



---

# Herbert Willsmore

## Final Examination

[ca. 1974]

---

— 1 —

### Question 1

Since one cannot, legally or morally, as a true experimental design would have it, select a number of individuals, divide them into two groups and provide services to one and withhold them from the other, the question looms: where do we get a control group? It would not be possible to find a control group whose characteristics and contextual setting is the same as a randomly selected group of DR clients. How could such individuals be found? If indeed they were found, does one somehow clandestinely monitor their progress over time to compare it with DR clients?

There is perhaps one feasible method of performing an evaluation which has somewhat of an experimental nature. Of course, the main feature of experimental design we are interested in capturing is that of comparing a group who received the services of DR with a group that did not. One way to find two such groups would be to seek out a particular class of disabled people who have in common an encounter with some public or private agency. One example would be a rehabilitation hospital. A search of the records of several rehab hospitals could generate a list of individuals with similar characteristics who passed through the hospital several years ago. These people could be contacted (although a good many will not be found) and their progress evaluated with respect to finding employment. The experimental group would be those served by DR. The control group would be those who have never had any contact with DR. Information on the services provided to those served by DR could be gathered from DR files. Information on the status of certain

— 2 —

measures or indicators tied to DR's goals on those not served by DR would rely on their memory of their situation several years ago, whatever time were selected as the beginning of the experimental program. In addition to measures tied to DR's objective function (preparing clients for employment) measures of a person's improvement in ability to live independently and perform activities of daily living could be chosen and used to evaluate progress of both groups in this area.

There are several disadvantages to an evaluation design outlined above. The primary limitation this design has in terms of shedding light on the central issues of DR is that it evaluates the effect of the program on only one class of client, i.e., orthopedically disabled. Extrapolating the results from this class of clients to DR's entire caseload would not be appropriate.

Another limitation of the above design in terms of scientific method is the fact that there is no actual "before" period to gather data. The above design relies on the subject's memory and his candidness. This can impose considerable uncertainty to the findings.

One additional problem presented by the above design is the notion that circumstances which kept an individual from coming into contact with DR may also have retarded their progress along whatever measures are chosen. If such were the case there would be considerable bias present. Such a bias cannot be dealt with unless it can somehow be measured; which is unlikely.

— 3 —

## Question 2

None of the evaluation methods discussed in class had any inherent mechanism which automatically indicated to the evaluation or, more appropriately, to the policy or decision makers whether a program's outcome is enough or not enough. None of the methods offers an absolute scale which can be used to make such a judgment. The proper use of each method requires that the evaluator have some understanding about the process he is attempting to evaluate. From his understanding of the process he selects indicators which will signal the effects of a program and in some methods the degree or magnitude of effects.

Cost/benefit and cost effectiveness evaluation, along with consumer participation, are the only methods which offer the evaluator an absolute scale (dollars) with which to judge a program's outcome. The valuation process is like the process of selection and specification of criteria in all the methods discussed in class, far from totally objective. In the case of cost/benefit analysis, a discount rate must be chosen. Costs and benefits must be identified. Often because certain costs or benefits are not easily measurable in dollars, they do not enter into the cost/benefit ratio. Once the c/b ratio has been determined, there remains the question of what is enough; a ratio of 1/8 or 1/1? What would be an appropriate ratio for one type of program may not be appropriate for another.

Cost effectiveness takes a different approach in that it weighs programs against one another by setting an output or outcome goal and choosing the one that fulfills that goal at the lowest cost. Here we still have a problem of objectivity. For example, is the goal (in the case of the vocational rehab program) simply to move a certain number of persons into employment

— 4 —

or is it to provide a certain level and quality of services and move people only into jobs which make use of their greatest potential and abilities. Optimization much like cost effectiveness uses the criteria of achieving a given goal at least cost. The answer to the question of what should be optimized is not inherently answered by the method.

Consumer participation is no more or less inherently objective than the other methods. The valuation here comes from those the program is designed to help. If their perspective and expectations for the program are not in conflict with the society as a whole, then this can be a useful source of evaluation.

Although neither experimental design nor causal modelling offer a basis for deciding whether the program's outcome is "enough," they do offer an opportunity to determine just what, in terms of outcome or improvement, is attributable specifically to the existence of the program. Causal modelling is somewhat weaker on this score than experimental design because one cannot rely solely on statistical correlation for identifying causal relationships. One must have a firm set of postulates about the process he is evaluating. In causal modelling and other application of multivariate regression correlation coefficients and indicators of the ratio of explained variation to total variation and hypothesis testing are the major criteria employed. Judging whether a correlation coefficient or  $R^2$  is large enough or choosing a level of significance for testing are subjective considerations with no hard and fast rules.

Site visits, systems analysis, comparative/historical participant observation and case study are methods which inherently call upon the evaluator to identify criteria. While many of the other methods generally specify

the nature of criteria to be used in the evaluation, these do not. These methods may be more appropriate for describing and understanding a process and in the case of participant/observation and site visit helping the evaluator get a "feeling" for what is happening. Systems analysis allows one to see relationships between various components of the program. A major drawback of all in this group except perhaps for systems analysis is that the results achieved or conclusions drawn from the evaluation can be quite inconsistent. That is, two different people evaluating the same program could conceivably come away with two diametrically opposed conclusions.

## Question 4

The claim is based mainly on the aversion people have to giving up their freedom of choice and their money. The government, through the political process, identifies areas of need which presumably cannot be satisfied by a market situation, i.e., public goods, and areas of need which arise from humanitarian commitment. If left to the free enterprise system public goods would be underproduced (spillovers and externalities) or not produced at all. Money is taxed away from individuals and devoted toward improving the public welfare. The call for accountability through evaluation is based on the notion that perhaps more money than is justified on the grounds of producing "public goods" or fulfilling a humanitarian commitment is being taken out of the free enterprise system and put into the public sector. People want to know what they are getting for their dollar and whether it is worth a dollar.

People are beginning to ask whether there aren't programs which aren't even producing a dollar's worth of output for each dollar spent or whose output is not viewed as essential public endeavors. They are beginning to ask whether it wouldn't be better to shift the money devoted to such programs to programs which could produce a greater output per dollar invested, or even to reduce taxes and put the money back into the hands of individuals. These are legitimate concerns in my opinion.

Evaluation offers one means of judging whether the output of a program is worth the money it costs. Of course, cost/benefit analysis offers the most concrete way of valuing performance. But there is a danger in depending solely on one method for evaluating all programs. Some programs whose outputs are quite highly valued yet to which it may be impossible to attach a dollar figure may fare poorly under the scrutiny of cost/benefit analysis, e.g., defense. Each approach to evaluation has its advantages and disadvantages. But evaluation per se will never, in my opinion, be able to replicate the precision of accountability of a market system.

One alternative to evaluation which attempts to capture some of the properties of a market system is a voucher strategy. The problem here is that not all of the services provided by the government could be achieved through a voucher system, e.g., highways, dams, defense. Achievement of a common goal which will benefit many at the same time is unlikely under a voucher system. One additional problem is that the production of the good or service may not be forthcoming from the private market because the profit to risk ratio is too high or because an excessively high initial investment or start up cost is required.

One alternative that has been put forward and which I would recommend

for future emphasis is consumer participation in planning, administering and evaluating programs. In a sense consumer participation also captures certain aspects of a market system. Consumers will fight for programs they feel are doing the most for them. They will also oppose or demand changes in programs which they view as ineffectual or even harmful, putting them out of business so to speak. In some respects to encourage accountability through consumer participation encourages political mitigation of the myriad of preferences in a democratic society. The problem here is that survival of the fittest (those who wield the greatest economic

and political power) is the determining force. This could lead to considerable inequity and chaos. However, in light of the fact that the ideal objective method will never be obtained, consumer participation would seem an appropriate subjective perspective to include with that of the evaluator and decision maker.