ON THE CONDUCT OF AN INVESTIGATION

by

E. H. Federer and W. T. Federer

BU-855-M*

September 1984

Abstract

Notes on steps involved in conducting a survey and in running an experiment are discussed. They were used in connection with two evening lectures on a program given by the Cornell Feline Health Center. The program was held during the week of 11-17 July 1982 for individuals interested in raising and caring for cats. The program was given under the auspices of the Cornell Adult University (CAU).

---

* Number in Technical Report Series of the Biometrics Unit, Cornell University, Ithaca, NY 14853.

ON RUNNING AN EXPERIMENT

Edna H. and Walter Federer

All of us are experimenters and we perform experiments from time to time. Sometimes we perform an experiment to find an answer to some specific question but most often just to satisfy our curiosity. No matter what the reason for running an experiment, all experiments will be better and more efficient if we follow prescribed procedural rules. Each step in the planning and in the conduct of an experiment is important. If there is a failure in one step of the procedure, it often affects or completely dooms the entire experiment. Remember the old adage "A chain is no stronger than its weakest link"? It holds for experiments just as well as for the chain.

There are several ways of writing a set of steps to follow in setting up, conducting and summarizing the results from an experiment. One set of steps follows:

(i) Determination of precisely (not vaguely) what one wants to do, why one wants to do it, and for what purpose are the results to be used.
(ii) Carefully evaluate what is decided in (i) and determine if the information is available elsewhere.
(iii) Careful design and planning of the proposed investigation in (i).
(iv) Resources available, standardization and calibraton of measuring instruments and equipment to be used.
(v) Conduct of experimental investigation.
(vi) Summarization of results from experiment and interpretation of results obtained from the experiment.
(vii) Making a record of the results and conclusions.

Let's look at the steps in more detail.

DETERMINATION AND JUSTIFICATION OF SPECIFIC OBJECTIVES OF EXPERIMENT

When an idea or question makes its first appearance in our minds, it begins as a poorly formulated one. As we think more about it, the idea or question begins to take form, to become clearer, and to become more definite or specific. To help this process we should ask ourselves:

a) Precisely what is it that I want to do?
b) Why would I want to do this?
c) Of what use will the results from this experiment be?
d) Why shouldn't someone else do this experiment?
e) By the way, precisely what is an experiment?

Many, many experiments, whether done in the home or in the most scientific of laboratories, that have been or are being conducted, are poorly planned and ill-designed because the experimenter did not consider the above questions. Many experiments appear to have as their main objective: to keep the experimenter busy. By being very busy, the salary check keeps coming in some cases.

Suppose that we want to know the effect of three different commercial brands of cat food (Brands A, B, and C) on the gain in weight from the time a kitten starts eating cat food until the kitten is eight months old. We want this information to use in raising cats as we would like to use the cat food producing the highest weights of kittens at eight months of age. We would need to do this because no one else would be interested in doing it for our cattery. In order to obtain the information, we need to perform an experiment, which is an investigation of questions or conditions over which we have control in the planning, in the conduct, and in the nature of the investigation. That is, we select the commercial cat foods to be compared, we decide when and where the experiment will be conducted, we decide which kittens will be used in the experiment, and we decide how the investigaton will be conducted. We have control of the entire investigation except its outcome. If we knew the outcome, there would be no purpose in conducting the experiment.

EVALUATION OF PROPOSED EXPERIMENT

After one has determined what problem to work on, and the need for doing it, one should sit back and do a thoroughly critical evaluation. One of the first things to ask is, "Why should I, of all people, be doing this experiment?" Also is the experiment really worth doing? Other questions to ask are "Can't I obtain this informaton elsewhere?"; "Does my veterinarian or some other 'cat person' have this information?"; "Should I call the Cornell Feline Health Center?"; or "Should I go to the library and search to see if this information is available in published literature?" Other similar questions should be asked of oneself about the proposed experiment. One should be certain that results will be obtained, that these results will be useful, and that the results are reliable.

Many experiments in all walks of life are repetitions of ones that were conducted elsewhere without the knowledge of the experimenter. In many cases, a perusal of published literature would have obviated the necessity for performing an experiment. If the proposed experiment is really worth your time, it should provide new information that is useful for the objective of the experiment. Simply redoing what someone else has already done is busy-work. Occasionally, we may wish to confirm what others have reported. In this case, we would have a reason for redoing an experiment other than because of ignorance or merely to keep busy.

After the critical evaluation, one needs to decide whether or not the experiment is to be performed. If the decision is a "go" decision, one then needs to plan how the experiment is to be conducted.

DESIGN OF THE EXPERIMENT

One needs to have a thorough, precise, and complete plan of how the experiment is to be conducted prior to doing anything. All too often scientists and others plan their experiments during the course of the experiment, or at most they give very little thought to planning an experiment. As a result many experiments are poorly planned, are inefficient, and yield little if any new information.

First of all we need to determine how to conduct an experiment so that it will be useful in practice. In other words, what is our experiment representative of? The area, material, or entities to which an experiment is representative are known as the population or universe for which the experiment is a representative sample and the results of an experiment are useful in predicting what will happen in the universe of the experiment when the results are applied. Going back to comparing the effect of three commercial brands of cat food, A, B, and C, on the gain in weight of kittens, the kittens selected for the experiment must be representative of all future unborn kittens in our cattery to which the results will apply, and the conditions of the experiment must be such that they are applicable to the conditions in our cattery. The type of experimental material used in any experiment must be representative of the universe to which the results of the experiment are to be applied.

The selection of which treatments, i.e. the items of interest to the experimenter, are to be used is an extremely important aspect of planning or designing an experiment. For the above example, the three cat foods to be used were A, B, and C. The selection of the treatments to be used in an experiment is known as the treatment design. In many medical and other experiments, the results of the experiments were useless or misleading because adequate points of reference, standards or controls, were not included. In the above cat food experiment, one of the three should be a standard to compare the other two with. If all three are unknown, they all may be very poor cat foods and we would not know it without a point of reference. Still others were misleading because the conditions of the experiment were not applicable to the universe for which they were meant.

The kind, number, and type of records to keep should be decided during the planning stage. One should not, as many experimenters do, decide this during the course of the experiment. Such a procedure results in incomplete and omitted records that should have been made for the success of the experiment. Remember, once the experiment has been completed, there is no way to retrieve omitted information.

Another aspect of the planning stage is how large is the experiment to be and is there sufficient experimental material available to conduct the experiment as planned. The amount of experimental material to which one treatment has been applied is called an experimental unit or unit of the experiment. The number of experimental units to which one treatment has been applied is called the number of replicates (repetitions) for the treatment. If we have t treatments and we are to have r replicates of each treatment, we require tr experimental units to conduct the experiment. We would like to use $r = 1$ because of cost but cannot because of the variation from experimental unit to experimental unit. The larger the number of experimental units, the more reliable our results will be but the greater the cost. We have to balance desire for reliable results against cost. One should not start an experiment and then have to abandon it because of lack of experimental material and finances.

The responsibility for conducting the experiment, for taking the records, and for storing them in a safe place should be assigned. There should be no doubt as to a person's responsibility. If there is, one can bet that some records will not be kept or that the experiment will not be conducted in a proper manner. One should ascertain that sufficient personnel are available to conduct the experiment throughout the entire period in which the treatment is conducted.

The plan or design of an experiment is known as the experiment design. It refers to the arrangement of treatments in an experiment. The treatments should be arranged in an experiment in such a manner as to make contrasts among treatments as precise as possible and to have the comparisons as precise as possible for extraneous effects. One way to make treatment comparisons as precise as possible is to group or block the experiment material in such a way as to have as little variation between experimental units within a block as is possible. It doesn't matter how much variation there is between blocks. What we want is to have the experimental units within a block as alike as possible. Then, we assign the treatments to the experimental units within a block randomly. For example, number the experimental units; then place their numbers in a container, shake up the numbers, and blindly draw out numbers. The first number drawn is assigned to treatment one, the second number drawn is assigned to treatment two, etc. until all the treatments have been assigned an experimental unit in the block. We do this to be fair to all treatments. All have an equal chance of being allotted to a particular experimental unit. In several blocks with a new randomization in each block, the kind of experimental units a treatment is assigned tends to average out over low and high experimental units. If there is no blocking, one simply numbers all the tr experimental units for the t treatments. Since each treatment receives r experimental units, we simply take the first r experimental units drawn from the container and assign them to treatment one, the second r experimental units to treatment two, etc., until all experimental units have been assigned. We should not use a procedure which biases the results in favor of one treatment or another. A large number of medical experiments have been performed in such a manner as to favor a particular treatment. Several experimenters reported a medical treatment to be effective but when the experiment was conducted correctly, it was found that the treatment had no effect or was even detrimental.

An experiment design which has all t treatments in each block, which has r blocks, and which has the treatments randomly allotted to the e.u's within each block is known as a randomized complete block experiment design. It is a plan of procedure, hence the word design; it is a plan or design for an experiment, hence the words experiment design; it has all the t treatments in each block, hence the word complete; and the treatments are randomly allocated to the t experimental units within each of the r blocks, hence the word randomized. For example, let us return to the three cat foods problem. Suppose that we plan to use three male kittens from a litter and we plan to use five such litters. We also plan to use three females from a litter and to use five such litters. This would give us ten blocks or litters with three experimental units each, the three kittens. We number the kittens from 1 to 3 in block or litter one, and do likewise in each of the other nine litters. Then we blindly pick the number out of a container and assign the first number, say 2, to treatment A, the second number, say 3, to treatment B, and the last number in the container, 1, to treatment C. We put the numbers back in the container and repeat the process for block two, etc. until all blocks have been used.

There are many, many types of experimental designs, constructed to satisfy specified conditions. Some experiments call for more complicated designs, and if so, it is advisable to consult a statistician who understands designing experiments.

STANDARDIZATION OF EXPERIMENTAL PROCEDURES AND CALIBRATION OF MEASURING INSTRUMENTS

During the conduct of any experiment, procedures should be standardized and kept constant throughout the experiment. Changing procedures in the middle of an experiment can make the results invalid. The equipment used should be correctly calibrated and it should remain so during the course of the experiment. One should not change equipment unless both the old and the new equipment have been calibrated to give the same measurements. One should ascertain that the equipment required will be available for the experiment. Other resource needs should also be planned for before the experiment is conducted.

The experiment needs to be conducted so that personal biases of the experimenter do not affect the results of the treatments. We all have our biases and consciously or unconsciously they may creep into our procedures and bias the results in favor of one of the treatments.

CONDUCT OF THE EXPERIMENT

The previous discussion related entirely to planning the experiment. After the planning has been completed, the experiment is conducted. Careful adherence to the procedures for conducting the experiment must be followed. The measurements should be made when planned and records of these measurements recorded and stored in a safe place. If the records are lost, the entire experiment turns out to be just a tragic experience. One should be on the lookout for unforeseen contingencies as well as for unusual and unplanned-for events in the experiment. The experimenter needs to be a keen observer. Knowledge of the entire conduct and of events that occurred will be useful in interpreting the results from an experiment.

Many important findings in science were the result of the curiosity and alertness of the experimenter during the course of the experiment. One must be aware that an unusual event is happening in order to capitalize on it. Many unusual events go unnoticed in experiments simply because the experimenter is not alert to what is happening. An experimenter needs an intimate knowledge of the experimental material and of the conduct of the experiment in order to be able to summarize the results from the experiment and draw conclusions for future use.

SUMMARIZATION AND INTERPRETATION OF THE RESULTS FROM EXPERIMENTS

After the experiment has been conducted and the results obtained, it is necessary to summarize the results. There are many statistical methods for summarizing results from experiments but perhaps the easiest to use is graphing. By plotting the results of an experiment, the information contained may be obvious at a glance. Unusual results may also become evident. An explanation of why the events are unusual would be desired. One often computes arithmetic averages or obtains a percent. More complicated prodcedures are available, and it is suggested that a statistician be consulted when such procedures are desired.

After a correct and careful statistical analysis of the results from the experiment, the experimenter is in a position to draw conclusions from the results (data) for use in the universe of the experiment. Here again keenness of observaton of the experimenter is required. He must be alert to what has happened in the data, and from this alertness, the information in the results can be extracted and used to interpret the data.

RECORD OF RESULTS AND CONCLUSIONS

One of the most appropriate methods of having a record of the results and conclusions of an experiment is to write a report. Once one starts writing the report, it often becomes clear that more summarization of the data needs to be made. It also helps to clarify one's thinking. Writing reports is difficult at first for most people but the writing becomes easier with each passing report. Another method is to record the experimenter's conclusions on tapes with the tapes and records being stored in a safe place. Preparaton of a final report of the experiment should always be made with the entire procedure being thoroughly described and with the results and conclusions being included. Other readers must know of the procedure followed in an experiment in order to decide whether or not they want to believe the reported results.

ON CONDUCTING A SURVEY

Edna H. and Walter T. Federer

1. Introduction

survey: is an investigation to obtain information on a specified set of questions and/or phenomena.

population: is composed of sampling units, the smallest unit on which sample information is obtained in a survey.

census: a survey of all sampling units, members, in the population.

sample: a survey of a fraction of the sampling units in a population; it is also called a sample survey.

2. An Example of a Survey

Disease and Death Survey by Cornell Feline Health Center 1975-1981. (See pages 73 and 74 of "Felis *domesticus*: A Manual of Feline Health.")

3. Steps in a Survey

Complete description of population to be surveyed: including a definition and description of the elements, sampling units, making up the population; this should be written down in a clear and precise manner.

Description of sampling frame: A sampling frame is a complete description of all the sampling units, elements, of the population. It frequently will be a listing of all members of the population.

Determination and description of the desired information in precise terms: Vagueness about type of information and the procurement of extraneous information should be avoided.

Determination of availability of information from other sources: A check should be made to determine if information can be obtained from other sources. If it can be procured elsewhere, there is no need to do a survey.

Procurement of pertinent information: Considering the previous two steps, only relevant information pertaining to the objectives of the survey should be obtained.

Determination of sample survey design: In a population, there are often many subpopulations making up the total. The sampling units within a subpopulation may be more alike than are sampling units from different subpopulations. Any natural groupings within a population can be used to increase the efficiency of a sample survey design. The chief characteristic of a design is that it should be representative of the population. Random sampling procedures are one way to assure representativeness.

Sample size: The larger the fraction of a population surveyed in a representative manner, the more reliable will be the results. However, time and cost considerations must be considered along with reliability and repeatability of results from a survey.

Construction of questionnaire or reporting form: (to be discussed in more detail later).

Training interviewees: Effect of approach, dress, age, race, sex, and occupation should be considered. Interviewers should obtain interviewees' responses and should not influence their answers in any manner.

Determination of how to handle the problems of "unavailable", "refuse to answer", and "unable to answer" sampling units: In order for any survey to be reliable, it is necessary to have a high response rate from a representative sample. The approach must be such that this can be obtained. Despite this, it is inevitable that some responses cannot be obtained in any survey. Decisions must be made prior to the start of a survey of how the unattainable sampling units are to be handled and perhaps replaced.

Conduct of survey: The previous plans made for the survey should be scrupulously followed. Checks should be made on interviewers to ascertain that steps are vigorously adhered to.

Summarization of results and written report: Results should be summarized correctly, quickly, and efficiently, and a written report prepared. Timeliness of results for most surveys is a must.

4. Questionnaire Construction and Recording Answers

Relevant questions only: The questionnaire should contain only questions related to objectives of survey.

Length of questionnaire: It should be as short as possible.

Clear, unambiguous, unbiased, and pertinent questions: Short, uninvolved questions are best.

Recording responses: Much interviewing and summarizing time can be saved by choosing the correct form. Checks are quicker to make than writing out a response.

Consult an expert: Expert help may not always be available but anyone who has conducted a survey may be of assistance.

Pretesting the questionnaire: Before any survey is made, the questionnaire or reporting form should be pretested on a small sample not in the regular sample. After revision, it may be necessary to pretest again.

Consideration of items to simplify conduct of survey and summarizing results: Mark sense forms for responses can greatly facilitate summarization of results. Precoding answers greatly facilitates summarization.

5. Types of Interviews

Personal interviews: With properly constructed questionnaires and properly trained interviewers, this type is considered to be the most reliable and to elicit the highest response rate. It is, however, the most costly.

Telephone interviews: Conducting an interview over a telephone is a very popular and not too costly type. An interviewer can contact a relatively large number of people in a relatively short period of time. Call-backs can be made to obtain responses from individuals not available on first calls. Nonresponses are more numerous than with personal interviews. The population must be limited to those people who can be contacted by telephone. Those who cannot be contacted by telephone cannot be a part of the population.

Mail interviews: In this type of interview, a questionnaire is mailed to the interviewee; it is completed and returned by mail. Nonresponse for this type of survey can range from zero to 100 percent. Although they are known to be unreliable, mail questionnaires are widely used. Low cost related to this type is the main factor for their popularity. Mailing lists are often available, and a form is sent to all or a part of the list.

Newsletter, newspaper, magazine, door-to-door leaflets: This type of interview is the cheapest of all forms and is included in the publication at relatively little cost. If return postage is guaranteed, this will increase the cost but if not, then the only cost, if any, is the extra space taken up by the questionnaire.

Inducements: For any of the above, a monetary gift or information inducement might be included. Those participating in the survey could receive a summary of the findings of the survey. Return postage is one form of inducement.

Combinations: In order to obtain more reliable results and keep costs low, a combination of the above could be used. For example, the sample could be selected and mail questionnaires sent to all members of the sample. Those who did not respond would then be contacted by telephone. Then, one might decide to use some form of inducement or send an interviewer to personally interview the nonrespondants. In this way, one could end up with a low nonresponse rate.

6. Interviewing Techniques

Direct method of questioning: All the above has related to formulating a question and asking the interviewee to respond directly to the question.

Indirect method of questioning: This method may take various forms. For example, a relative or friend is asked to respond to questions about you. Or a series of questions related to the main question may be asked. These are designed to obtain a response to a main question which is never asked directly.

Nonanonymous response:The response to a question is known for each interviewee. If sensitive, embarassing, or incriminating questions are asked, an interviewee usually does not want to have his response known. If the interviewee believes his response to be anonymous, an answer may be obtained. If he believes that his response can be associated with him, no response or a nonincriminating response will be obtained.

Anonymous response: Techniques have been devised to maintain the anonymity of an interviewee's responses. To be effective the interviewee must be convinced of anonymity.

7. Techniques for Maintaining Anonymity

Envelopes with no identifying marks: Interviewees frequently believe there are invisible indentifying marks on an envelope.

Sealed or locked box where a questionnaire with no identifying marks may be placed: This is sometimes called a "black box" procedure.

Randomized response: In this procedure, an interviewee is presented with two questions, one of which is to be answered by a random process. The interviewer never knows which question was answered as he only receives one answer. To illustrate, suppose the two questions are A and B, where A is a sensitive question. One randomizing device is to put ten identical tags in a container with the tags being numbered 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9. Suppose nonsensitive question B is "Is the number you drew 7 or 8?" Now, you answer yes or no but the interviewer does not know if you answered A or B. He knows that if you answer B, in 2 out of 3 cases you will answer yes. Knowing this he is able to estimate the proportion in the sample who answer yes to question A but cannot determine which question you answered. Hence, your response is anonymous.

Block total response: With this procedure, an interviewee is given a set of 2, 3, or more questions. One total for the set of questions is given. Responses for individual questions are not known to the interviewer. Since a total can arise from many sets of responses, responses to individual questions are anonymous. Basically, there are two types of block total response procedures. The first is a supplemented block total which has the following sets of questions when a total of 4 questions (A, B, C, and D) are involved. Suppose A is a sensitive question. Then the sets are:

A and B	A and D
A and C	A, B, C, and D

The interviewee gives a total for whichever set of questions he is to answer. One is able to obtain estimates for each question in the total sample but one cannot identify the response to a particular question. The second type is called a balanced block total. The sets of questions to be answered are:

A and B	A and D	B and D
A and C	B and C	C and D

From the set (block) totals, one can obtain estimates of responses to each question. Here again an interviewee's answers to individual questions are anonymous.

8. Glossary

block total response: an interviewee answers two or more questions and gives a total for all questions. Individual questions are not answered. Different sets of questions are given to different interviewees.

census: complete enumeration of every entity in a group.

nonprobability sample: chance of selecting a sampling unit is unknown.

population or universe: all the elements (sampling units) of a group or aggregation. All the cats in New York State could be the universe.

probability sample: chance of selecting a sampling unit is known.

randomized response: interviewee answers one of two questions but only he/she knows which one was answered. A random device is used to determine which question is answered.

sample or sample survey: enumeration of a fraction of the group.

sampling frame: a complete description of all the sampling units in the population. A list of all sampling units would be a sampling frame.

sampling unit: the element of the group.

survey: an investigation to obtain information on a specified question or phenomenon in some prescribed group or aggregation.