of evaluation research, each formulated as a question. For example, Definition--How is evaluation defined?; Functions--What are the functions of evaluation? The scheme could be articulated to incorporate missing categories and the questions reformulated to apply to a particular program being evaluated. Although impaired as a classification--some dimensions are topical (e.g., Objects), some are epistemological (e.g., Definition), and others are phenomenological or existential (Functions)--the scheme proves a useful example of organizing knowledge of a domain in light of the kinds of things known via answers to questions of those kinds.

Nevo presents a conceptualization of evaluation by reviewing the literature for the answers it gives to the ten questions. Note that the conception had already been constructed by the act of classifying the (kinds of) questions; then the individual studies were reviewed by the categories thus established.

So, too, an evaluator can organize his/her understanding of a program being evaluated--can conceptualize the program--by first classifying the kinds of topical questions that can be asked about it. That yields the kinds of topical knowledge that can be had about the program: knowledge of topic X, topic Y, topic Z.

Then a second, categorial classification can be used together with the topical scheme, specifying the kinds of questions that can be asked (and thus the categories of knowledge that can be had) about each one of the aspects of the program: Questions A+B+C about aspect X, ABC about aspect Y, and Z. For example, questions about the

nature, and the functions, and the effects of the teaching in the program, and of the program's goals, and of the program's staff. That yields knowledge of various kinds about things of various kinds that constitute the program being evaluated.

In the end, the merit of a classification scheme rests on a pragmatic criterion--briefly, how useful it proves to be for the purposes intended. These purposes are defined by the evaluator (and sponsor, etc.) in the particular evaluation circumstance. Therefore the suggestion here is not to study classifications or to search out some good scheme somewhere. Rather, the didactic suggestion runs: Classify the questions. The evaluator is the one who must do that for the domain being evaluated.

If that is done, novel and original (kinds of) questions are likely to appear, as well as the important obvious ones not to be overlooked. Furthermore, the evaluator will be conceptually involved in the evaluation task, establishing the framework for thought and action, construing the question-answer and determining the significance of the evaluation study. Classification is a conceptual act, and the actor is the evaluator.

Other approaches

By contrast to the conceptual approach of classification, what might be called a "practical and political" way of identifying questions is suggested by Cronbach (1982). In planning the study, the evaluator first entertains the widest possible range of questions, and then selects among them. Wide range is assured by consulting not only the policy-shapers and program sponsors but also social critics and scholars, members of minority and elite communities, professionals operating the program and citizens served by it.

It is not enough to derive questions from program goals, since these do not provide for unwanted outcomes (p. 221). The evaluator anticipates the possible outcomes that "partisans hope for and skeptics fear" (p. 210)--as well as the reverse, we might add; and then anticipates the competing interpretations for each. "These rival hypotheses suggest research questions" (p. 219). The selection among questions is then determined by assessing four characteristics: the degree of uncertainty about the answer; the promise of reducing that uncertainty; the comparative cost of the inquiry; and the leverage that the answer would have on policy or operations (pp. 225-226).

A number of other approaches are suggested for researchers in general as they formulate the question for their individual study. Campbell, Daft and Hulin (1982) provide, hesitantly, a detailed list of "proactive behaviors that might enhance problem finding and question generation" (p. 116). One strategy that might apply to evaluation research is to "break established mindsets" through various group and individual activities such as brainstorming. Another is to state the converse of the conventional wisdom on the subject, assume the converse to be true, and see what research questions follow. Two

general activities are to take whatever question emerges and (a) discuss it with colleagues and (b) reformulate it in three alternative ways, examining the pros and cons of each.

Still other, and slighter, pieces of advice are listed by McGuire (1973) and Webb (1961), and by the authors of any methods text who list the usual do's-and-don'ts (do ask a significant question; don't ask an unclear question...).

It should be evident that classification serves in all these cases as well, as it does also in cases where the evaluator is not free to identify the question but is presented with one to answer or is to negotiate the questions with sponsors. From whatever source, the various questions are fitted within and construed by the classification scheme that the evaluator has taken care to produce beforehand--in the determined effort to conceive of the enterprise and to maintain conceptual control over it.

Construing the given question within this scheme permits most of the same advantages to accrue to this question as accrue to the ones that the evaluator sole might have identified thereby. And it yields the further advantage of clarifying for all parties the sense of both question and answer. For instance, it prevents a categorial switch, where an answer is construed in a category other than that of the question. The classification scheme can be used to see to it that answers of a kind are given for questions of a kind. Then there can be no question as to the meaning of the results of the evaluation study. (The results may nonetheless be rejected, meaningful and true though they be; that is a matter of question-answer pragmatics, to be addressed by another suggestion.)

What strikes the uninformed outsider is the rhetoric about questions in the evaluation literature. It hails the role of questions; it rues the neglect of the question; it prescribes attention to the question; and, having in these one or two sentences writ QUESTION thus large, it moves on to say no more about it. Classification supplies a way to do what the rhetoric exhorts be done.

Nick Smith observes: "One present difficulty in evaluation which is seldom noted is that evaluators are often unclear about the nature of their questions" (1981, p. 61). --Classification is a way to become clear about the nature of questions that can be asked and that are being asked.

Henry Levin observes: "One of the most neglected areas of evaluation generally is that of proper identification of the problem" (1983, p. 34). --Classification is a way to identify the problem properly.

Levin goes on: "Before one begins to address the problem, one must attempt to specify with great clarity and insight the nature of

the problem that ought to be addressed" (p. 35). --Classification is a way to specify the nature of the problem with clarity and insight.

Paul Wortman observes: "The first step in the conduct of a cost analysis involves the specification of the problem.... Evaluation research, however, has been program, rather than problem, focused" (1983, p. 248). --Classification is a way to give specification to the problem and problem-focus to evaluation.

Samuel Ball observes that Peril No.6 of program evaluation is Having to Answer the Wrong Questions. "Evaluators are expected to provide slick, easy answers to slick easy questions" (1982, p. 171). --Classification is a way to pin the question down and make the answer stick to it.

Classification is a useful way to go about finding and formulating the question for evaluation research. Naturally, no one need feel called upon to do it that way.

Posing the Question

Before posing the question, analyze it.

Once the question is posed it takes on a dynamic and force of its own in the thought and action of the investigator. Yet the question may be impaired or unfit; it must first be examined and vetted for inquiry. Then it can be posed in the dress that analysis reveals as most properly suiting the investigation.

The analysis of questions is far easier than their classification. Only two things need to be looked at. These may be described informally as (sets of) sentences that come before the question, and sentences that come after. Those that come after are answers, which are the object of the third suggestion discussed in the subsequent section.

As for sentences that come before the question, they are of two types: the presuppositions of the question, and the presumptions of the questioner.

Presuppositions

Presuppositions are sentences (expressing propositions that are) entailed by the question-sentence. They are sentences that must be true for that the question can have a true answer. If these sentences are not true, the question is invalid. It cannot be validly posed because it cannot be truly answered. Logicians say in that case that the question does not arise. (For precise formulations of theories of question-presuppositions, see Belnap, 1969 and Keenan & Hull, 1973.)

Here are two favorite examples cited by linguists and logicians. "Have you stopped beating your wife?" presupposes:

a. you have a wife;

b. you used to beat her;

c. either you have stopped beating her or you are still beating her.

"Is the King of France bald?" presupposes:

a. there is a present King of France;

b. the King is eitner bald or not-bald.

Naturally, anyone can go on to ask such a question anyway and then go on to get an answer. The answer will not be a true one and no one will be the wiser. A certain amount of published research reports untrue answers to invalid questions, and no one is the wiser--wiser neither for knowing the answer nor for not-knowing that the answer is untrue.

The analytic questions to ask here are simple ones:

1. What does this question presuppose?

2. Is that which it presupposes known to be true?

If the presuppositions are not known to be true--they are false or indeterminate (not known to be true or false)--then the research question cannot be asked. Another question must be chosen for evaluation.

Particularly useful here is a classification scheme that arranges its questions in a hierarchy or sequence, for it displays those prior questions whose answers serve in turn as presuppositions to the present question of interest. If those prior questions have not been answered, then the subsequent questions cannot be asked; for, the truth of the prior question-answer proposition provides just the condition for asking the next question--to say nothing of answering it. (The scheme also exhibits the subsequent questions that may be asked in a programmatic evaluation project.)

Naturally, any researcher can go on to ask the more significant and high-level questions without troubling self with the trivial, low-level ones. Indeed, such is the strategy of preference for becoming a researcher of repute. Thus it transpires that a certain amount of published research addresses the subsequent questions to which the prior ones remain unanswered and necessarily serve as indeterminate and possibly false presuppositions to an invalid question for which an untrue answer has been proposed as the conclusion from significant research. But that is of no concern to anyone because no one knows that. The problem rather is precisely one of not-knowing that the answer is untrue, howsoever significant. That remains not-known because the very posing of the question had presumed

the truth of the presuppositions--that is, those sentences whose truth is a condition for there to be a true answer to the question.

The rush towards the frontiers of knowledge is slowed only when intelligence begins to filter in from the cognitively depressed hinterland, where toil the faceless researchers over their lowly questions, left behind by the path-breakers in their ascent to the cognitively lofty pursuits. These minor studies from odd, neglected lines of research can prove to invalidate the entire series of conclusions by force of demonstrating the falsity of the presupposition to prior questions in the series. A far more sophisticated and easy way is to examine the truth of the presuppositions before posing the question.

Presumptions

Presumptions are those sentences (expressing attitudes that are) implied by the questioning-act. They are sentences that must be true (accurate of the attitudes) for that the question can be genuinely posed. If they are not true, the question is fraudulent. It cannot be genuinely asked because the questioning-attitude is absent. That means at the least that either (a) that question is not being asked, but some other; or (b) no question at all is being asked, but some other act or performance or exchange is being engaged in.

Naturally, a good many questions in everyday life and in specialized pursuits, not alone in the research enterprise, are fraudulent and so are fraudulently answered. Here too no one is the wiser for it. But the evaluator has to know the score, what's going to count, what coinage the questioner will accept--counterfeit or not.

The first (primary) presumption is that the questioner believes that the presupposition/s to the question is true. It may be false, but no matter: Asking the question presumes its truth, i.e., commits the speaker to it as true, or expresses the speaker's belief that it is true. The analytic question here is, Does the questioner (Q) believe that the presuppositions to the question are true?

Other presumptions follow and may be variously enumerated and formulated. They include the following.

1. Q believes that there is a (true) answer.

2. Q does not know the answer.

3. Q needs/wants/desires to know the answer.

4. Q believes that the respondent can supply the answer.

5. Q is willing to do what it takes to find the answer.

(Note that in answering the question, the respondent also shares in the presumptions: Respondent believes that Q believes..., etc.)

Another formulation, together with examples of questions that violate the conditions, is provided by Knight (1967, p. 571):

1. an unknown based on fact. "When you did stop beating your wife?"; "Who is the Tsar of Russia?"

2. a desire to know the unknown. Questions posed merely to disturb violate this condition.

3. faith that such knowledge exists and can be acquired. "Where are the snows of yesteryear?"

4. courage to accept the consequences of attempting to know. "What is the purpose of life?" or other questions where certain answers would be rejected a priori.

Any question that violates one of these conditions is "noninquisitive or meaningless for purposes of inquiry" (p. 571).

Logicians and philosophers have supplied a few other names for questions that fail on one or another of these grounds. Nuel Belnap (1969; Belnap & Steel, 1976) identifies these types of questions:

1. irrelevant - no answer will have any logical relation to beliefs of the questioner/audience. As in the adage, "What does that have to do with the price of tea in China?"

2. trivial - the answer is already known (to the questioner). "Does 2+2 equal 4?"; "Are these words written in English?"

3. foolish - the question is known to have no true answer. "Which of the following exist: unicorns or chimeras?"; "Are these words printed in red or in green ink?"

4. dumb - the question has no direct answers whatsoever. "What are at least three truths from among the following: A and B?"

From Keenan and Hull (1973):

1. vacuous - of zero truth-value; the presupposition is neither true nor false. "Did the students who failed get drunk?" (when all students passed).

2. pathological - no answer is false. "Which man likes the girl he likes?"; "Which students came early and didn't come early?"

From Rescher (1982):

1. trivial - the presuppositions of the question afford an answer to it. "Does this question have an answer?"

2. premature - the truth-status of the presuppositions is unknown or indeterminate (but none is known to be false). "What is the reason or the cause for the existence of the world?"; "What are

the learning processes of extra-terrestrial inhabitants of our galaxy?"

3. inappropriate - every answer is false, since some presupposition is false. "How long is a novel?"; "Why is the moon made of green cheese?"

4. absurd - there is no answer at all, every answer is self-inconsistent. "Why is that tree inorganic?"; "Is 'No' the correct answer to this question?"

By contrast, a legitimate or proper question is "one whose presuppositions are all (known to be) true" (pp. 136-137).

Approaches to analysis

As simply though this narrative has rendered it, the task of analyzing the question may seem all too fancy and difficult. Indeed, theorists of questions do couch their treatments in occasionally remote and forbidding language. Nonetneless, these are the people who know most and best about questions. Most of the "practical" advice available comes from people who have not thought much at all about questions, and their advice is practically worthless.

There are, however, a few sources that make some of the same points, but in more familiar language, as do the erotetic logicians and linguists. For those who would seek help in analyzing their question for research, here are some readable yet reliable sources (still not as informed or sophisticated as the theorists cited).

A non-technical, old-fashioned kind of logical analysis is suggested by the historiographer David Fischer (1970) in an excellent chapter, "fallacies of question-framing." He reviews ten logical fallacies that a research question might exhibit, and he richly and humorously illustrates each by analyzing the questions in published historical studies. For example, to illustrate false dichotomous questons, Fischer cites a dozen published titles including: "Martin Luther--Reformer or Revolutionary?"; "The Robber Barons--Pirates or Pioneers?"; "The Removal of the Cherokee Nation--Manifest Destiny or National Dishonor?"

Other types include the fallacy of many questions ("When did racial segregation harden into its elabcrate mold?"); metaphysical questions ("Was the Civil War inevitable?"); and semantical questions ("Was the political structure of 17th-century America democratic or aristocratic?").

"Without questions of the right sort, the historian's empirical projects are consigned to failure before they are fairly begun" (pp. 3-4). With Fischer's list in hand, the evaluator can quickly spot the fallacies in the proposed question before he/she goes on laboriously and unwittingly to produce, as Fischer's hapless historians did, a fallacious answer. It is quite common to pose a fallacious question, e.g., a falsely dichotomous question.

Another old-fashioned kind of analysis is pleasantly set forth by the philosopher Emmet (1968) in a chapter on "asking the right questions." He reviews various ways in which questions can be wrong.

1. merely verbal issues. "Can a computer think?"; "Are kind people generous?"; "Is Economics a science?"; "Are we ever really free?"

2. suggesting the answer. "Do you approve of the proposed betrayal of our country?"; "Are you in favor of this doctrinaire egalitarianism?"

3. built-in wrong assumptions. "What is it that determines the real worth or value of an object?"; "How do mind and body int ct?"

4. built-in unanswerability. "Is the statement 'This statement is false' true or false?"; "What is Reality really like?"

"What we must try to do is to think what it is that the questioner wants to know: if we can discover this, even though perhaps only vaguely, it may be possible by a rephrasing of the question to lay bare or to remove the implicit assumptions and make it intelligible and answerable" (p. 104). At chapter's end Emmet provides exercises for analyzing illustrative questions, and at book's end he obligingly supplies the correct answers.

Lastly, there are fields of practical endeavor, such as opinion polling and survey research, where it is of the utmost importance to formulate questions properly. Payne's (1951) classic manual provides a clear and simple analysis of questions, ending with a checklist of 100 considerations for wording a question properly. The best current manual--thoroughly readable, richly illustrated, and grounded in research--is Asking Questions, by Sudman and Bradburn (1982). Their helpful suggestions for beginners add up to the unexpected lesson that asking questions rightly is quite difficult. even for the experienced survey researcher. "Unfortunately, because of a belief that question wording is a simple matter that does not require great skill or experience, many researchers do not devote the needed time and effort to pretesting the question" (p. 122). This same lesson--that asking questions is a complex skill--is drawn by Hargie, Saunders, and Dickson (1981) in a nice review on questioning as a social skill in numerous practical contexts.

The docile evaluator can profit from any of these sources, applying their hard-won lessons, their experience, skill, and research knowledge, to the task of analyzing the questions proposed for evaluation research.

The analytic task is not as difficult as the language of analysis may seem to make it. It requires asking only a couple of very smart questions about the proposed question. The obstacle does not lie in the task but in the attitude behind it. Even to undertake the task, the evaluator must be convinced of two odd beliefs:

1. It is the easiest thing in the world to ask a wrong research question, and most difficult to ask a right one--not to ask a world-shaking significant question, just to ask an ordinary one in the right way.

2. By far the major effort of research has to go into getting the question right--rather than designing the right method and getting the right answer.

The more common beliefs seem to hold that posing the question is of course important but nothing much need be done about it; what needs to be done is to answer the question--especially by performing approbated methodological and statistical machinations. The suggestion here is rather: Before answering the question, and even before posing it (for research), analyze it.

It seems a matter of mere common sense to undertake a leisurely reconnaissance of the (presuppositional & presumptive) grounds of the question before venturing forth upon them with the expensive and cumbersome train of research. Even Napoleon, of course, once failed to reconnoitre sufficiently and pressed the attack with Milhaud's heavy cuirassiers, all of whom, it will be recalled, arrived just at the peak of their plan to tumble one after another into the sunken road that cut unseen across the grounds clearly in face.

The faithful evaluator is not responsible for the grounds but only for knowing what grounds he/she must cautiously proceed upon. Before posing the question it is useful to scrutinize its grounds.

Answering the Question

Before addressing the question, construct a dummy answer.

Long before taking that first step towards finding the answer, the evaluator constructs a model of the answer. It is not the eventual answer that will emerge but a dummy of it. Making the dummy answer is harder than analyzing the question but easier than classifying the questions to ask. The dummy also takes the longest time and is the most vexing part of the question-answering process. For, it involves other people and their various opinions and intentions regarding the question.

Naturally, here as elsewhere anyone can go right ahead and get an answer without troubling to construct a dummy of it. The answer can be a true one, too, and also a genuine one. The problem is that the answer still won't be any good. The evaluator brings the answer in and people either ignore or reject it. So, first make a dummy answer and bring that in: Is this what you're looking for? Is this what you want to know? Can you use something like this? What would you do with this one? and how about that model over there?

Here as elsewhere, it is a matter of pragmatics--pragmatics as a theoretical aspect of the study of questioning, and pragmatics as a

practical aspect of question-answering. Scholars have not completely worked out the theoretical aspects of these pragmatics. But the practical aspects can be worked out by the pragmatic evaluator. Some helpful formulations will be reviewed for various pragmatic aspects of giving answers. Each formulation seems to point to the wisdom of making a dummy of the answer before the answer makes a dummy out of you.

Multiple answers

Only in an ideal world is the evaluator likely to be called upon to answer a single question, not to say ask a single question by his/her lone self. Even were there to be a single question it would likely change from start to finish and in between to boot, becoming different questions at various stages of the evaluation. And, even were there to be a single question remaining one and the same throughout, still there are various people who are posing it--probably different questions--and still other various people who are awaiting the proposed results--probably different answers.

The evaluator may be faced with giving multiple answers to various people asking different questions. Regrettably, the pragmatic evaluator is burdened from the outset with finding out from all sorts of people such things as:

1. who is asking which question?
2. who is going to get the answer?
3. who wants the answer?
4. what do they already know?
5. what do they propose to do with the answer?
6. what answer do they anticipate, and which one will they (a) accept and (b) reject?

It is useful to find these out before setting out to answer the question. The upshot will undoubtedly be to change the question identified for research--to abandon it as posed, to modify and reformulate it, to multiply its variants, to add and delete some others.

The evaluator accordingly has to fiddle with the research design; and if free advice has been followed, there will not have been much of a design already prepared at this stage. Methods will have to be refined, resources reallocated, emphases shifted or reversed, plans subtly laid to give this and that answer in this and that way to this and that person. Make a dummy of the answer before making a mock-up of the design.

Unwelcome answers

Without constructing a dummy answer, the evaluator will bring the

arduously elicited results into the expectant arena of events to receive such welcomes as these.

1. That's obvious!

2. That's absurd!

3. That's irrelevant!

These niceties of reception reveal that the answer has surprisingly proven to relate to the audience's assumption-grounds in particularly unwelcome ways. Murray Davis (1971, p. 327) specifies these as follows.

1. obvious. Instead of denying, the answer affirms some aspect of the audience's assumption-ground. "Husbands often influence their wives' political behavior." It says: "What seems to be the case is in fact the case; what you always thought was true is really true." The audience's response is: "That's obvious!"

2. absurd. Instead of denying some aspect, the answer denies the whole assumption-ground. "Social factors have no effect on a person's behavior." It says: "Everything that seems to be the case is not the case at all; everything you always thought was true is really false." The audience responds: "That's absurd!"

3. irrelevant. Instead of denying or affirming, the answer does not speak to any aspect of this assumption-ground at all. "Eskimos are more likely than Jews to..." It says: "What is really true has no connection with what you always thought was true." The audience responds: "That's irrelevant!"

We may guess that these welcomes reveal, not that the answer is wrong, but that the question was the wrong one to answer. It may have been, logically speaking, a dumb question (Belnap, 1969); and it probably was, pragmatically speaking, a dumb one to answer. Far smarter to make first a dummy answer than to answer a dumb question. Then the dummy can proceed to tell everybody everything that they always wanted to know but were too dumb to ask.

Interesting answers

As a result, the audience will welcome the answer with, "That's interesting!" As Davis (1971) defines it, an interesting proposition denies the truth of some part of the routinely (but weakly) held assumption-ground of the audience. It is interesting not because it tells them some truth that they did not already know, but that some truth that they already know is wrong (p. 327).

As an "Index of the Interesting," Davis provides a classification of twelve kinds of propositions about a phenomenon, according to logical categories attaching to it or its relation with other phenomena. Each proposition is of the form: What seems to be X is in

reality non-X (p. 313), where X represents the characteristic in the logical category. For example, here are three of the twelve categories, each illustrated by a proposition from Sigmund Freud.

(1) Composition. What seem to be assorted, heterogenous phenomena are in reality composed of a single element. Freud proposed that the behaviors of children, primitives, neurotics, and adults in crowds, as well as dreams, jokes, and slips of the tongue and pen, are various manifestations of the same instinctual drives.

(2) Generalization. What seems to be a local phenomenon is in reality a general phenomenon. Freud proposed that sexual impulses are a major influence on the behavior not only of adults (which was fairly obvious) but also of children (which was not so obvious).

(3) Co-existence. What seem to be phenomena which cannot exist together are in reality phenomena which can exist together. Freud proposed that love and hate are compatible.

Davis proposes that this Index can be used as a criterion to determine whether or not a particular proposition is interesting (p. 327). Construing propositions as formed of a question-answer pair, we can appreciate that Davis has provided a useful classification of interesting answers to questions.

We may also appreciate that to give an interesting answer is a matter apart from giving a correct answer. The evaluator's audience can reject the answer's value while affirming its truth. In the audience are program advocates and opponents, experts and lay persons, and so forth, whose relevant beliefs and hopes are mutually contradictory. The assumptions of the one already constitute a denial of the assumptions of the other. Which answer is interesting to whom?

Experts, for example, will find interesting an answer that strikes laypersons as obvious. It denies expert assumptions while affirming lay ones. It says: "What everybody, except experts on the subject, think is true is in fact true" (Davis, 1971, p. 331). That's interesting! A boring answer tells experts that what they think is true is true. To them, That's obvious!

To give an interesting answer, then, the evaluator must first have learned what each of the various parts of the audience already know and assume, with respect to that particular aspect of the phenomenon which is in question (e.g., the aspects distinguished in Davis's classification). Let the dummy go around and find these things out beforehand.

Influential answers

A related practical formulation has been worked out for evaluators by Cronbach (1982), in two chapters on choosing the questions and providing the answers. In planning to provide the answers, the evaluator is urged to face these questions:

1. Will each fraction of the audience (a) attend to the answer, (b) understand it, and (c) find it credible?

2. Will the answers alter the preconceptions of the audience?

3. Will the answers enrich and elevate the dialogue leading to decisions? (p. 11)

The valence of the answer must be considered in relation to the values of the audience. Cronbach points out that in some cases (e.g., Laetrile for cancer) a positive answer will have great effect on unbelievers but a negative one will leave enthusiasts unmoved. As a general case, Cronbach suggests assessing the possible questions by the degree to which the anticipated answers promise to reduce uncertainty and to exert leverage. "The planner is to ask: 'How much influence is each of the conceivable answers to this research question expected to have?'" (p. 226).

We can appreciate that, to answer Cronbach's questions, the pragmatic evaluator must discover and weigh the socio-political as well as cognitive positions and purposes of the various recipients of the answer. Otherwise, the great effort to obtain reliable results, even meaningful and true and good ones, will eventuate in the frustrated offer of a useless answer. Let the dummy do it first.

Cooperative answers

At what might be called a technological level, the pragmatics of man-machine question-answering have been worked out by Aravind Joshi and associates. They point out that giving a direct, logically correct answer is sometimes not helpful; cooperative answers are called for. And they have produced a computerized query system, CO-OP, that gives cooperative answers.

Cooperative answers are called for in the frequent case of discrepancy in the "mutual beliefs" of the user and the system (questioner and answerer), regarding the structure and content of the data-base. Joshi defines a cooperative answer as "both giving a truthful and informative response and 'squaring away' the discrepancies in mutual beliefs discerned during the interaction" (1983, p. 237). Varieties of cooperative responses include correcting the questioner's false presumption and providing supportive and/or suggestive information in addition to that requested. For example, when no linguistics courses had been offered:

Q. Which students got an F in linguistics in Fall 1980?

A1. None. (correct, direct, but misleading answer)

A2. I don't know of any linguistics course in Fall 1980. (corrective, indirect response, cooperative & informative)

(Note on terminology: As Joshi and Kaplan define it, "presumption" incorporates part of what others--including this paper--label as "presupposition," along with some of what others also label, with Joshi and Kaplan, as "presumption.")

CO-OP is a natural-language query system (Kaplan, 1981) that does just that--gives cooperative answers. It computes and checks presumptions detected in the question's linguistic structure; when it detects a false or discrepant presumption, CO-OP generates a cooperative response rather than the meaningless correct answer. Since the presumptions of a question can be partially ordered by an entailment relationship--the failure of some entails the failure of others--CO-OP responds most appropriately by correcting the "least-failing set" of presumptions (p. 137), rather than provoke a series of little corrective question-answers or, worse, a series of little correct, direct answers--"stonewalling" (p. 130). To illustrate stonewalling:

Q1. Which students got an F in linguistics in Fall 1980?

A1. Nil.

Q2. Did anyone fail linguistics in Fall 1980?

A2. No.

Q3. How many people passed linguistics in Fall 1980?

A3. Zero.

Q4. Was linguistics given in Fall 1980?

A4. No.

In general, then, the principle runs that not only must the respondent not give a false answer, but also must not permit the questioner to infer from the true answer something that the respondent knows to be false (Joshi, 1982, p. 190). Not only must the respondent correct the questioner's presumptions but also supply additional information (supportive and suggestive) to make the answer helpful--both truthful and informative (Joshi, 1983, p. 237).

CO-OP is not the only machine that answers questions, nor the only one to do interesting things before answering. PLIDIS is used in the control of industrial water pollution in Germany (Berry-Rogghe, Kolvenbach, & Lutz, 1980). It is designed to give communicatively adequate answers to busy officials who are expert about pollution but naive about the question-answer system. PLIDIS checks their questions and corrects erroneous assumptions instead of giving a correct but meaningless answer.

The answer scheme implicit in a question is "a meaning structure" including presuppositions and presumptions (p. 146). For example, the

question, "Did Brecht check the samples from Lauxmann in Stuttgart in 1976?" presumes that Brecht is entitled to collect samples, and it 2 presupposes the existence of: the person Brecht, the firm Lauxmann, the location Stuttgart, the year 1976; samples from 1976, firms in Stuttgart, a firm Lauxmann in Stuttgart, and samples from the firm Lauxmann in Stuttgart in 1976 (p. 192). PLIDIS corrects false presuppositions as in this example:

Q. How many toxic elements were found in the November sample at VW?

A. There was no November sample.

It also corrects false presuppositions, and further it clarifies vague questions, as in this example:

Q1. Was there any cyanide in the Lutz sample?

A1. Lutz unknown, what syntactic class does the word belong to?--noun, adjective,...

Q2 It's a name.

A2. Name of what kind of thing?--company, person, place...

Q3. Person.

A3. Lutz is not entitled to collect samples. (correcting a false presumption)

We might say that, before giving an answer, PLIDIS takes care to discover what the questioner is asking, what he/she understands and means by the question, and what sort of answer will satisfy not just the question but especially the questioner. Whatever the question, what is it that the questioner knows (presumes & presupposes) and what is it that he/she wishes to know (an informative answer)?

We can appreciate that, if machines can be made to do as much, then the pragmatic evaluator, ever harried and burdened, can probably work out a way to give cooperative answers to even the stupid questions imposed on him for evaluation purposes. One way to do that is to make a dummy of the answer and bring it in for a little talk with the questioner--for a dialogue that will clarify the question for the questioner, correct his false assumptions, and help him to put into words what he really means to ask. Dummy answers are smart.

Significant answers

At a theoretical level, the most sophisticated formulation of question-answer pragmatics is Grewendorf's (1983) concept of "pragmatically significant answers." Briefly, an answer is significant when it is both informative and useful to the questioner in that situation in which the question is asked.

To be both informative and useful, the answer must take into account the knowledge and the purposes of the questioner in the question-situation. For example, here is one question (Q) with five different answers according to situation (S): "Where is Lutter & Wegner?"

S1. In Munich two friends are conversing about taverns and another friend joins in, realizes that L & W is a tavern, and asks Q. Answer: In Berlin.

S2. In Berlin someone gets into a taxi and says "Take me to L & W." The driver asks Q. Answer: Schlueterstr. 55.

S3. At the Wittenbergplazt in Berlin a pedestrian asks Q. Answer: Take the Underground to the stop BU and then ask again.

S4. On the corner of Kant and Leibnitz streets in Berlin a pedestrian asks Q. Answer: Go straight ahead, take the first left as far as the second traffic light and it's right there.

S5. From a point where one can see L & W a pedestrian asks Q. Answer: Over there.

Grewendorf points out that the range of significant answers is different in each of these situations, to the point that in no case does the given answer in any one situation belong to the range of alternative answers available for any of the other situations (pp. 72, 73). We can appreciate that all five of the answers given to this question are correct answers, and that any one of them may be informative. Yet only one of them is in addition useful, helpful, valuable--and that one is a different one in each of the five cases. That one is the pragmatically significant answer.

We may put it that the answer provides that which the questioner needs to know in order to complement his situational ignorance and to accomplish his situational purpose. The pragmatic evaluator is, regrettably but necessarily, charged with finding these things out before finding out the answer. He/she must, then, make assumptions and estimates as specif by Grewendorf (1983), about:

(a) the credulity of the questioner (The questioner believes what the answerer says to be true);

(b) the questioner's state of knowledge, in this situation;

(c) given (b), the questioner's purpose, in this situation;

(d) given (abc), the value for the questioner, of a given possible answer in this situation. (pp. 76, 78)

Then answerer compares and weighs the possible answers, then gives a pragmatically significant one according to this pragmatic postulate:

> Choose among the answers which you think are true that one for which the expected/assumed (by you) pragmatic significance is greatest. (pp. 79, 80)

Our own pragmatic postulate would suggest that the answerer do these things before going out to get the answer and, indeed, before even setting out the question (as a question for research). This pragmatic postulate runs:

> Before addressing the question, make a dummy of the answer.

This dummy can save everyone a good deal of money, time, effort, and tears. Bringing in a dummy might even give the evaluator a reputation for being smart.

Reporting the Answer

Before stating the answer, state the question.

If research is a question-answering process, then its result is a question-answer proposition. In this proposition is represented the knowledge that research yields. The meaning of that knowledge is not given by the statement that is the answer; meaning is located in the question-answer pair that constitutes the proposition--the knowledge proposed by research.

Aristotle put this epistemological point in, as usual, a simple and magisterial statement: The kinds of things that we know are as many as the kinds of questions we ask; and it is in the answers to these questions that our knowledge consists (Posterior Analytics, 89b). The implication too is simple: State the results of the study in a question-answer proposition. The implications of not doing so are complex and awful.

It is a familiar fact of language that a question can have several answers. It is familiarly overlooked that an answer in turn can have several questions. To which question does this statement stand in relation of answer? Without knowing that question (not another), that answer cannot be known. What will be known in that case is the proposition that the recipient construes by force of devising a question to which the statement may then be adduced in answer.

Such is the activity that a recipient or reader of the evaluation report necessarily engages in--but not necessarily wittingly--in the effort to understand the research. What does this study mean? What do the results show? etc.

And such is, consequently, the provenance of some vast and hopeless confusion over the research knowledge in any given domain or on any given topic: A wonderful diversity of questions is attached to

one and the same given answer. Reviewers then enter to contribute clarifying syntheses by the device of reversing the situation: Attaching to one and the same given question, of their customarily unexpressed formulation, a wonderful diversity of answers; and then declaring the one of their choice that best represents the state of knowledge on that issue--i.e., reversing things once again so as to adduce someone else's answer to their question.

To anticipate the client's need for meaning, and to forestall his/her perplexed reaction (But what does it mean? What is it good for?), the evaluator can do one simple thing: Report the results in a question-answer proposition.

It is a very simple matter. First, state the question: "What color is the program?" Then state the answer: "The color of the program is red." This elementary task takes extraordinary discipline, not often observed in research reports; and it yields an oddly powerful measure of knowledge in a most economical way, traits again not often observed in research reports.

Especially where evaluator and client have together explored the grounds of the question (as suggested above) during the act of asking it, then the two of them are more likely to share one understanding of the question-answer pair, the knowledge proposed by the evaluation study. Others too who have not shared in the questioning will yet have the advantage of construing the answer, which they read or hear, in terms of the question that precedes and complements it in the report of the study.

At the start of the report, the evaluator may set out all the assumptions of the question. At the end he sets out question and answer together, in question-answer statements. This catechetical strategy protects the end-results of the study, just as previous stategies protect its beginnings. Together they enhance the meaning of the evaluation project--the conceptual and practical sense that it must be made to have, not only for those who engage in it but also for all those others who are implicated in it, presumably to their betterment.

We may note that it is not enough to state the question on the first page of the report and then the answer on the last page. There is many a slip 'twixt question and answer. And the reader is not the only one who can lose hold of the connection while threading his way through the study: Analysis of published research (Dillon, 1983, 1984b) reveals plenty of instances where no question is stated but answers are given; or where an answer is stated for a question that is not given and no answer to the stated question is given.

The answer alone ("results and conclusions") makes no sense. It makes sense to state the question before the answer, in a question-answer proposition.

Conclusion

In general, it makes good pragmatic sense to expend at least as much effort on finding the right question as on finding the correct answer. For, the correct answer to a wrong question will be rejected--and rightly so--just as an incorrect answer to the right question.

How do you go about finding the question for evaluation research? (--so did I ask myself, specifying, "What do I know about questions that might be helpful for the practising evaluator?") Some answers are: curiosity, luck, inspiration, genius. But these are not ways of going about doing it. Four systematic ways are:

1. Before identifying the question, classify the questions that can be asked.

2. Before posing the question, analyze it.

3. Before addressing the question, make a dummy of the answer.

4. Before stating the answer, state the question.

These strategies are some of the practical fruits of conceiving research as a question-answering process, and its resulting knowledge as a question-answer proposition.

References

Ball, S. (1982). The perils of program evaluation. Contemporary Education Review, 1, 166-173.

Belnap, N. D. (1969). Questions: Their presuppositions, and how they can fail to arise. In K. Lambert (Ed.), The logical way of doing things. New Haven, CT: Yale University Press. Pp. 23-37.

Belnap, N. D. & Steel, T. B. (1976). The logic of questions and answers. New Haven, CT: Yale University Press.

Berry-Rogghe, G. L., Kolvenbach, M., & Lutz, H. D. (1980). Interacting with PLIDIS, a deductive question answering system for German. In L. Bolc (Ed.), Natural language question answering systems. Munich: Hanser. Pp. 137-216.

Campbell, J. P., Daft, R. L., & Hulin, C. L. (1982). What to study: Generating and developing research questions. Beverly Hills, CA: Sage.

Cronbach, L. J. (1982). Designing evaluations of educational and social programs. San Francisco: Jossey-Bass.

Davis, M. S. (1971). That's interesting! Philosophy of the Social Sciences, 1, 309-344.

Dillon, J. T. (1983). The use of questions in educational research. Educational Researcher, 12(9), 19-24.

Dillon, J. T. (1984a). The classification of research questions. Review of Educational Research.

Dillon, J. T. (1984b). The character of inquiry in curriculum. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Emmet, E. R. (1968). Learning to philosophize (rev. ed.). Harmondsworth, England: Penguin.

Fischer, D. H. (1970). Historians' fallacies: Toward a logic of historical thought. New York: Harper & Row.

Grewendorf, G. (1983). What answers can be given? In F. Kiefer (Ed.), Questions and answers. Dordrecht, Holland: Reidel. Pp. 45-84.

Hajek, P. & Havranek, T. (1978). Mechanizing hypothesis formation: Mathematical foundations for a general theory. Berlin: Springer.

Hargie, O., Saunders, C., & Dickson, D. (1981). Social skills in interpersonal communication. London: Croom Helm.

Joshi, A. K. (1982). Mutual beliefs in question-answer systems. In N. V. Smith (Ed.), Mutual knowledge. London: Academic. Pp. 181-197.

Joshi, A. K. (1983). Varieties of cooperative responses in question-answer systems. In r. Kiefer (Ed.), Questions and answers. Dordrecht, Holland: Reidel. Pp. 229-240.

Kahn, C. H. (1978). Questions and categories. In H. Hiz (Ed.), Questions. Dordrecht, Holland: Reidel. Pp. 227-278.

Kaplan, S. J. (1981). Appropriate responses to inappropriate questions. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), Elements of discourse understanding. Cambridge: Cambridge University. Pp. 127-144.

Keenan, E. L. & Hull, R. D. (1973). The logical presuppositions of questions and answers. In J. S. Petoefi & D. Franck (Eds.), Praesuppositionen in philosophie und linguistik. Frankfurt: Athenaeum. Pp. 441-466.

Knight, T. S. (1967). Questions and universals. Philosophy and Phenomenological Research, 27, 564-576.

Levin, H. M. (1983). Cost-effectiveness: A primer. Beverly Hills, CA: Sage.

McGuire, W. J. (1973). The yin and yang of progress in social psychology. Journal of Personality and Social Psychology, 26 446-456.

Nevo, D. (1983). The conceptualization of educational evaluation. Review of Educational Research, 53, 117-128.

Payne, S. L. (1951). The art of asking questions. Princeton: Princeton University.

Rescher, N. (1982). Empirical inquiry. London: Athlone.

Smith, N. L. (1980). Federal research on evaluation methods in health, criminal justice, and the military. Educational Evaluation and Policy Analysis, 2, 53-59.

Smith, N. L. (1981). Developing evaluation methods; Metaphors for evaluation. In N. L. Smith (Ed.), Metaphors for evaluation. Beverly Hills, CA: Sage. Pp. 17-49; 51-65.

Smith, N. L. (1982). Introductory guidelines for designing evaluations of teacher education programs. In S. Hord, T.V. Savage, & L. J. Bethel (Eds.), Toward usable strategies for teacher education program evaluation. Austin, TX: University of Texas. Pp. 101-129.

Sudman, S. & Bradburn, N. M. (1982). Asking questions: A practical guide to questionnaire design. San Francisco: Jossey-Bass.

Webb, W. B. (1961). The choice of the problem. American Psychologist, 16, 223-227.

Wortman, P. M. (1983). Evaluation research: A methodological perspective. Annual Review of Psychology, 34, 223-260.