REMARKS ON DR. AXELROD'S PAPER

W. Edwards Deming

Consideration of costs is basic to the statistician in sample-design. Cost-functions enter formulas for optimum allocation of interviewers, optimum number of call-backs, as well as formulas for optimum size of sampling unit, number of secondary sampling units per primary unit, and optimum probabilities of selection. Any paper that promises reduction in costs is therefore of interest.

The problem of how many call-backs to make raises first of all the question of just what is a call-back? Actually, the word can mean almost anything. It can mean productive call-backs, or it can mean call-backs made at the convenience of the interviewer. Some organizations pay for call-backs but do not get them. Call-backs require rules on when and how to call back, and they require supervision: otherwise costs soar, without improvement of quality of data.

Astonishing differences show up between interviewers when one lays out the work so that comparison of interviewers is possible. Interpenetrating samples offer the possibility of detecting faults and need of retraining or need of better supervision. One may allot two interviewers to random parts of two subsamples, and calculate from the results the variance between and within interviewers.

Significant differences between interviewers, in the number and type of responses obtained, will indicate need of retraining, and need of different rules for calling back. They may also show that the supervision is not doing its job.

To show how astounding such comparisons may be, I may cite from my own experience the fact that some interviewers have found significantly more people at home at first call than other interviewers found in the same area. That some interviewers will find more people 65 and over, or more young children, is of course well known, as well as the fact that some interviewers have significantly more or fewer successes, and significantly more or fewer refusals.

One system that increases the effectiveness of call-backs is to allow the first interviewer only 3 call-backs. If the 3d trial is unsuccessful, she turns the case back to the supervisor, who allots the incompleted case to a special interviewer. Interviewers are careful not to build up too many failures, and in consequence they use judgment about when to call. (The number 3 is not important: the principle of a limited number is.)

Nothing is so hazardous as an experiment: it is too easy to draw wrong conclusions. I am glad that Dr. Axelrod is cautious and does not draw conclusions from his experiment. Actually, if he would continue the experiment, and increase the number of call-backs, he would, I believe, come to the conclusion that 5 or 6 call-backs are desirable if one wishes to maximize the amount of information per unit cost.

One can always reduce the cost on anything, by reducing quality. One can cut out call-backs entirely and save money, but what does it do to the quality?

I have known of research organizations that claim to do probability sampling, but which in actuality ask an interviewer to start at a designated point and to ring door-bells until she has completed a quota of interviews. This system requires no call-backs, but it has 3 faults: (1) It is not probability sampling, as there is no way to ascertain anyone's probability of selection. (2) What this system does is to load the sample with people that stay at home. (3) It fails to utilize the investment already placed in the survey to locate the random starting-points.

Miss Sybil Carof has been doing some work with me at Young & Rubicam on call-backs. We find that tabulations on general attitudes, and even on important demographic characteristics, are about the same with 3 attempts as with 6. I think that people forget that one does not carry out a survey to acquire information that you can buy from the Census for $2.50. One carried out a survey to learn about specific characteristics that are not in the Census.

Interviews			Women working outside the home	
Call-back	Number	Cumulated	Number	Cumulated
0	520	520	86	86
1	334	854	116	202
2	218	1072	101	303
3	87	1159	50	353
4	36	1195	22	375
5	24	1219	11	386

For example, take female home-makers that work outside the home for pay.

-- Initial call   17% female home-makers
-- After 2 call-backs, 28%
-- After 5 call-backs, 32%

The nearest Census figure on this characteristic is 34% for women in households that work outside the home for pay.

One can see from this table how the results change with call-backs. Any marketing policy aimed at households with income from working women, or aimed at working women, might go off the track if based on a survey that did not carry call-backs through the 5th or 6th recall.

Here is another example. How often do you use coconut?

-- Initial call, 1 to 3 times per month, 41%
-- After 2 call-backs, times per month, 32%
-- After 5 call-backs, times per month, 29%

Another point. Any interview, be it the initial call, or a call-back, costs money. With proper rules and supervision, the interview obtained on a call-back need be but little more expensive than the initial call. I question Dr. Axelrod's statement that the marginal cost of call-backs increases considerably. I am not sure that it need be this way.

Moreover, I would return to my statement that call-backs should be directed, not at random times; and that they should actually be made, not just billed.

I remarked a while ago that an experiment can be a hazard. It is better, I believe, to depend so far as possible on theory. Theory, if we have the right theory, gives the answers as well as the limitations on the range of validity of the answers that the theory provides.

For example, some years ago, I showed by theory and arithmetic that the amount of information per dollar increases steadily through the 6th call-back for any characteristic that varied from 50 to 100, or from 100 to 50, between people hard to find at home, and people that stay at home most of the time. The computations assumed that a call-back cost 67% more than the initial call. Any other realistic costs, and any other reasonable allowance for variability between people seldom at home and people much at home must reinforce these conclusions.

Theory is usually much cheaper than an experiment, and not so hazardous. You can correct a theory: but you can not correct a bungled or incomplete experiment except by repeating it. Moreover, people are more adept at finding faults in theory than in finding mistakes in experimental data.