PN-AAT-076
40603

Handbook for Family Planning Operations Research Design

Andrew Fisher
John Laing
John Stoeckel

The Population Council

The Population Council
One Dag Hammarskjold Plaza
New York, New York 10017

Library of Congress
Cataloging in Publication Data

Fisher, Andrew, 1941–
Handbook for family planning operations research design.

Bibliography: p.
1. Birth control—Research—United States—Handbooks, manuals, etc. I. Laing, John. II. Stoeckel, John E. III. Title.

HQ763.6.U5F57 1983 363.9'6'072073

83-13731

ISBN 0-87834-051-3

Any part of this Handbook may be copied or adapted to meet local needs without permission from the authors or the Population Council, **provided that the parts copied are distributed free or at cost (not for profit).** Any commercial reproduction requires prior permission from the Population Council. The authors would appreciate receiving a copy of any materials in which the text from the Handbook has been used.

Printed in the United States of America

Andrew Fisher, Sc.D., is Associate and Regional Advisor for Family Planning Research, The Population Council, Bangkok, Thailand.

John Laing, Ph.D., is Associate, The Population Council, and Visiting Lecturer and Research Associate, The Population Institute, University of the Philippines, Manila, the Philippines.

John Stoeckel, Ph.D., is Senior Associate and Regional Advisor for Population and Development, The Population Council, Bangkok, Thailand.

Contents and Outline for Preparing a Research Proposal

Preface vii
Acknowledgments viii

Introduction 1

1

1 Problem Identification and Definition 3
A Problem Identification 3
B Problem Definition 5
1 Incidence and Prevalence
2 Geographic Areas Affected
3 Characteristics of Population Groups
4 Probable Reasons for the Problem
5 Possible Solutions
6 Unanswered Questions

2

2 Justification for Selecting the Research Problem 8
A It Is Current and Timely
B A Large Population Is Affected
C A Special Population Group Is Affected
D It Relates to Ongoing Program Activities
E There Are Broad Social, Economic, Health, or Political Implications
F It Is Considered Important by Many Different Groups

3

3 Objectives and Hypotheses 10
A Ultimate Objectives 10
1 Reduce Morbidity, Mortality, or Fertility
2 Influence Policy
3 Improve Service Delivery
B Immediate Objectives—Who will do, how much of what, to whom, when, where, and for what purpose? 10
C Hypotheses 11
1 Relationships among Variables
a Independent Variables
b Intervening Variables
c Dependent Variables
2 Direction of Relationship
3 Conditions of Relationship

4

4 Operational Definitions 15
A Operational Definitions of Variables 15
1 Indicators
2 Variable Categories
a Mutually Exclusive
b Totally Inclusive
B Operational Definitions of Terms 16

5

5 Study Design 19
A Reliability and Validity 19
B Validity Threats 19
1 History
2 Selection
3 Testing
4 Instrumentation
5 Maturation
6 Mortality
C Types of Study Designs 21
1 Nonexperimental Designs
a Posttest-Only Design
b Pretest–Posttest Design
c Static-Group Comparison
2 Experimental Designs
a Pretest–Posttest Control Group Design
b Posttest-Only Control Group Design
3 Quasi-Experimental Designs
a Time Series Design
b Nonequivalent Control Group Design
c Separate Sample Pretest–Posttest Design

6

6 Sampling 28
A Selecting Area and Population 28
B Selecting Cases 28
1 Probability Sample
a Simple Random Sampling
b Systematic Sampling
c Stratified Sampling
d Cluster Sampling
e Multistage Sampling
f Probability Proportional to Size (PPS)
2 Non-probability Sample
C Sample Size Determination 31

7

7 Data Collection 35
A Quantitative Data 35
1 Structured Interviews
2 Service Statistics
3 Self-Administered Questionnaires
4 Secondary Data Sources
B Qualitative Data 37
1 Unstructured Interviews
2 Focus Group Discussions
3 Direct Observation of Operations
4 Content Analysis of Written Materials
C Data Quality Checks 39
D Confidentiality of Information 40

8

8 Tabulation and Analysis of Data 41
A Preparing Tabulations 41
1 Coding Data
2 Editing Coded Data
B Plan for Data Analysis 42
1 Attributes of the Data
a Central Tendency of the Data
b Variance in the Data
c Differences within the Data
d Relationships within the Data
2 Types of Analytical Procedures
a Variable Transformations
Recodes
Counts
Scaling
Conditional Transformations
Other Mathematical Transformations
b Univariate Analysis
Nominal Measurement
Ordinal Measurement
Interval Measurement
c Time Series Analysis
d Comparisons
e Bivariate Relationships
Between Nominal Variables
Between Ordinal Variables
Between Interval Variables
f Multivariate Analysis
All Interval Independent Variables
Both Interval and Categoric Independent Variables
All Categoric Independent Variables
g Cost-Effectiveness Analysis
h Use-Effectiveness Analysis

i Fertility Analysis
C Dummy Tables 51

9

9 Reporting Research Findings 52
A Interim Reports 52
B Final Reports 52

10

10 Limitations of the Study 54
A Design and Analysis Limitations 54
B Special Situational Factors 54

11

11 Resources and Facilities 55
A Available Resources and Facilities 55
B Study Budget 55
1 Salaries and Benefits
2 Materials, Supplies, Computer Services
3 Travel
4 Miscellaneous Expenses

12

12 Study Work Schedule 57

13

13 Appendixes 59

14

14 Title Page and Abstract 60

Bibliography 63

Preface

This Handbook is designed to help you develop and write a detailed family planning operations research proposal. The Handbook begins with an outline for preparing a research proposal. Each major section of the outline is explained further in the pages that follow.

Although the Handbook is oriented toward developing family planning research proposals, it also can be used for preparing proposals in other health and social welfare areas. It is not intended to be a textbook on research methods but a guide for preparing a well-written and detailed research study proposal. The Handbook assumes some familiarity with the terms and concepts of research design and statistics as well as some experience with implementing field studies. The examples used in the Handbook are drawn from our experience and are based on situations that often occur in family planning programs.

Acknowledgments

The Handbook was written as training material incidental to the US Agency for International Development (AID), Office of Population Contract No. DPE-0632-00-1029-00, Family Planning Operations Research in Asia. After several revisions, the Handbook was used as the basic text given to participants attending operations research workshops conducted by the authors in several Asian countries. The primary purpose of the Handbook is to assist workshop participants with the development of a complete family planning research proposal. The authors are particularly grateful to the participants for their helpful evaluations of the Handbook. We also appreciate the careful reviews provided by professional colleagues at AID, the Population Council, universities, and research institutes.

Introduction

Most national health and family planning programs are intended to have an impact on birth, death, and morbidity rates. These programs are large in expenditure, personnel employed, geographic coverage, and component activities. They are complex undertakings that require detailed planning at all levels, close coordination of program efforts, careful training and supervision of personnel, and continuous evaluation of program development and impact. Operations research is a primary means available to support these essential planning, coordinating, training, supervising, and evaluation functions.

Operations research is concerned with the day to day "operations" of programs. It is intended to provide managers, administrators, and policymakers with the information they need to improve service delivery activities and plan future ones. It seeks practical solutions to problem situations and viable alternatives to unsatisfactory operating methods. It compares one service delivery approach against another in terms of impact, cost-effectiveness, and client acceptability.

All research of whatever kind is set in motion by the existence of a problem situation. A problem is a perceived difficulty, a feeling of dissatisfaction with the way things are, a discrepancy between what someone thinks should be and what is. Operational health and family planning research provides answers to perceived program problems. Typically, operations research has two basic objectives:

1. To diagnose current problem situations and examine underlying causes behind activities that are successful or unsuccessful.
2. To test new approaches to service delivery that are more cost-effective than existing approaches.

Operations research should be thought of as a continuous process of program evaluation, experimentation, and information feedback. The range of potential research topics is as large as the number of program activities. The study methods that can be used for evaluation and experimentation are also large in number. Usually operational health and family planning research studies can be categorized under one or more of the following headings:

Field Experimental Studies. Field experiments involve the introduction of a new approach to service delivery. These studies are sometimes called "intervention studies" because they seek to change or intervene in an existing service delivery system. They usually employ a prospective experimental or quasi-experimental research design in order to evaluate the effect of a new program intervention. The new intervention might be a new category of field worker, a new approach to communications, a new community participation strategy, or a new decentralized system of administration.

Cost-Effectiveness Studies. These studies compare different program approaches in terms of the relative cost to obtain a unit of output. For example, two approaches to obtaining IUD acceptors might be compared in terms of the relative cost to obtain a full 12 months of continuous IUD use. Another example might be the comparison of two different approaches to communicating with rural audiences. The comparison would try to determine which approach is able to communicate health and family planning messages to the largest number of people for the least cost.

Qualitative Studies. There are many administrative and managerial aspects of health and family planning programs that do not lend themselves to quantification. For example, it is often difficult to quantify the acceptability of services provided to clients, the structure of an organization, the process of administrative decision making, or the work patterns

of field staff. In such cases, nonquantitative research methods must be used. Such methods include participant observation, focus group discussion, and content analysis of documents. Often the best operational research studies use these methods and are able to provide an in-depth analysis of problem situations.

Quantitative Studies. Many diagnostic studies make use of quantitative methods to evaluate particular aspects of a problem situation. Such studies tend to be cross-sectional in nature and make use of a variety of statistical techniques such as multiple classification analysis (MCA).

These study categories are not mutually exclusive. Very often a researcher will make use of several different research methods. Numerous examples of operations research studies could be listed under each one of these broad categories. For the purpose of illustration only, we list below nine topics that often are the focus of operations research studies:

1. **Family Planning Training Programs.** Studies that examine the content and methods of training, or evaluate the effect of training programs on field worker performance.
2. **Information, Education and Communication (IEC).** Studies that compare different communication approaches in terms of message understanding, message retention, cost-effectiveness, and ability to reach a large audience.
3. **Service Delivery Impact.** Field experiments that prospectively test and compare selected aspects of service delivery in terms of prevalence rates, continuation rates, client satisfaction, availability and accessibility.
4. **Planning and Management.** Qualitative assessments of the management of program personnel and funds, or assessments of the flow of service statistics information among program managers as well as the utilization by managers of the information.
5. **Management Information Systems.** Studies that seek to design and test simple, low-cost data collection and feedback procedures.
6. **Commercial Channels of Distribution.** Comparative studies of private and government service delivery systems in terms of client satisfaction, client utilization, program coverage, and cost-effectiveness.
7. **Individual and Community Incentives.** Comparisons of the impact of different levels and types of individual and community incentives.
8. **Method User Perspectives.** Studies that examine health and family planning programs from the perspective of the user. Studies that assess the reasons why users switch from one method to another, or from one source of supply to another.
9. **Private Organizations.** Studies that evaluate the role and effectiveness of private organizations in the delivery of health and family planning services.

Obviously there are many other topics that could become the focus of an operations research study. Whatever the topics selected, the primary function of operations research is to improve the availability, accessibility, and acceptability of health and family planning services.

Problem Identification and Definition

1

A Problem Identification

There are countless problems with family planning and health programs. Finding a problem is not hard, but identifying one for the purpose of research is not always easy. One of the most important tasks of research is to identify and define clearly the problem you wish to study. If you are uncertain about the research problem, if you are not certain in your own mind about what you want to study, you may be sure that others who read your proposal will also be uncertain. A well-defined problem leads naturally to the statement of research objectives, to the hypotheses, to a definition of key variables, and to a selection of a methodology for measuring the variables. A poorly defined research problem leads to confusion.

All research is set in motion by the existence of a problem. A problem is a perceived difficulty, a feeling of discomfort with the way things are, a discrepancy between what someone believes **should be** and **what is.** While problems are the initiating force behind research, not all problems require research. A potential research situation arises when three conditions exist:

1 A perceived **discrepancy** between what is and what should be.
2 A **question** about why the discrepancy exists.
3 At least **two possible and plausible answers** to the question.

The last point is important. If there is only one possible and plausible answer to the question about the discrepancy, then a research situation does not exist. Consider the example given below.

1 Example of a Nonresearch Problem

Problem Situation: A recent survey in District A found that 1,000 women were continuous users of contraceptive pills. But last month's service statistics indicate that none of these women are using contraceptive pills.

Discrepancy: All 1,000 women **should be using** contraceptive pills, but all 1,000 women **are not using** contraceptive pills.

Problem Question: What factor or factors are responsible for 1,000 women discontinuing their use of contraceptive pills?

Answer: A monsoon flood has prevented all new supplies of pills reaching District A, and all old supplies have been exhausted.

In the above example, a problem situation exists, but the reason for the problem is already known. Therefore, assuming that all the facts are correct, there is no reason to conduct research on the factors associated with pill discontinuation among the 1,000 women. On the other hand, there may very well be a need to conduct research on the question of why the supply and logistics system is incapable of providing contraceptives to women during the monsoon. Study the next example.

2 Example of a Research Problem

Problem Situation: District A is always flooded during the monsoon season. Recognizing this problem, the national family planning program established a new supply logistics system for the district. Each pill user is given a four-month supply before the monsoon begins. During the monsoon, small motorboats are available to transport new supplies to selected distribution centers accessible to village-level family planning workers. Despite these new measures, this year service statistics indicated that there are no pill supplies in District A.

Discrepancy: The new logistics system **should be able** to assure a continuous supply of pills, but this year **there are no supplies.**

Problem Question: Why has the new supply logistics system been incapable of delivering contraceptive pills to users?

Possible Answers:

1. An order for new pill supplies was not placed in time before the monsoon rains.
2. The riverboats used to transport the supplies are out of order.
3. Field-workers were not told about the new system and failed to give users a four-month supply of pills before the monsoon.

In this example, there are several possible and plausible reasons for the problem situation. One or more of these reasons might be correct. Therefore, this is a potential research situation.

In some situations, it is relatively easy to identify the problem, to define it, to hypothesize the reasons for it, and to conduct research to determine which reason is correct or more nearly correct. The reasons for the supply and logistics problem in the above example could probably be determined fairly easily and certainly would not require an expensive research study. Other problems, such as the one in the next example, are not so easy to identify or to study.

3 Example of a Research Problem

Problem Situation: A recent family planning survey revealed great differences between villages in the rate of contraceptive prevalence. Despite the fact that all villages receive the same level of health and family planning services, some villages have a prevalence rate as high as 80 percent, while others have a rate as low as 6 percent.

Discrepancy: All villages **should have** approximately the same rate of contraceptive prevalence, but in fact **there is great variation** between villages.

Problem Question: What factors are responsible for the areal variation in contraceptive prevalence rates?

Possible Answers:

1. Villages differ in their socioeconomic environments. Some villages are agricultural; some are fishing communities. Some villages are Hindu; others are predominantly Muslim or Buddhist. Some villages have access to markets in towns; others do not. Some villages have schools, health clinics, electricity, and a good water supply; others do not have these facilities. These socioeconomic differences affect the level of contraceptive practice.
2. Villages differ in institutional support for contraceptive acceptance. In some villages, local influentials strongly support the national family planning program. In other villages, they do not support it. In some villages, there are active Mothers' Clubs that support family planning. In other villages, there are

no Mothers' Clubs. These differences in institutional support for family planning affect the level of contraceptive practice.

3 Village-level health and family planning workers differ in their effectiveness. Some workers are highly motivated and very active in their assigned areas. Other workers are less motivated and less active. These differences in worker effectiveness affect the level of contraceptive practice.

While the problem situation presented above is fairly clear, the reasons for the problem are complex. Three reasons have been given, but it is likely many more could be stated. In situations such as this, the researcher must devote considerable time and attention to identifying the problem situation. The aim is to focus the research on the most important aspects of the problem.

B Problem Definition

Identifying a problem situation is the first essential step in designing a research proposal, but it must then be followed by a process of **problem definition.** The research problem identified must now be defined in terms of its occurrence, intensity, distribution, and other measures for which data are already available. The aim is to determine all that is currently known about the problem and why it exists. While it is always possible to guess why a problem exists, guesses are often wrong and usually they do not provide a firm basis for designing a research study. A far better way to define a problem situation is to review relevant literature, examine current service statistics, seek educated opinions from persons concerned with the problem, and obtain probable reasons for the problem from social, economic, or health theory. A careful review of existing sources of information on a problem and an epidemiologic diagnosis help the investigator determine:

1 **Incidence and Prevalence.** An epidemiologic diagnosis should always be made of problems related to health and family planning. In other words, how widespread is the problem? What is its distribution? How often does it occur? An epidemiologic diagnosis will help establish the parameters of the problem.

2 **Geographic Areas Affected.** It is important to know if there are particular geographic areas affected by the problem. Does the problem generally occur in rural areas only? Does it also affect cities? Is the problem restricted to mountain areas, coastal areas, or island areas?

3 **Characteristics of Population Groups.** Are there special population groups affected by the problem, such as mothers, children, teenagers, newly married couples, or women over age 35?

4 **Probable Reasons for the Problem.** A review of information on a problem should suggest a number of probable reasons why the problem exists. What is the current thinking about the reasons for the problem? Is there general agreement among many people about the reasons, or are there many different, conflicting views?

5 **Possible Solutions.** Many projects and programs may have been directed at the problem in an attempt to overcome it. What types of solutions have been tried in the past? How successful have past efforts been? What approaches to solving the problem seem to work? What approaches seem not to work?

6 **Unanswered Questions.** From the review of information on the problem, what seem to be the unanswered questions about the problem? What aspects of the problem need to be researched further?

Reviewing what is already known about a problem situation is an essential part of the research process. A good review of information will suggest the social, economic, political, and health importance of the problem. It will help to narrow the focus of the proposed research. It will indicate the major theoretical concepts and operational variables other researchers have considered important. It will suggest possible research hypotheses that need to be tested. Finally, it will help the investigator avoid reinventing the wheel, or in other words, conducting research on a problem that has already been researched countless times in the past with fairly consistent findings from each study.

1 Example of a Research Problem Identification and Definition Statement

1 Although substantial progress toward reducing fertility rates was made during the last decade in several Southeast Asian countries, significant program constraints still exist and threaten further fertility declines. In Indonesia, the relatively high contraceptive prevalence rate on Java and Bali of 42 percent must be viewed against the relatively low rate of 17 percent in the Outer Island I area and only 6 percent in the Outer Island II area. In the Philippines, only 25 percent of all contraceptive practice involves the use of modern methods. In Malaysia, new sterilization acceptors constitute less than 6 percent (1979) of all new acceptors and new IUD acceptors less than 3 percent. In Thailand, the vasectomy program is lagging despite some remarkable acceptance rates demonstrated when private sector teams have taken services directly to villages.

2 Most program administrators within Asia are aware that a new level of effort and direction will be needed in order to sustain the trends established in the 1970s. In Indonesia, attention needs to be given not only to the Outer Island areas but also to the urban areas, particularly Jakarta. Although Jakarta has a health and family planning service delivery system designed to cover all areas of the city, the contraceptive prevalence rate is only 21 percent of currently married women, or in other words, approximately half of the rate for all of Java and Bali. Other large urban areas in the country also lag behind the rural areas.

3 While the factors responsible for the difference in contraceptive prevalence rates between urban and rural areas in Indonesia are not known, it has been suggested that urban areas lack the traditional structures of community organization and leadership that help support the family planning program in rural areas. Another possible reason for the difference might be lack of awareness, particularly among new migrants to urban areas, of sources of contraceptive supply. Many of the new migrants have low levels of education, do not have access to sources of information, and are unaware of the location of health and family planning service centers. Also, family planning city workers often find it difficult to provide continuous services to new migrants at their homes since these people tend to shift their residence periodically from one area of the city to another. It is likely that, among acceptors, these shifts result in discontinuation, or at least in an interruption of contraceptive practice.

4 In the past, several attempts have been made to increase the urban rate of contraceptive prevalence by providing more accessible services to residents. In Jakarta, mobile vans have been used to bring services to outlying areas of the city. Another approach has been to distribute condoms and pills through such commercial outlets as small shops and tea stalls. While both of these approaches are useful, they also have significant drawbacks. Mobile vans are expensive to operate, and they can reach only a relatively small proportion of the population of potential acceptors. Commercial outlets are helpful for current acceptors, but they are not very effective as a means to motivate new acceptors. Also, these outlets cannot provide needed maternal and child health services, such as immunizations and pre- and postnatal care.

5 What is needed in urban areas is some means of (1) reaching large numbers of current and potential acceptors and (2) providing them with a range of health and family planning services. An important unanswered question for the national family planning program is how this can best be accomplished.

6 One approach that has been suggested but so far not tried is to train a special category of urban health and family planning personnel who would contact acceptors and potential acceptors at their places of employment. With the assistance of employers and unions, the urban health and family planning worker could establish service delivery centers at factories, government offices, hotels, construction sites, markets, and other places of employment. It is likely that this approach would be less costly than the use of mobile vans; it would be capable to reaching large numbers of people; and the range of services that could be provided would be greater than that provided by commercial outlets.

2 Comments on the Example

In the first paragraph of this example, a broad problem situation is identified. Substantial progress toward reducing fertility rates in several Southeast Asian countries has been made, but continued

progress is threatened by significant program constraints. Several examples of program constraints are then given. These examples serve to indicate the geographic areas affected by the problem.

In the second paragraph, the problem of program constraints is narrowed down to just one country—Indonesia. Also, attention is focused on just one program problem—the difference in contraceptive prevalence rates between urban and rural areas. A discrepancy is indicated. Jakarta has a health and family planning service delivery system designed to cover the entire city. This system should result in prevalence rates equal to the rate in rural areas. But the prevalence rate in Jakarta is only half the rate found in all of Java and Bali. By the end of the second paragraph, we know the specific problem situation—urban/rural differences in contraceptive practice. We know the location of the problem—Indonesia, urban areas, particularly Jakarta. We know the magnitude of the problem—21 percent prevalence in Jakarta compared with 42 for all of Java and Bali. We know the specific population of interest—urban residents.

In the third paragraph, several possible reasons for the problem situation are suggested. It is not known which one of these reasons is correct or more nearly correct.

The fourth paragraph outlines several possible solutions to the problem of low prevalence rates in urban areas. Mobile vans have been used. Commercial outlets have been used. Each of these approaches to the problem is helpful, but each is rejected as inadequate.

The fifth paragraph suggests what is needed to overcome the problem, and the sixth paragraph outlines a new and previously untested approach to overcoming the problem.

What To Do—Problem Identification and Definition

1 Follow this general procedure in identifying and defining a problem situation:
   - **a** Start with a simple statement of the problem situation.
   - **b** Add details as you review the literature and investigate the problem in greater depth.
   - **c** Move back to simplicity by focusing on the most important aspects of the problem that are **researchable.**

2 Make a first attempt at identifying the problem situation by using the following format:

Problem Situation: (Write a small, simple paragraph that identifies the problem)

Discrepancy: (State the discrepancy between **what is** and **what should be**)

Problem Question: (Write the central problem question)

Possible Answers: (Write **two or more** plausible answers to the problem question)

3 From available research literature, current service statistics, educated opinions, and other sources of information, try to add detail to the problem situation you have just identified. Look for theoretical concepts and operational variables that you may have missed. List these concepts and variables on a piece of paper as you come across them. Try to answer the following questions:
   - **a** What is the incidence and prevalence of the problem?
   - **b** What geographic areas are affected by the problem?
   - **c** What population groups are affected by the problem?
   - **d** What are the findings of other research studies?
   - **e** What has been done to overcome the problem in the past?
   - **f** How successful were past efforts to overcome the problem?
   - **g** What seem to be major unanswered questions about the problem?

4 With the information you have collected from a literature review and other sources, rewrite your statement identifying and defining the problem. Use the format of Problem Situation, Discrepancy, Problem Question, and Possible Answers. Add details that help to define the problem, but organize the information. Try to establish the boundaries of the problem. Focus your attention on the most important, researchable aspects of the problem. Be clear and specific.

5 Have one or more friends read your final statement identifying and defining the problem situation. Have the friend tell you what he or she thinks the problem is. If your friend is unclear about the problem situation, if your friend cannot describe the discrepancy between what is and what should be, then go back to the beginning and start all over again!

2 Justification for Selecting the Research Problem

Now that you have identified and defined the problem situation, it is necessary next to justify the importance of the problem. Research often is expensive and time consuming. Ask yourself, why is the problem I wish to study important? Can you justify your selection of the research problem? Can you convince others about the importance of the problem?

A Example Justifying the Selection of a Research Problem

The difference between urban and rural contraceptive prevalence rates in Indonesia relates directly to the larger issues of rapid urban growth and to inadequate and ineffective service delivery to urban residents. Throughout South and Southeast Asia, urban population growth rates far exceed those in rural areas. Typically, urban areas grow at an annual rate of about 4–8 percent, while rural areas grow at a rate of 2–3 percent. Most of the urban growth is due to migration from rural areas, but as the population base of cities expand, fertility also contributes to the growth. While migrants come in search of jobs and a better life, often they find that health, housing, and sanitary conditions in cities are worse than what they left behind. In most large cities, effective means to deliver basic health and other services have not been found.

The relatively low contraceptive prevalence rate in Jakarta compared with rural areas is a problem recognized by the government family planning program (BKKBN) and top government policymakers. This problem is an indication that health and family planning service delivery to urban residents is inadequate and ineffective. New models of service delivery based on the special situation of metropolitan areas and the life-style of urban residents must be found and tested. Particularly important is the need to reach new migrants to the city and those who are in the lower socioeconomic strata. For many of these people, health and family planning services are neither available nor accessible. So long as this situation continues, there will be significant differences in contraceptive practice between urban and rural areas.

B Comments on the Example

The justification for selecting the research problem is presented in several different ways. First, it is pointed out that the problem of differences between urban and rural contraceptive prevalence rates affects many areas in South and Southeast Asia and is not just restricted to Indonesia and Jakarta. In other words, the problem is widespread. Second, this is a current and timely problem that is likely to become worse as the population base of cities continues to expand. Third, the problem is recognized as an issue of great concern by the BKKBN officials. Fourth, the problem relates to ongoing service delivery efforts, in particular to the inadequate and ineffective health and service delivery efforts in urban areas. Finally, the research justification statement suggests that, if new urban-based models for service delivery are not found, the contraceptive prevalence rate will continue to be lower in urban areas than in rural areas.

What To Do—Justifying the Selection of a Research Problem

1 In justifying the importance of a research problem, it is helpful to ask yourself a series of questions and then try to answer each of them.

- **a** Is the problem you wish to study a current and timely one? Does the problem exist now?
- **b** How widespread is the problem? Are many areas and many people affected by the problem?
- **c** Does the problem affect special groups, such as mothers or children?
- **d** Does the problem relate to ongoing program activities?
- **e** Does the problem relate to broad social, economic, and health issues, such as unemployment, income maldistribution, the status of women, education, and maternal and child health?
- **f** Who else is concerned about the problem? Are top government officials concerned? Are medical doctors or other professionals concerned?

2 Review your answers to these questions, and arrange them into a paragraph that justifies the importance of the research problem.

3 Objectives and Hypotheses

The objectives and hypotheses of a research study should flow logically from the earlier sections identifying the problem situation, defining the parameters of the problem, and justifying its importance. In this section, the research purpose is narrowed and focused. Specific objectives are written that describe the expected contributions and outcomes arising from the study and the variables that will be measured. Also, specific, testable hypotheses are formulated that give the relationship between variables.

A Ultimate Objectives

Most research studies have a statement of ultimate objectives that describes the expected implications or contributions arising from the study. The ultimate objective of many health and family planning studies is to contribute in some way toward reducing morbidity, mortality, or fertility. This of course is a rather unrealistic objective. It is unlikely that any single study will do much to reduce morbidity, mortality, or fertility. A somewhat more reasonable ultimate objective is that the study will provide program administrators and policymakers with information useful for improving service delivery activities or reformulating policy.

Ultimate objectives help to relate the proposed research to broad social, economic, and health concerns. They state the reasonable and expected contributions of the study. In this way, they contribute to the justification of why research on the problem situation is required. Note, however, that ultimate objectives are **expected** and **hoped for** contributions. In a sense, they represent wishful thinking. The investigator does not promise that the contributions will occur.

1 Examples of Ultimate Objectives

1. Study of Areal Variation in Contraceptive Practice in Bangladesh.
   "The ultimate objective of this study is to help program administrators to design new educational strategies for motivating couples in villages with low contraceptive prevalence."
2. Study in Indonesia of Urban and Rural Differences in Contraceptive Prevalence Rates.
   "The ultimate objective of this study is to increase the availability and accessibility of health and family planning services to urban residents and thus to raise the urban rate of contraceptive prevalence."
3. Evaluation Study of Different Training Methodologies.
   "The ultimate objective of this study is to improve the effectiveness of field-workers in the delivery of health and family planning services."
4. Survey Study of the Family Planning Knowledge, Attitudes, and Practices of Community Influentials in Sri Lanka.
   "The ultimate objective of this study is to increase the involvement of community influentials in health and family planning program activities."

B Immediate Objectives

In contrast to ultimate objectives that state what is **expected** to happen, immediate objectives state what **will** happen. Immediate objectives relate directly to the research problem situation.

They indicate the variables that will be examined and measured. An immediate objective represents a promise by the investigator that certain activities will take place and specific variables will be examined.

Whenever possible, and certainly in all evaluation studies, the immediate objective should be stated in **behavioral** terms—in other words, in terms of what will be done. The objective should specify: **Who will do, how much of what, to whom, when, where,** and **for what purpose.**

If your study includes an intervention, such as a training program or an information, education and communication (IEC) campaign, the nature of the intervention should be described fully after you have stated the immediate objectives. In describing the intervention, you should elaborate on the questions of who will do, how much of what, to whom, when, where, and for what purpose. In other words, try to be as specific as you can in describing the intervention program.

1 Examples of Immediate Objectives

1 Study of Areal Variation in Contraceptive Practice in Bangladesh.
"The Bangladesh Institute of Population Research will conduct, between January 1983 and December 1984, a survey study in 75 villages to determine the social, economic, institutional, and program-related factors associated with the variance in contraceptive practice among currently married women aged 15–44."

2 Study in Indonesia of Urban and Rural Differences in Contraceptive Prevalence Rates.
"In the first phase of this study, the Indonesian BKKBN will conduct in 1983 an exploratory study in Jakarta and in two rural Kecamatan to determine the institutional and program-related factors associated with the variance in contraceptive practice among currently married women aged 15–44."
"In the second phase of this study, to be conducted in 1984, the BKKBN will evaluate a new model of health and family planning service delivery in terms of (a) increased contraceptive practice among currently married women aged 15–44 and (b) greater utilization of primary health care facilities among people in the lower socioeconomic strata."

3 Evaluation Study of Different Training Methodologies.
"The Training Division of the Family Planning and Maternal and Child Health Project in Nepal will conduct an evaluation study in 1983 to determine the difference in field-worker performance between those workers who received a classroom-based training program and those who received a field-based training program."

4 Survey Study of the Family Planning Knowledge, Attitudes, and Practices of Community Influentials in Sri Lanka.
"In 1983, the Ministry of Plan Implementation will conduct a survey in 100 randomly selected villages to determine the family planning knowledge, attitudes, and practices of community influentials."

2 Comments on the Example

Note that, in each of these objectives, the investigator promises to undertake **a specific behaviorial activity:** to conduct, to evaluate, to survey, or to determine. The situation or context of the research is also specified. **Who** will conduct the study **where** and **when** is stated. Finally, and most important, the key variables of interest are indicated—social, economic, institutional, and program factors; contraceptive practice; utilization of primary health care facilities; field-worker performance; training; and so on. Each of these variables will need to be carefully defined in operational terms later, but at least as stated here in the immediate objectives, they serve to indicate the focus of the proposed research.

C Hypotheses

A hypothesis is a statement about an expected relationship between two or more variables that permits empirical testing. While ultimate objectives give the anticipated contributions arising from a study and immediate objectives state (in behavioral terms) what will be done or measured, hypotheses specify the expected relationship among the variables.

Study hypotheses serve to direct and guide the research. They indicate the major independent and dependent variables of interest. They suggest the type of data that must be collected and the type of analysis that must be conducted in order to measure the relationship.

A single hypothesis might state that variable A is **associated** with variable B, or that variable A **causes** variable B. Sometimes a hypothesis will specify that, **under conditions X, Y, and Z,** variable A is associated with or causes variable B. A well-written hypothesis focuses the attention of the researcher on specific variables.

When writing hypothesis statements, it is important to keep in mind the distinction between **independent** and **dependent variables.** An independent variable **causes** or **determines** or **influences** the dependent variable. The basic relationship between these two types of variables can be represented as in Figure 3.1. This model shows a **direct relationship.** In other words, whenever the independent variable changes, there is a change in the dependent variable. The dependent variable **depends** on the independent variable.

In most family planning operations research studies, it is unusual to find a direct relationship of this type. Almost always, the relationship is **indirect.** The independent variable acts upon the dependent variable through **intervening variables.** These intervening variables serve either to increase or to decrease the effect the independent variable has on the dependent variable. Intervening variables are sometimes referred to as **test** or **control** variables. Many, if not most, family planning research studies are based on a model that includes intervening variables (see Figure 3.2).

For example, in an effort to increase contraceptive acceptance (the dependent variable), a family planning program might initiate a large information, education, and communication (IEC) program (the independent variable). Alone and by itself, an IEC program cannot increase contraceptive use. The program must act on and operate through a set of intervening variables, which in turn cause or determine or influence contraceptive use. There may be many of these intervening variables, but the most likely ones that would be influenced by an IEC program would be people's knowledge about family planning, their attitudes toward the use of birth control methods, and their beliefs about family size. A possible research model for an evaluation study of the effects of an IEC program is shown in Figure 3.3.

In family planning operations research, the hypotheses of most importance usually are those that consider program activities as the independent variable. The family planning input, such as an IEC effort, is the independent variable. The objective of the research is to determine the extent to which the independent variable affects the dependent variable. The dependent variable is usually the central concern of a research proposal's problem statement. The intervening variables are important because they tend either to increase or to decrease the strength of the

FIGURE 3.1
A direct relationship between independent and dependent variables

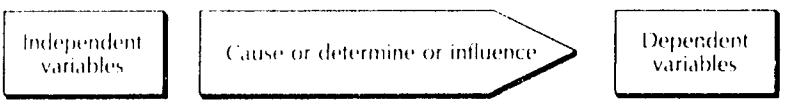


FIGURE 3.2
Intervening variables

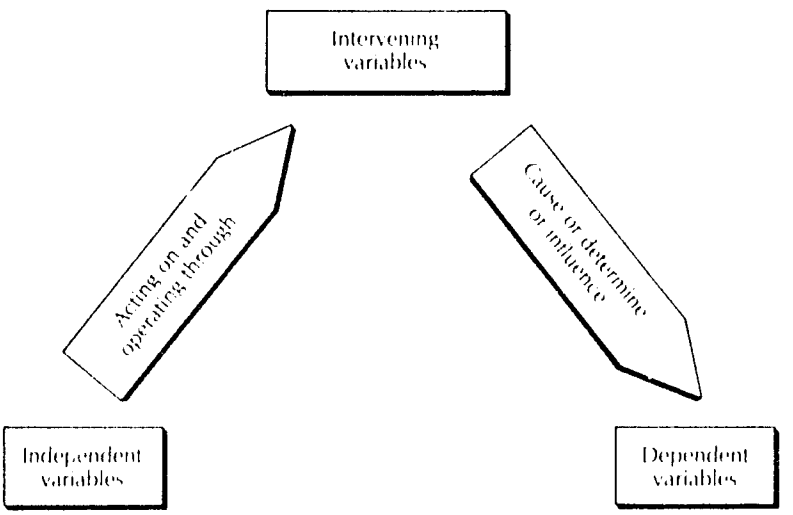
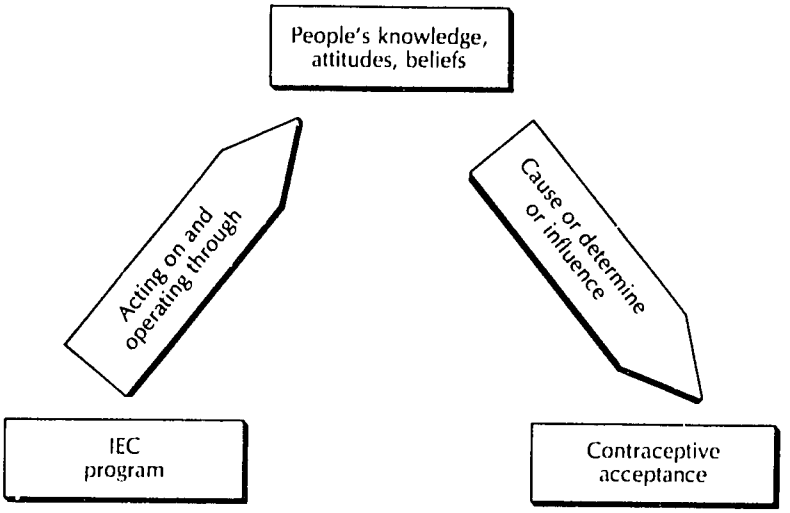


FIGURE 3.3
The effect of intervening variables in an IEC program



relationship between the program (the independent variable) and the outcome effect (the dependent variable).

Most research studies will examine many independent variables and many intervening variables but only a few dependent variables. In writing study hypotheses, always think in terms of the expected relationship between variables. Think first about the central problem your study will address (the dependent variable). Next consider what factor or factors (the independent variables) might cause, determine, or influence the problem situation. Finally, ask yourself if the relationship between the independent and dependent variables is direct or indirect through a set of intervening variables.

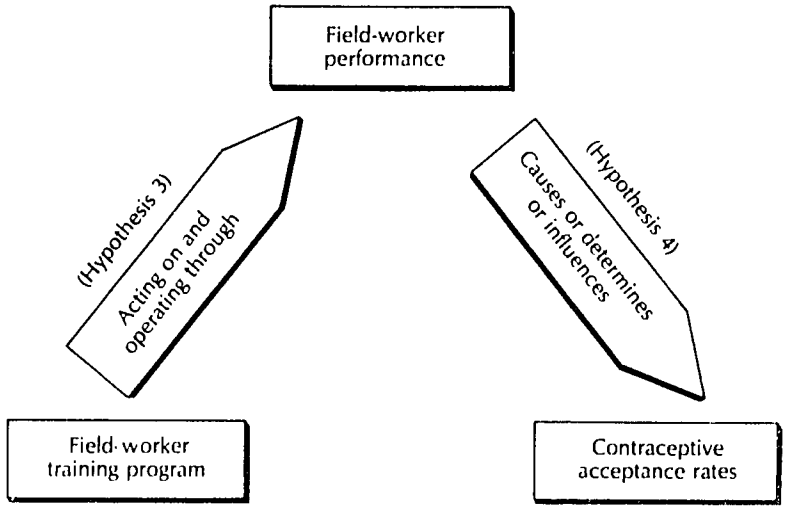
1 Examples of Hypotheses

1. Villages with many modern characteristics will have a higher rate of contraceptive practice than villages with few modern characteristics.
2. Contraceptive practice will be higher in villages where the family planning field-worker is viewed as credible than in villages where the field-worker is not viewed as credible.
3. The work performance of health and family planning staff who have received a five-week, field-based training course will be higher than the work performance of similar staff who have received a four-week, classroom-based training course.
4. Higher levels of field-worker performance will lead to higher rates of contraceptive acceptance among eligible couples.
5. An information and education program using radio, television, and printed material will increase people's knowledge about family planning and lead toward more favorable attitudes about the use of contraceptive methods.
6. People with higher levels of family planning knowledge and more favorable attitudes toward the use of contraceptive methods will be more likely to be users of contraceptive methods than people who are not as knowledgeable about family planning and do not have as favorable attitudes toward the use of contraceptive methods.

2 Comments on the Examples

Note that in each hypothesis there is a statement of an expected relationship between two or more variables. In the first hypothesis, the expected relationship is between modern characteristics of villages (the independent variable) and contraceptive practice (the dependent variable). In the second hypothesis, the relationship is between the credibility of field-workers (the independent variable) and contraceptive practice (the dependent variable). The third and fourth hypotheses go together. In the third hypothesis, the relationship is between a five-week, field-based training program (the independent variable) and field-worker performance (the intervening but also dependent variable). In the fourth hypothesis, field-worker performance acts as an independent variable and contraceptive acceptance rates become the dependent variable. Taken together, the model for these two hypotheses would look like Figure 3.4. Similarly, hypotheses five and six go together. Family planning knowledge and attitudes are the intervening but also dependent variables in hypothesis five, but they become the independent variables in hypothesis six.

FIGURE 3.4
A model for two hypotheses about the effect of a field-worker training program on contraceptive acceptance rates



To summarize, all research studies should include a statement of (1) ultimate objectives, (2) immediate objectives, and (3) hypotheses (except in the case of purely exploratory or descriptive studies). Each of these statements serves a different purpose:

1 **Ultimate objectives** state the **anticipated (hoped for) contributions** of the study.
2 **Immediate objectives** state **what will be done** in the study.
3 **Hypotheses** state the **expected relationship** between two or more variables.

What To Do—Writing ultimate objectives, immediate objectives, and hypotheses

1 List the major variables of your study under the headings: Independent Variables and Dependent Variables. Intervening variables can be listed under both headings. The dependent variable(s) should relate directly to your problem statement.
2 Review the list of dependent variables, and then write a statement of ultimate objectives that relates to them. Ask yourself the following question: "If I knew the factors that caused, determined, or influenced the dependent variable(s), how would this knowledge help policymakers, program administrators, or others?" How will the results from the study help improve service delivery, improve training programs, assist in the design of educational materials? In other words, what are the **anticipated contributions** of the study?
3 Now write the immediate objectives for your study. Focus your attention on the specific actions that you will perform. Exactly what do you plan to do? What variables do you plan to measure? Write behavioral objectives that answer the following questions:
   a Who will do?
   b How much of what?
   c To whom?
   d When?
   e Where?
   f For what purpose?
4 If your study includes an intervention, use the same questions listed above to describe the intervention.
5 Finally, write the hypotheses that your study will test. Review once again your list of independent and dependent variables. How are these variables related? Are there special conditions that must be present before they are related? Write hypothesis statements in **positive,** not negative, terms. Write a hypothesis for each major relationship that you expect to test in your study.

Operational Definitions

4

After formulating the study objectives and hypotheses, the next step in the research process is to define operationally the key variables and terms of the study. Operational definitions serve two very essential purposes: (1) They establish the rules and procedures the research investigator will use to measure variables. (2) They provide unambiguous meaning to terms that otherwise can be interpreted in different ways. Every research proposal must include operational definitions of major variables and terms.

A Operational Definitions of Variables

Suppose that a dependent variable of a study is family planning knowledge. Before this variable can be measured, it is necessary first to establish the operational procedures that specify how the measurement will be made and at the same time define what the researcher means by the three words **family planning knowledge.** The variable must be defined in terms of events that are **observable by the senses.** The observable events serve as an **indicator** of the variable, family planning knowledge. Alone and by itself, knowledge is not observable by the senses. You cannot touch knowledge, or see knowledge, or smell it, taste it, or hear it. What is needed is an observable event that indicates knowledge. Usually, such an indicator is based on a series of questions. For example, we might ask a respondent, "Have you ever heard about contraceptive pills?" "Can you tell me how often a woman should take the contraceptive pill in order to avoid pregnancy?" "Have you ever heard about the IUD?" "Can you tell me where a woman can get an IUD?" Each of these questions serves to indicate whether the respondent knows about certain aspects of family planning. Asking a question and hearing a response is an observable event that can be measured.

A research study might ask 20 knowledge questions. Each time a respondent gives an answer that indicates knowledge, the researcher could record a score of 1 (one). Every time an answer is given that does not indicate knowledge, the researcher could give a score of 0 (zero). For each respondent, the researcher could add the total number of correct answers to the 20 questions and create a family planning knowledge score. This score would range from 0 correct answers to 20 correct answers. Persons with a score of 0 would be said to have no family planning knowledge. Persons with a score of 20 would be said to have high knowledge. What this procedure does is to define operationally the variable, family planning knowledge. In the research proposal, the variable might be defined operationally as follows:

Family Planning Knowledge = The number of correct answers a respondent gives to 20 questions on family planning.

Of course, this is not the only way the variable could be defined operationally. We might wish, for example, to establish **categories** of family planning knowledge, distinguishing between those respondents who have **high** family planning knowledge, **medium** knowledge, **low** knowledge, and **no** knowledge. Each of these levels is a category of the variable, and each requires an operational rule that tells us how to assign any given respondent to a category. One way of operationally defining the categories might be the following:

High Knowledge = Correct responses to 15 or more of the 20 questions.
Medium Knowledge = Correct responses to at least 8 but not more than 14 of the 20 questions.
Low Knowledge = Correct responses to at least 1 but not more than 7 of the questions.
No Knowledge = No correct answers to any of the 20 questions.

Notice that the four categories of the variable are **mutually exclusive** (they do not overlap). By the operational rules established, a person cannot be placed in the category "High Knowledge" **and at the same time** be placed in the "Medium," or "Low," or "No" category. The categories are also **totally inclusive.** There are only four categories. There is no fifth or sixth category that a respondent might possibly fit into.

In some instances, we may not want to be quite so specific in defining before data collection the exact ways in which a variable will be split into categories. Sometimes it is preferable to determine the category "cutting points" after data have been collected and the response distribution for the variable can be examined. As a general rule, it is best to have approximately an equal number of respondents in each category. Thus, in the above example, each of the four categories of the variable family planning knowledge should have approximately 25 percent of the respondents in the study population. If it is necessary to examine the response distribution of a variable before the procedures for establishing categories can be determined, then in the operational definition section of a study proposal the category names can be specified, but with a note indicating that each category will comprise approximately equal numbers of respondents.

All variables must have at least two or more categories, otherwise they are not variables but constants. Whenever you are operationally defining a variable, it is always better to divide the variable into many categories instead of just a few. In the examples given above, the variable family planning knowledge ranges from 0 to 20. That range gives a total of 21 separate categories. Subsequently, in the second example, we collapsed these 21 categories into just four—high, medium, low, and no knowledge. If we wanted to, we could go even further and collapse the four categories into just two:

Knowledge of Family Planning = A correct response to 1 or more of the 20 questions.
No Knowledge of Family Planning = No correct answers to any of the 20 questions.

If you start with many categories, it is always easy to collapse these down to just a few. But do not make the mistake of starting with just a few categories, because subsequently you cannot expand them into many. Collapsing the categories of a variable from many to just a few usually is done after data collection has been completed and the frequency distribution of the variable can be examined.

1 Examples of Operationally Defined Variables

Family Planning Acceptance = The initiation of use of any modern method of contraception.
Field-Worker Performance = The number of job-related tasks a field-worker undertakes out of 20 defined tasks.
Family Planning Attitude = The number of positive responses to 10 questions indicating approval or disapproval of the use of contraceptive methods.

B Operational Definitions of Terms

Recall that a hypothesis is a statement about an expected relationship between two or more variables. Just as it is necessary to define variables operationally, so too is it necessary to define the terms that indicate the nature of the relationship between the variables. For example, in many hypothesis statements, you will find such terms as:

more than	greater than
less than	larger than
higher than	bigger than
lower than	smaller than

You will also see in hypothesis statements such words as:

safer	significant
acceptable	expanded
improved	increased

Each of these terms can have a variety of meanings, and thus each requires an operational definition for the research proposal. The basic problem with such terms as **more than** or **increased** is that they denote a comparison but do not indicate the standard of comparison. We need to know **how much more** and **increased by how much.**

Suppose a study has the following simple hypothesis:

> A five-week, field-based training program will increase the family planning knowledge of health and family planning workers who have taken the program.

The training program is the independent variable. In the hypothesis, this variable is already defined, at least partially, as five weeks long and field based. Family planning knowledge is the dependent variable. We already have defined this variable as the number of correct responses to 20 questions. What is left is to define the word **increase.** If we do not define this word, it will be impossible for us to know when the hypothesis has been proved or disproved. In other words, we need a standard of comparison that will tell us **increase by how much.** One way to define **increase** might be the following:

Increase = Among trainees, a mean family planning knowledge score on the posttraining test that is greater than and significantly different from the mean family planning knowledge score of a control group of field-workers who did not receive the training program.

Note that this operational definition not only tells us the meaning of **increase** but also gives us the procedures that will be used to measure the increase. The mean family planning knowledge score of trainees (in an experimental group) will be compared against the mean family planning knowledge score of a control group. The hypothesis will be accepted only if the trainees' mean score is greater than and significantly different from the mean score of the control group. Finally, to be absolutely clear, we also should define the word **significantly:**

Significantly = A probability equal to or greater than .95 that the trainees mean score is higher than the control group mean score.

To summarize, operational definitions establish the rules and procedures an investigator plans to use to measure and give meaning to variables and terms. The definition is in terms of indicators that are **observable events.** We must be able to **ask** a question, **hear** a response, **see** a behavior, **record** an action, **measure** an attribute. The definition establishes **categories** for variables. The categories must be **mutually exclusive** and **totally inclusive.** Operational definitions also establish the **standard of comparison** the investigator will use to either accept or reject a hypothesis.

1 Examples of Operationally Defined Terms

Family Planning User = Any currently married woman aged 15–49 and any currently married male aged 15 and older who says he or she has used a modern method of contraception at least once during the last month.

Modern Method of Contraception = Any one of the following pregnancy prevention methods: pills, IUD, female sterilization, male sterilization, depo provera, condom, diaphragm.

Improved Worker Performance = Any government family planning field-worker who recruits more new family planning acceptors, holds more village meetings, or distributes more oral rehydration solution packets than he or she did six months earlier.

Modern Village = Any village that has three or more of the following facilities: electricity, a government health clinic, a paved road within half a mile, a primary school, a bank, a post office, irrigation for 50 percent or more of the farmland.

What To Do—Writing Operational Definitions

1 Write an operational definition for each variable on your list of independent and dependent variables.
2 Write an operational definition for each term used to connect the variables.
3 For each definition you write, ask yourself:
   a Are the rules and procedures for measuring the variable clear?
   b Have mutually exclusive and totally inclusive categories for the variable been established?

Study Design

5

A Reliability and Validity

The design of a study is the investigator's plan of action for answering the research question(s). The objective in selecting a study design is to minimize possible errors by maximizing the **reliability** and **validity** of the data.

Reliability refers to the consistency, stability, or dependability of the data. Whenever an investigator measures a variable of interest, he or she wants to be sure that the measurement provides dependable and consistent results. A reliable measurement is one that if repeated a second time will give the same results as it did the first time. If the results are different, then the measurement is unreliable.

Validity refers to data that are not only reliable but also true and accurate. Put another way, validity is the extent to which a measurement does what it is supposed to do. If a measurement is valid, it is also reliable. But if it is reliable, it may or may not be valid. For example, suppose an investigator asks a respondent, "How old are you?" The respondent replies, "I am 40 years old." The investigator then asks a second question to check on the consistency or dependability of the age measurement: "In what month and year were you born?" The respondent replies, "In September 1941." If it is now February 1982, the investigator calculates that indeed the respondent is 40 years old.

In this example, two questions have been asked. Each is designed to determine how old the respondent is, and each question gives the same results. The results are consistent, stable, and dependable; they are reliable. But suppose later on the investigator happens to see the respondent's birth certificate, which shows that the true date of birth was September 1938. The investigator then concludes that, although the first two questions gave reliable results, they did not give valid or true results.

B Validity Threats

In the selection of a research design, one criterion to use is the extent to which the design controls for threats to validity. In other words, a researcher wants a design that will give true and accurate information—a design that will avoid factors that might invalidate the study. While there are many such factors, some of the more common threats to validity are explained below.

1 History

Sometimes events occur during the life of a project that tend either to increase or to decrease the expected outcomes of the project. History effects refer to these events. The events are not part of the project; they are not planned or anticipated events. They just happen, and they produce an effect that influences the study results

For example, suppose that a group of field-workers have been given a special training course that is designed to improve their performance. Six months after the course, an evaluation is conducted, and the results indicate that there has been no improvement in field-worker performance compared with the pretraining period. The evaluator might conclude that the training program was a failure. But a further examination of the situation reveals that soon after

the training program was completed, all field-workers decided to go on strike for three weeks in an effort to raise their salaries. The strike represents a history threat to the validity of the evaluation. Obviously, the strike was not a planned or anticipated project event. It just happened, and the effect was to influence the performance of the field-workers. The question then arises, was the training program **really** a failure?

If we could somehow control for the effect of the strike, would an evaluation of the training program still reveal that the field-workers' performance did not improve? In other words, what is the true and valid effect of the training program? The ability to answer this question depends on the study design the evaluator has selected.

In a second example, suppose that an evaluation is conducted to determine the effect on people's knowledge and attitudes of a family planning motivational campaign. A precampaign survey might be conducted, and the results then compared with the postcampaign survey results. The comparison might indicate substantial increases in people's knowledge about family planning and more favorable attitudes toward birth control. Very likely the evaluator would conclude that the motivational campaign was a success and its objectives were fulfilled. But suppose that during the campaign the president of the country made a major radio and TV address to the nation on the subject of rapid population growth and the need for family planning. The question arises, was it the motivational campaign that produced the increases in knowledge and changes in attitudes, or was it the president's address, or possibly was it both? The president's address represents a history effect. Unless there is some way in the study design to control for this effect, it is impossible to determine the true and valid results of the motivational campaign. In other words, the separate effect of the campaign cannot be distinguished from the history effect of the president's speech.

2 Selection

A very common threat to validity occurs whenever the people selected for the control group differ greatly from the people selected for the experimental group. For example, suppose that a special family planning incentive program is offered to people in one district (the experimental group) but not to people in another district (the control group). After a year, the incentive program is evaluated in terms of the family planning prevalence rate in both districts. The evaluation shows that there has been a significant increase in the experimental district over the control district. In this situation, the evaluator might conclude that the incentive program was a success: It increased the prevalence rate. But upon further checking, the evaluator discovers that there are large and marked differences in the characteristics of the people in the two districts. In the experimental district, the people tend to be younger, better educated, more urban, and somewhat better off economically than the people in the control district. Because of these differences, the people in the experimental district are probably more likely than the people in the control district to accept family planning, whether or not there is an incentive program. Once again, the question of validity arises. What is the true effect of a family planning incentive program? If both the experimental and control districts had been similar in the beginning, it might be possible to provide some answer to this question. But because there were initial differences, the effect of the incentive program cannot be distinguished separately from the difference due to a selection effect.

3 Testing

Whenever a pretest is given, it tends to have an effect on the posttest. People who are given a pretest (trainees, for example) are likely to remember some of the questions and some of the errors they made when they take the posttest. They are likely to do somewhat better on the posttest than they did on the pretest. The difference or better performance on the posttest might have nothing to do with a program (such as a training course) but instead be due to the effect of the pretest. Sometimes in knowledge, attitude, and practice (KAP) surveys, the same group of people are repeatedly asked the same questions each year for several years. After awhile, they begin to remember the correct answers to many of the questions asked, such as, "Can you name three modern methods of contraception?" Their responses may be due to their familiarity with the KAP survey and not to their increased knowledge of family planning as a result of an educational program. Whenever a test is given repeatedly to the same group of people, there is every likelihood of a testing effect or threat to validity.

4 Instrumentation

Whenever a measurement instrument (such as a questionnaire) is changed between the pretest and the posttest, this change is likely to result in an effect that is independent of any effect due to a program or project. For example, if knowledge of family planning is measured on the pretest by the question "Have you ever heard of family planning?" and then on the posttest by the question "Can you name three modern methods of contraception?" there will very likely be a difference in the percentage of people with "knowledge." The difference may be due to the change in the questionnaire instrument. Similarly, an instrumentation effect may be involved when interviewers become more experienced. An experienced interviewer may obtain more complete information from a respondent than an inexperienced interviewer. The additional information obtained may be due to the fact that the interviewer has become more skilled in asking questions or observing events, and not due to the effect of a program intervention. Changes in the way information is collected or changes in the instrument of data collection can result in a threat to validity.

5 Maturation

Over time, people (and things) change. In long training programs, it is not unusual for trainees to become tired, hungry, bored, or discouraged. In longitudinal studies, particularly those that extend over a period of years, it is not unusual to find that respondents become more experienced, more knowledgeable, wiser, sometimes more resistant, and, of course, older. In other words, people mature over time. The maturation process can produce changes that are independent of the changes a program intervention is designed to produce. If a posttest is given to a group of trainees who are tired, hungry, or bored, we might expect lower scores, on the average, than would occur if the trainees had not been tired, hungry, or bored. If these lower scores are then compared with the pretest scores, an evaluator might find there was very little difference and thus conclude that the training program was not successful. But is this really the case, or are the results just due to a maturation effect, another threat to validity?

6 Mortality

In cohort studies (also called panel studies) where the same group of people are followed over time, there is almost always some dropout or loss of cases. It is often impossible to locate for a second survey all the same people who were interviewed in the first survey. Some cases will be lost to follow-up. Mortality effects refer to these losses. If the people who cannot be contacted are very different from those who can be contacted, there are likely to be great differences between the results of the first and second surveys. These differences may be due to the loss of cases rather than to the effect of a program intervention.

Each of the threats to validity (history, selection, testing, instrumentation, maturation, and mortality) must be considered carefully when designing a research study. Unless controlled (accounted for), each validity threat represents a possible explanation for why a program produced or did not produce an effect. In other words, it is always necessary to distinguish an effect due to a program intervention from an effect due to a validity threat.

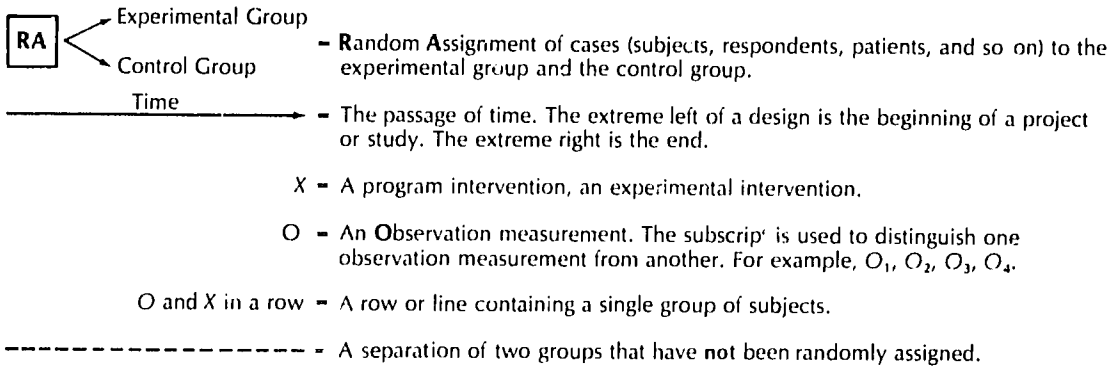
C Types of Study Designs

In the following pages, several different research designs are presented and discussed. They range from nonexperimental to quasi-experimental to true experimental designs. The designs presented here are certainly not the only ones that exist, but they are some of the more frequently used designs in family planning operations research. In describing each of these, it is helpful to use a common notation developed by Campbell and Stanley.* This notation is explained in Figure 5.1.

In this design, all subjects are randomly assigned (RA) from a single population to the experimental group and the control group. Both the experimental and control groups receive an initial measurement observation (the pretests O_1 and O_3). The experimental group then

* See reference in bibliography.

FIGURE 5.1
Notation for study designs



Example of a Study Design

		Time		
RA	Experimental Group	O_1	X	O_2
	Control Group	O_3		O_4

receives the program intervention (X), but the control group does not receive this intervention. Finally, a second set of measurement observations are made (O_2 and O_4). We would expect that, since the experimental group received a special program intervention (X), O_2 would be greater than O_4. Also, since both the experimental and control cases were randomly assigned, we would expect that O_1 would be equal to O_3 on such key variables as age, sex, parity, and education.

The design shown above is a **true experimental design.** Because it is one of the strongest in terms of controlling for threats to validity, it acts as a point of reference for the designs discussed below.

1 Nonexperimental Designs

There are several nonexperimental designs commonly used by family planning researchers. These designs are most appropriate for collecting descriptive information or for doing small case studies of a particular situation. They are not recommended for evaluation studies that attempt to determine the effect of a program intervention, but they may be useful in diagnostic studies designed to determine the reasons why a problem exists.

a Posttest-Only Design

	Time
Experimental Group	$X \quad O_1$

In this design, a program intervention (X) has been introduced, and sometime after its introduction, a measurement observation (O_1) is made. Since there is no control group, there is no possibility of comparing the O_1 measurement with any other measurement. All that the O_1 measurement can do is provide descriptive information. The threats to validity of history, maturation, selection, and mortality are not controlled and therefore are factors to consider. This design would be most appropriate if an investigator wanted to know, for example, the demographic characteristics of IUD acceptors or sterilization acceptors. The X would represent the treatment (the IUD insertion or the sterilization operation). The O_1 would consist of information collected from a group of acceptors. While numerous problems are associated with the posttest-only design, particularly if a comparative analysis is desired, much useful program information can be obtained if multivariate data analysis techniques are used (see Chapter 8, Section B.2.f).

b Pretest–Posttest Design

	Time
Experimental Group	$O_1 \; X \; O_2$

In this design, there is still no control group, but at least there is an earlier measurement observation (O_1) that allows the research investigator to examine changes over time. But it is important to be aware that the pretest-posttest design is subject to several threats to validity—history, testing, maturation, and instrumentation.

c Static-Group Comparison

	Time →
Experimental Group	X O_1
Control Group	O_2

Unlike the other two designs, this one adds a control or comparison group. The experimental group receives a program intervention (X) followed by a measurement observation (O_1). This measurement observation is then compared against a second observation (O_2) from a control group that did not receive the program intervention. Note, however, that the two groups are separated by a broken line (- - - - -), which indicates that a random process was not used to create the two groups. This design might be employed if patients at one clinic were used as a control group to make comparisons with patients at another clinic where a special program intervention had taken place. The primary problem with this design is the threats to validity of selection and mortality. Initially, the two groups might differ greatly on the basis of such variables as age, parity, and education. Also, it is always possible that many of the patients in the experimental group who received the program intervention (X) may subsequently have been lost to follow-up for the O_1 observation. Thus, the differences between O_1 and O_2 might be due to selection or mortality effects and not to the program intervention (X).

2 *Experimental Designs*

a Pretest-Posttest Control Group Design

	Time →
RA — Experimental Group	O_1 X O_2
Control Group	O_3 O_4

This design has already been presented earlier and discussed briefly. This is a true experimental design. The primary distinguishing feature is the **random assignment** of cases to the experimental and the control group. Random assignment helps to assure that the two groups will be equal before the program intervention is introduced. In other words, if the randomization has been done properly, there will be a known probability of difference between the groups. Because the two groups are essentially equal at the beginning, any difference observed later on between O_2 and O_4 can be attributed to the effect of the program intervention (X). This is one of the strongest designs in terms of controlling (or accounting for) threats to validity. At the same time, it is a design that is difficult to implement in field situations. Often it is not possible to assign cases randomly to experimental and control groups. For ethical or programmatic reasons, it is sometimes impossible to deny program interventions to one group while offering them to another.

b Posttest-Only Control Group Design

	Time →
RA — Experimental Group	X O_1
Control Group	O_2

This is also a true experimental design and similar to the one discussed above except there is no baseline (pretest) measurement observation. Since cases have been assigned randomly to the experimental and the control group, these groups are assumed to be similar before the program intervention. This design allows the investigator to measure the effect of a program intervention on the experimental group by comparing that group with the controls. But this design does not allow an investigator to determine the extent (or magnitude) of change within the experimental group because a baseline pretest measurement was not taken.

3 *Quasi-Experimental Designs*

In many field research situations, it is simply impossible or certainly very costly and difficult to meet the random assignment criteria of a true experimental design. At the same time, researchers want to avoid the problems of threats to validity associated with nonexperimental designs. A

reasonable compromise often can be made by selecting a quasi-experimental design. These designs do not have the restrictions of random assignment. At the same time, they tend to control for many threats to validity.

a Time Series Design

$$\text{Experimental Group} \quad \overset{\text{Time}}{\longrightarrow} \quad O_1\ O_2\ O_3\ X\ O_4\ O_5\ O_6$$

The time series design is similar to the nonexperimental pretest-posttest design except that it has the advantage of **repeated measurement observations** before and after the program intervention (X). If, for example, we find that there is no difference between O_1, O_2, and O_3, but then a sudden increase occurs between O_3 and O_4, which is subsequently maintained in O_5 and O_6, we can conclude with some degree of confidence that the sudden increase was probably due to the effect of the program (X). This type of increase might be shown in graph form as in Figure 5.2.

Suppose in this example the program intervention (X) is a large-scale educational campaign designed to increase people's knowledge and understanding of family planning. Before the educational program, observations O_1, O_2, and O_3 reveal that only 30 percent of adults in a particular geographic area know and understand the meaning of family planning. After the program, observations O_4, O_5, and O_6 reveal that 50 percent of the adults know and understand family planning. The most likely explanation for the change is the educational campaign. But another explanation might be a history effect or threat to validity that suddenly happened between O_3 and O_4.

Although history and instrumentation effects are possible, the times series design helps the researcher avoid making mistaken or false conclusions. Study Figure 5.3. Note that there is a steady and constant percentage increase over time. The increase is no greater between O_3 and O_4, when the program intervention (X) took place, than it is at any other time. Obviously, in this situation the investigator could not claim that the program made a difference. This is the

FIGURE 5.2
A sudden increase in a measured variable after a program intervention

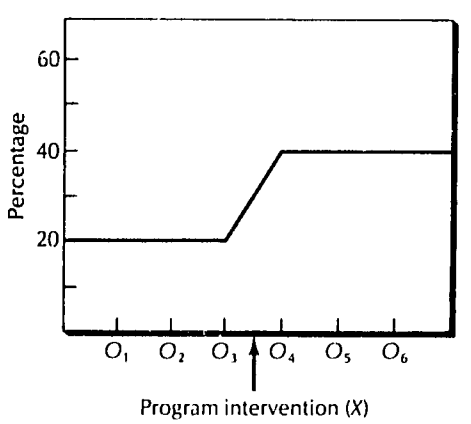
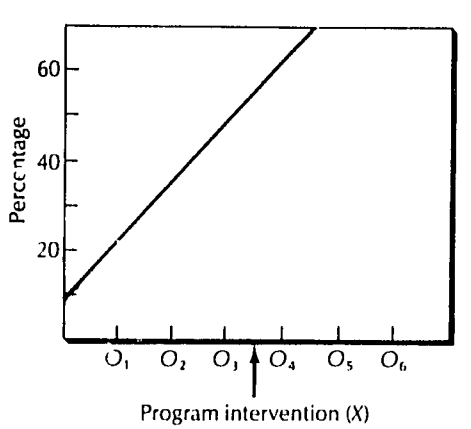


FIGURE 5.3
Steady increase in a measured variable before and after a program intervention



advantage of the time series design. The multiple measurement observations give us a trend. If the investigator had used the pretest-posttest design and compared only O_3 against O_4, he or she might have mistakenly concluded that the program had made a difference.

Now examine Figure 5.4. The trend in this example is regular and consistent increases and decreases in the percentage. The program intervention did not seem to make a difference in this trend. But, once again, if the investigator had used a pretest-posttest design and compared only O_3 against O_4, he or she might mistakenly have concluded that the program had had an impact.

The time series design is a good one to use when you have access to regularly collected information, such as monthly service statistics. You can plot the information on a graph and then note the point at which a special program intervention occurred. If there is a sudden change in the plotted line that is unusual and not similar to the trend before or after a program intervention, you can be fairly sure that the sudden change was due to the program.

Consider Figure 5.5. In this xample, suppose that each month sterilization acceptors account for 40 percent of all new acceptors. Measurement observations O_1, O_2, and O_3 all confirm the 40 percent figure. Then a major information and education campaign is conducted. Suddenly, sterilization acceptors account for 70 percent of all new acceptors, as shown through measurement observation O_4. But soon after the campaign, sterilization acceptors drop back to the 40 percent level. The program intervention clearly had an impact, but the impact was not long lasting. Once again, had the investigator used a pretest-posttest design and compared only O_3 against O_4, he or she would mistakenly have concluded that the information and education campaign had had an enormous impact. But the investigator would have missed the important point that the impact was only temporary.

To summarize, although the time series design does not include a control group and does not control for history and possibly instrumentation threats to validity, it does allow for a more detailed analysis of data and program impact than the pretest-posttest design because it gives information on trends before and after a program intervention. It is a particularly appropriate design to use when it is possible for a researcher to make multiple measurement observations before and after a program intervention.

FIGURE 5.4
Regular increases and decreases in a measured variable before and after a program intervention

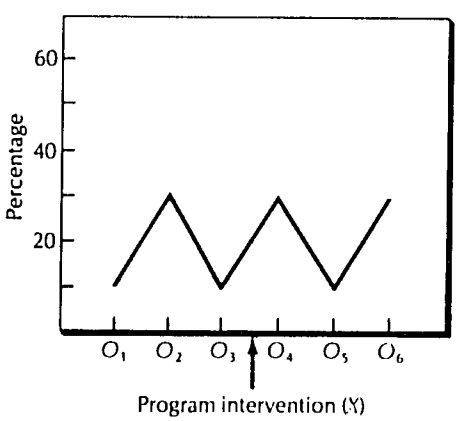
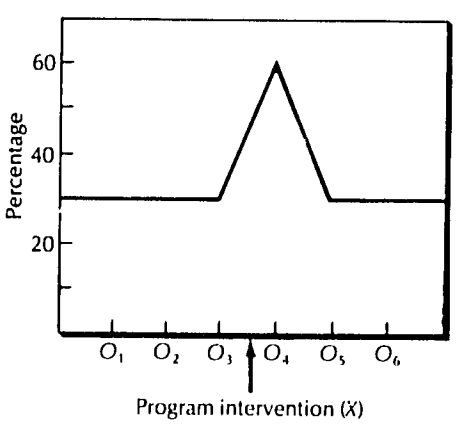


FIGURE 5.5
Temporary impact of a program intervention



b Nonequivalent Control Group Design

	Time →
Experimental Group	O_1 X O_2
Nonequivalent Control Group	O_3 O_4

In field research settings, it usually is possible to compare an experimental group with a similar but not necessarily equivalent group. For example, suppose you want to study the effect of a refresher training course given to family planning field-workers. Groups of 20 field-workers might be called into the training center for a three-week retraining program. These trainees could be given a pretest (O_1), followed by the three-week training course (X) and then the posttest (O_2). For purposes of comparison, a similar group of field-workers who had not yet been given the refresher training course could be administered the same pretest at the same time (O_3) and the same posttest (O_4). In this situation, we could use the two pretests (O_1 and O_3) to assess the extent to which the two groups of field-workers were truly similar. Then we would compare the two posttests (O_2 and O_4). We would expect that O_2 would be greater than O_4 because of the effect of the refresher training (X).

The nonequivalent control group design is a particularly good one to use in evaluations of training programs. It is also a good design to use when a program intervention is introduced into one area (say, one district) and you want to compare the program effects in that district against a similar, but not necessarily equivalent, neighboring district. In using this design, you should give careful attention to the analysis and comparison of O_1 against O_3 (the pretest baselines). In particular, you should look for selection effects or major differences between the experimental and the nonequivalent control group that might help explain differences (or lack of differences) in the posttest O_2 and O_4 comparisons.

c Separate Sample Pretest–Posttest Design

		Time →
(RA) →	Pretest Group	O_1 X
(RA) →	Posttest Group	X O_2

This design is very frequently used in family planning research. Essentially, it involves doing a baseline pretest (O_1) with a randomly selected sample from a study population. Subsequently, a program intervention (X) is introduced, and then a posttest measurement (O_2) is made using a second randomly selected sample from the same study population. This design is a good one to use to avoid testing effects; however, it does not control for history, maturation, mortality, or possibly instrumentation effects.

The different designs discussed above are just a few of the many available. Whichever design you select, it is important to keep a few basic guidelines in mind:

1. Whenever possible, try to create experimental and control groups by assigning cases **randomly** from a single population study group.
2. When random assignment is not possible, try to find a comparison group that is as **nearly equivalent** to the experimental group as possible.
3. When neither a randomly assigned control group nor a similar comparison group is available, try to use a **time series design** that can provide information on trends before and after a program intervention (X).
4. If a time series design cannot be used, as a minimum and before a program starts, try to obtain **baseline (pretest) information** that can be compared against postprogram information (a pretest–posttest design).
5. If baseline (pretest) information is unavailable, be aware that you will be limited in the type of analysis you can conduct. You should consider using **multivariate analytic techniques.**
6. Always keep in mind the issue of validity. Are the measurements **true?** Do they do what they are supposed to do? Are there possible **threats to validity** (history, selection, testing, maturation, mortality, or instrumentation) that might explain the results?

What To Do—Designing a Study

1 The objectives of a research study determine to a large extent the type of research design required. Review the objectives of your study, and consider the following questions:

   a Are you planning to do an evaluation of a program? If your answer is yes, consider using an experimental or a quasi-experimental design.

   b Are you interested only in describing a particular event? If the event has already taken place, you can use a posttest-only design ($X\ O_1$) and consider multivariate analysis. If the event has not yet taken place, you might be able to use a pretest–posttest design ($O_1\ X\ O_2$), a time series design ($O_1\ O_2\ O_3\ X\ O_4\ O_5\ O_6$), or a nonequivalent control group

$$\begin{array}{ccc} O_1 & X & O_2 \\ \hline O_3 & & O_4 \end{array}$$

   c Will you be conducting an experiment to determine the effect of a particular approach on family planning service delivery? If your answer is yes, then you should seek to obtain a randomly assigned control group or at least a comparison group that is similar to the experimental group.

2 On the basis of your answers to the above questions, diagram with *O*s and *X*s the type of design you think is most appropriate for your research study. Remember that many variations are possible to the designs presented earlier.

3 Review the list of threats to validity, and determine which ones are controlled and which are not.

6 Sampling

A Selecting Area and Population

In family planning operations research, the information that is collected refers most commonly to individuals. For instance, the effect of offering a new type of contraceptive may be studied by making it available to one group of eligible couples but not to another. Later, in postproject interviews, information can be obtained from the wives in both groups about their contraceptive practice. If the assignment of couples to experimental and control groups is randomized, such a design will permit a test of the hypothesis that availability of the new contraceptive method affects contraceptive practice. In such a study, the availability and use of the new contraceptive method are viewed as characteristics of the interviewed wives. Each wife may be referred to as an experimental **subject.** Since the wives are interviewed, they may also be called **respondents.** More generally, they may be referred to simply as **cases,** a useful term, since not all studies are experiments or surveys.

Cases need not be individual persons; they may be aggregations or groups of persons. The relationship between the introduction of a new contraceptive and the practice of contraception described above could be studied at the community level. If the cases to be studied are entire villages (instead of individuals), the new contraceptive could be defined as "available" if it is offered by a clinic within a specified distance of the village; contraceptive practice could be measured in terms of the prevalence rate for the entire village (the percentage of married women of reproductive age using the method). Cases may also be organizational units. For instance, a study of clinics could be designed to determine the effect of offering a new contraceptive method on the method mix of new acceptors.

Usually the type of case to be studied in a particular research project is clear from the nature of the project and its objectives. In a study to determine the effects of an intervention on knowledge and attitudes of target couples, it is clear that the appropriate cases to be studied will be target couples. In a study of the effects of training on field-workers' knowledge and skills, the appropriate cases for study are field-workers. In a study intended to compare high and low performing clinics, the appropriate cases are clinics. Some studies may involve more than one type of case. A study to test the effects of a new training program for field-workers may be designed to permit observation of its effects on both the trainees and the eligible couples served by them. Such a design would call for analysis at two levels (at the trainee level and at the couple level), which would require separate data collection and processing for each level.

B Selecting Cases

Once the types of cases to be studied have been determined, the next step is to decide how the individual cases shall be selected. For this purpose, it is useful to distinguish between the **total population** and the **sample** from the population that will actually be observed. The population is composed of **elements,** each of which is a potential case. Some studies involve such small populations that all the elements can and should be included in the sample. For example, in a study of central management of a national program, there may be a total of only 20 or 30 top managers available for the study. In this case, a 100 percent sample (the entire population of managers) would be most desirable. Similarly, if there are only five mobile teams in a program, a study of the operations of such teams should probably involve all five. Sometimes, 100 percent sampling is appropriate even for studies involving large populations. For instance, in a study of

the effects of an intervention on the age distribution of acceptors, it may be possible to analyze from clinic records the data on all acceptors, especially if such data are computerized. Similarly, it may be possible to use the entire population of clinics as the sample to study clinic performance.

Usually, however, research problems involve so many persons or places or service points that it is necessary to select a sample of less than 100 percent to represent them. For a survey of family planning acceptors at program clinics during 1982, the population of concern would be the total number of persons registered as acceptors at program clinics during 1982. If this number is large—say, 800,000—it would be excessively costly, time consuming, and unnecessary to interview them all. The usual practice is to select and interview a relatively small sample of acceptors—a representative subset of the total population. If the sample is truly representative of the population and consists of several hundred cases or more, the information obtained from it will be very similar to the information that would be obtained from a study of the entire population.

There are two basic types of samples: (1) **probability samples,** and (2) **non-probability samples.**

1 Probability Sample

A probability sample is more likely than a non-probability sample to be truly representative of the larger population from which it is drawn and is therefore usually considered preferable to a non-probability sample. The essence of probability sampling is that each element of the larger population (i.e., each couple, each field-worker, each clinic, etc.) has a **known probability of being selected.** (If each element has an equal probability of selection, the sample is known as a **self-weighting** sample, and the findings from analysis of the sample data can be taken as representative of the general population.) There are several ways of drawing probability samples, the most common being the following.

a Simple Random Sampling

In this sampling procedure, each element of the larger population is assigned a unique number, and a table of random numbers or a lottery technique is used to select elements, one at a time, until the desired sample size is reached. This procedure can be quite tedious in drawing large samples.

b Systematic Sampling

This is a modification of simple random sampling, which is ordinarily less time consuming and easier to implement. The estimated number of elements in the larger population is divided by the desired sample size, yielding a **sampling interval** (let us call it *n*), and the sample is drawn by listing the population elements in an arbitrary order and selecting every *n*th case, starting with a randomly selected number between one and *n*.

While simple random sampling requires a list of all population elements (called the **sampling frame**) in advance, systematic sampling permits concurrent listing and sampling. This can greatly reduce costs. Of course, it is preferable that the sampling frame already be available, since systematic sampling can then be used to obtain precisely the sample size desired. If the size of the population is not known in advance, the size of the sample also cannot be known in advance. For instance, if you want to draw a sample of 20 couples from a village, you must know the total number of couples in the village population. If that number is not known, it is necessary to estimate the number. If you estimate that there are 200 couples, the sampling interval should be 10, meaning that every tenth case is interviewed during the listing process. If your estimate is too high, you will draw a smaller sample than you intended; if it is too low, your sample will be larger than intended. It is always better to have a larger sample than required instead of a smaller one than needed to do the research. Also, it is a good idea to budget for a somewhat larger sample than you need.

c Stratified Sampling

Stratification may be used in conjunction with either simple random sampling or systematic sampling to ensure the desired representation of specific subgroups. For example, if it is especially important that the age distribution of respondents in the sample be exactly the same as in the larger population, the list of elements in the larger population can be arranged by age category—that is, stratified by age—and then samples for each age strata drawn separately. If the sample fraction is the same for all age strata, the age distribution will be the same in the sample as in

the larger population. It is often useful, however, to apply different sampling fractions in different strata. For instance, in a study of family planning acceptors, it may be desirable to sample equal numbers of acceptors of pills, condoms, and sterilization, even though these methods account for different proportions of all acceptors. This could be accomplished by stratifying by contraceptive method and then applying different sampling fractions to the various strata.

d Cluster Sampling

Sampling can be simplified greatly be selecting **clusters** of elements (using simple random, systematic, or stratified sampling techniques) and then proceeding to study all elements in each of the sampled clusters. Most commonly the clusters are geographic units (e.g., provinces, districts, villages) or organizational units, such as clinics, Mothers' Clubs, or training groups. For instance, in order to obtain a sample of rural married women of reproductive age (MWRA), first clusters of rural villages could be sampled and then all MWRA interviewed in the sampled clusters. The chief drawbacks of this technique are: (1) it complicates the mathematics of calculating the sampling variance and therefore of testing for statistical significance; and (2) it usually requires larger sample sizes for achieving statistical significance at any given level.

e Multistage Sampling

In drawing samples from very large and diverse populations, the technique most commonly used is multistage sampling. The population is first divided into naturally occurring clusters, and then a sample of these clusters is drawn. This is called the first stage of sampling. The second stage consists of sampling individual elements (respondents, cases) from within each of the sampled clusters. Sometimes more than two stages are required. For example, provinces might be selected at the first stage, districts within the provinces at the second stage, villages within the districts at the third stage, and finally, individual respondents within the villages at the fourth stage. At each stage, simple random, systematic, or stratified techniques might be used.

f Probability Proportional to Size (PPS)

This procedure is a variation on multistage sampling in which the probability of selecting a cluster is proportional to its size and an equal number of elements is sampled within each cluster. The PPS method of sampling is useful when the clusters vary greatly in size. In such situations, PPS often reduces both the sampling variance as well as the data collection costs.

2 Non-probability Sample

Non-probability sampling (also known as **convenience sampling**) refers to the selection of a sample that is not based on known probabilities. Non-probability sampling may be either **accidental** (through the selection of whatever cases happen to be available), or it may be **purposive** (through the planned selection of specific types of cases). Non-probability sampling procedures are not valid for obtaining a sample that is truly representative of a larger population group. Almost always, non-probability samples tend to overselect some population elements and underselect others. In the absence of known probabilities of selection, there is no precise way to adjust for such distortions.

Despite these drawbacks, there are instances where trying to obtain a truly representative probability sample may be too difficult or too expensive. In such instances, non-probability sampling provides the only viable means for collecting data. For instance, if a government ministry has been attempting to integrate family planning into the activities of its field-workers, it might be worthwhile to mount a study of the performance of such workers. If the research budget is too small to permit extensive travel to visit workers in many parts of the country, it may be decided to sample field-workers in areas relatively nearby even though their experience may differ somewhat from that of the field-workers in more remote areas.

Sometimes it is impossible or too expensive to find enough cases of a particular type by using probability sampling methods. For instance, if a study is intended to examine the effect of a particular radio program on family planning perceptions and attitudes of listeners, it would be costly to draw a probability sample of listeners. This would require, first, a complete enumeration of the broadcast area (or of a representative sample of subareas) to determine who has heard the program, and then, selection of a representative sample of listeners. If the funds available do not permit such a costly sampling scheme, the only alternative may be to go from house to house, interviewing listeners as they are found, until the desired sample size has been

reached. The findings from such a sample would have to be treated with caution, since its composition would probably be affected by variations in case of access. Rural listeners and slum dwellers, for example, might be underrepresented because they are relatively harder to find and interview than urban and more middle-class people.

If non-probability sampling is necessary, efforts should be made to minimize bias. If any of the characteristics of the general population are known, the sample can be drawn in such a way as to ensure proportionate representation of the various categories of respondents. For example, if it is known that one-third of the population lives in urban areas and two-thirds in rural areas, the sample cases can be selected purposively from urban and rural areas in the same proportions. If the age distribution is known and is thought to be related to the key variables of the study, the sampling instructions might specify precise numbers to be interviewed in each age category. Such an approach to purposive sampling is called **quota sampling,** since quotas for specific subgroups are established at the outset.

If the main purpose of a study is to maximize the variety of qualitative information from a group of respondents (for instance, to get a comprehensive view of how a particular strategy is implemented, to learn about a broad range of problems to be solved, or to get a wide variety of proposed solutions to existing problems), non-probability sampling may be the **best** approach. It permits the researcher deliberately to select a diversity of respondents and to seek persons who are most likely to provide articulate answers to complex questions. Typically, anthropologists use purposive samples when they obtain information about communities from selected **key informants**—persons who are knowledgeable, articulate, and drawn from the various subgroups in the community.

C Sample Size Determination

The size of the sample is determined primarily by two things: the availability of **resources,** which sets the upper limit of the sample size; and the requirements of the proposed **plan of analysis,** which sets the lower limit. For instance, the availability of trained interviewers to do a survey in a specified period of time may be insufficient to permit more than a certain number of interviews, in which case the sample size cannot exceed this number. On the other hand, the sample size must be large enough (1) to allow for reliable analysis of cross-tabulations; (2) to provide for desired levels of accuracy in estimates of proportions; and (3) to test for the significance of differences between proportions. If the resources available for a study are inadequate to obtain a sample of sufficient size, then the researcher must either find additional resources or revise the plans for data analysis.

If you intend to analyze **cross-tabulations** of variables, you should keep two points in mind when determining sample size. First, each category of an independent variable included in a cross-tabulation should contain at least 50 cases, since percentages calculated on the basis of fewer than 50 cases tend to be unreliable. The minimum sample size required to ensure at least 50 cases in each category of the independent variable is obtained by dividing 50 by the proportion (P_s) of total cases you expect in the smallest category of the variable:

$$m = \frac{50}{P_s}$$

The second point to keep in mind is that the **expected** number of cases in each cell of a table should be at least five. The sample required to ensure a minimum expected cell size of five is obtained by dividing five by the product of the proportions falling in the smallest categories of the variables in the cross-tabulation table.

To illustrate these two points, consider the sample size required to analyze a cross-tabulation of contraceptive practice and educational attainment shown in Table 6.1.

Education would be the independent variable in Table 6.1. Hence, in order to have a sample large enough to ensure that at least 50 cases will fall into the smallest category of the variable educational attainment (that is, into the category "college," which accounts for only 5 percent of the total number of cases), we need to divide 50 by 5 percent ($n = 50/.05 = 1{,}000$ cases required). To find the minimum sample size needed to ensure an expected cell frequency of at least 5, we must divide 5 by the product of the proportions falling in the smallest categories of the two variables (.05 for the category "college" and .20 for the category "past user"):

$$n = \frac{5}{(.05)(.20)} = 500 \text{ cases required}$$

TABLE 6.1 Cross-tabulation of contraceptive practice and educational attainment

	Educational attainment			
Contraceptive practice	College (5 percent)	Secondary school (15 percent)	Primary school (40 percent)	No schooling (40 percent)
Current user (50 percent)				
Past user (20 percent)				
Never user (30 percent)				

Since the sample size required must meet **both criteria** (50 cases in each variable category and 5 expected cases in each cell), the larger estimate of $n = 1{,}000$ should be selected instead of the smaller estimate of $n = 500$.

To calculate the minimum sample size required for accuracy in **estimating proportions,** you need first to answer the following questions:

1. What are **reasonable estimates** of key proportions to be measured in the study? For example, if you are studying contraceptive prevalence, you should try to guess what prevalence rate you will obtain. If you cannot guess what it will be, the safest procedure is to assume it is .50 (50 percent), which maximizes the expected variance and therefore indicates a sample size that is sure to be large enough.
2. What **degree of accuracy** do you want to have in your study? How far can you allow the sample estimates of key proportions to deviate from the true proportions in the population as a whole? For instance, if you find that the sample estimate of the prevalence rate is .50, do you want to be confident that this finding is accurate within 1 percent or 5 percent (usually referred to as the .01 and .05 level, respectively)? If you seek a high degree of accuracy (such as .01), your sample will need to be much larger than if you seek a relatively low degree of accuracy (such as .05).
3. What **confidence level** do you want to use? How confident do you want to be that the sample estimate is as accurate as you wish? Customarily, the 95 percent confidence level is specified.
4. What is the **size of the population** that the sample is supposed to represent? If it is greater than 10,000, the precise magnitude is not likely to be very important. But if it is less than 10,000, the required sample size may be smaller.
5. If you are seeking to measure the difference between the two subgroups with regard to a proportion, what is the **minimum difference** you expect to find statistically significant? For instance, if you are comparing the contraceptive prevalence of an experimental group and a control group, and you find a difference of only 5 percentage points, do you expect a difference this small to be statistically significant? The smaller the difference you expect to be significant, the larger your subsample sizes will have to be.

On the basis of your answers to these five questions, you can calculate the sample size needed to measure a given proportion with a given degree of accuracy at a given level of statistical significance by using a simple formula, provided that the total population size is greater than 10,000:

$$n = \frac{z^2pq}{d^2}$$

Where:

n = the desired sample size (when population is greater than 10,000).
z = the standard normal deviate, usually set at 1.96 (or more simply at 2.0), which corresponds to the 95 percent confidence level.
p = the proportion in the target population estimated to have a particular characteristic. If there is no reasonable estimate, then use 50 percent (.50).
$q = 1.0 - p$.
d = degree of accuracy desired, usually set at .05 or occasionally at .02.

For example, if the proportion of a target population with a certain characteristic is .50, the z statistic is 1.96, and we desire accuracy at the .05 level, then the sample size is:

$$n = \frac{(1.96)^2(.50)(.50)}{(.05)^2}$$
$$= 384$$

If we use the more convenient 2.0 for the z statistic, then the sample size is:

$$n = \frac{(2)^2(.50)(.50)}{(.05)^2}$$
$$= 400$$

Note that the numerator in this case is 1.0. This means that when you assume the proportion is .50 and set a 95 percent confidence level by using z equal to 2.0, then the formula for sample size is simply:

$$n = \frac{1.0}{d^2}$$

If N (the entire population) is less than 10,000, the required sample size will be smaller. In such cases, calculate a final sample estimate (n_f) by using the following formula:

$$n_f = \frac{n}{1 + \left(\frac{n}{N}\right)}$$

Where:

n_f = the desired sample size (when population is **less than** 10,000).
n = the desired sample size (when the population is **more than** 10,000).
N = the estimate of the population size.

For example, if n were found to be 400 and if the population size were estimated at 1,000, then n_f would be calculated as follows:

$$n_f = \frac{400}{1 + \left(\frac{400}{1{,}000}\right)}$$
$$= \frac{400}{1.4}$$
$$= 286$$

If you wish to test differences (d') between two subsamples regarding a proportion and can assume an equal number of cases ($n_1 = n_2 = n$) in the two subsamples, the formula for n' is very similar to the one given above:

$$n' = \frac{2z^2pq}{(d')^2}$$

For example, suppose you want to compare an experimental group against a control group with regard to the proportion of persons using contraception. If you expect p (the proportion using) to be .40 and you wish to conclude that an observed difference of .10 or more is significant at the .05 level, then the size of the sample can be calculated as follows:

$$n' = \frac{2(1.96)^2(.40)(.60)}{(.10)^2} = 184$$

Thus, we need 184 experimental subjects and another 184 control subjects. Note that, as with the earlier example, the formula can be simplified to the following if p is set at .50 and z at 2.0:

$$n' = \frac{2}{(d')^2}$$

If this formula were applied to the example above, the sample size would be calculated as:

$$n' = \frac{2}{(.10)^2}$$
$$= 200$$

If you cannot assume equal subsample sizes or if the sizes of the populations from which the subsamples are drawn are much less than 10,000, the procedure for estimating sample sizes is more complicated. In such cases, or in any others where you are confused by sampling, you should consult a statistician.

What To Do—Sampling

1 Decide first whether you want to draw a sample and, if so, whether it should be a probability sample or a non-probability sample. In making this decision, take into account the objectives of the study, the extent to which you want the findings to be representative of a larger population, and such resource factors as cost, time, and personnel.

2 Calculate the size of the sample required for your study. You may use the formulas given above, but you should also seek the assistance of a statistician if possible. The statistician will need to know your estimates of the proportions to be tested, the degree of accuracy you seek or magnitudes of differences you wish to test, the confidence levels you wish to use, and the approximate size of the population from which the sample will be drawn.

3 Bear in mind the following basic principles:

a A larger sample will yield more accurate results but will be more costly than a smaller sample.

b A probability sample will provide quantitative data more representative of a larger population than will a non-probability sample, but a non-probability sample can be designed in such a way as to maximize insightful qualitative data from relatively small samples.

c If your proposed analysis calls for studying particular subgroups of your sample, your sample size will need to be expanded accordingly. For example, to study characteristics of a group of acceptors, you may need a sample of only 400 acceptors, but if you want to extend the analysis to the acceptors of particular methods, the subsample sizes will be too small to yield significant findings unless the total sample size is increased. Even if statistical significance is not considered very important, there should be at least 50 cases in the smallest subgroup to be studied if you want to obtain even moderately reliable percentages.

Data Collection

7

There are many different ways to collect data. The approach selected depends on the study objectives, the study design, and the availability of time, money, and personnel. An important consideration in deciding on the best way to collect data is whether the study is intended to produce relatively precise **quantitative** findings or to produce **qualitative** descriptive information. Most operations research studies are concerned with the quantitative measurement of program operations, but many also are (or should be) concerned with detailed qualitative information on processes (e.g., how a project is actually implemented in the field or how couples decide to use one contraceptive method in preference to the others available) or perceptions (e.g., potential acceptors' understanding of how a method is used and its possible effects or perceived problems with existing program services). Often study objectives call for both quantitative and qualitative information, which may imply that more than one data collection method should be used.

A Quantitative Data

The most common means used for collecting quantitative data on people is the structured interview. However, such information can sometimes be collected in other ways: from self-administered questionnaires, service statistics, or such secondary sources as the census, vital records, and other existing records and reports. If the sample is composed of geographic or organizational units (such as villages, districts, clinics, hospitals) rather than people, quantitative data usually can be obtained from service statistics and secondary sources. If information is not available from such sources, it may be necessary to obtain data on geographic or organizational units by interviewing people who are members of these units or knowledgeable about them.

1 Structured Interviews

Studies that obtain data by interviewing people are called **surveys.** If the people interviewed are a representative sample of a larger population, such studies are called **sample surveys.** If the sample is large enough to permit statistical analysis, it is customary to employ structured interviews rather than unstructured ones, since the former lend themselves better to quantitative analysis and the latter create serious data processing difficulties, particularly if the sample is large. A structured interview is one that employs a standard questionnaire (or interview schedule) to ensure that all respondents are asked exactly the same set of questions in the same sequence. The exact wording of each question is specified in advance, and the interviewer merely reads each question to the respondent. In designing a questionnaire and then using interviewers to administer it, there are several points to remember:

a **Use simple language** that will be understood easily by the respondents.

b Whenever possible, **precode the responses** to the questions so that the information can be transferred easily to computer cards or otherwise tabulated. This requires more effort when designing the interview schedule, but the savings of time and effort during the processing and analysis more than compensates for it.

c **Try to avoid embarrassing or painful questions.** If it is necessary to ask a sensitive question, word the question as tactfully as possible and avoid asking it near the beginning of the interview, when the respondent may not yet be completely relaxed.

d **Avoid asking for more than one item of information in a single question.** For instance, do not ask "Do you **and** your husband want another child?" If the respondent and her husband disagree about having another child, an answer of either "yes" or "no" will be impossible to interpret accurately. A response of "yes" could mean "I want another child," "My husband wants a child," or "Both of us want a child." It would be better to ask two separate questions: "Do you want another child?" "Does your husband want another child?"

e **Watch out for ambiguous wording** of questions. For instance, if **family planning** has already been defined, a question such as "Which method or methods of family planning do you use?" may seem clear enough. But if the respondents are wives, they may not mention use of vasectomy or condoms, since these are male methods. The question should be "Which method or methods of family planning do you **and** your husband use?"

f **Do not overload your interview schedule** with items that are not essential for your study. Keep it as short as possible to avoid tiring the respondent and to simplify the data processing and analysis.

g On the other hand, be sure to **include all questions necessary to provide sufficient information** on the variables you want to study.

h **Start with the easier questions,** moving on to the ones that are more sensitive or difficult to answer only after the respondent has had an opportunity to become accustomed to the interview situation. Respondents are likely to be somewhat tense or even suspicious at first, and a major task of the interviewer during the first few minutes is to establish rapport—to place the respondent at ease. This task is facilitated if the initial questions are not likely to cause embarrassment or be difficult to answer.

i **Ask all respondents each question in exactly the same way.** If the interview is to be conducted in more than one language, prepare full written translations into all major languages and instruct your interviewers to use those translations word for word. Do not permit free translations, except for languages with too few respondents to justify the cost of preparing written translations. To ensure comparability of wording among the various written translations, have them **back-translated** into the original language to verify that the meaning is retained. The back translation should be done by persons who are not familiar with the original wording of the questionnaire.

j **Pretest the questionnaire in an actual field situation.** There are several principles of pretesting that should be noted:

1 The pretest need not involve large numbers of respondents; 30–50 is often enough if the respondents are sampled (purposively) in such a way as to ensure that the expected heterogeneity of the study sample is reflected in the pretest sample. This means making sure that the pretest includes old and young, urban and rural, less educated and more educated, contraceptors and noncontraceptors, and so on.

2 Be prepared to conduct more than one pretest. If a pretest results in major revisions, it is a good idea to conduct a second pretest to be sure the revisions are satisfactory.

3 Pretesting should be completed before the training of interviewers. Often it is possible to use field supervisors to do the pretest. This gives them a clearer understanding of study objectives and better prepares them to assist in the training of interviewers.

4 The main purpose of the pretest is to ensure that the respondents are able to understand the questions and answer them usefully. Hence, it is not enough simply to interview the pretest respondents; each interview should be followed by a **debriefing,** in which the interviewer asks about the respondent's understanding of questions that are thought likely to be misunderstood or that appear to have caused difficulty during the interview.

k **Provide complete training for all interviewers.** The training should be designed to familiarize the interviewers with the intent and meaning of the questions, let them role play interview situations, and give them experience in actually conducting interviews in the field under supervision.

l If a prospective respondent is not available during the interviewer's first visit, **an appointment should be made for a call-back visit.** It is common to require at least two call-backs before dismissing a sample case as unavailable for interview.

- **m** If the sample is small, it may be necessary to find substitutes for cases that cannot be located. If so, the **interviewers must be given clear instructions for obtaining substitutes** that will ensure either a random sampling of substitutes or selection of substitutes who are similar to the cases originally selected. However, if the sample is large enough to tolerate some loss of cases, it is usually better not to draw substitutes.
- **n** Every effort should be made to **isolate the respondent** during the interview. If others are present, the respondent's answers may be influenced. For example, a respondent whose husband does not know she is using oral contraceptives may give the interviewer one type of answer to a question if she is interviewed alone, but quite another type of answer if her husband is present.
- **o** **Check all completed interview schedules** for errors, omissions, and discrepancies as soon after interviewing as possible. Respondents should be revisited to correct errors that cannot be otherwise resolved. It is best if the interviewer can check the questionnaire immediately after the interview so that the respondent can be consulted. After the interviewer has checked and corrected the questionnaire, it should be rechecked by the field supervisor. This checking process is known as **field editing.**

2 Service Statistics

Every national family planning organization generates a large volume of service statistics. The quality of these statistics varies from country to country and even within countries. Thus, they should be used with caution. Service statistics often help a researcher define the parameters of the problem he or she wants to study. In some cases, they can be used to compare the results of a particular study with nationwide figures. Often in operations research projects it is necessary to design supplementary forms to provide data that are not available from the regular service statistics.

3 Self-Administered Questionnaires

If money, personnel, and time allow, interviews are generally preferable to self-administered questionnaires. For most surveys in developing countries, self-administered questionnaires are difficult if not impossible since many respondents are not well enough educated to complete questionnaires themselves. There are other problems with self-administered questionnaires, even among educated respondents: Instructions or questions are more likely to be misunderstood without an interviewer to help explain them; portions of the questionnaire are more likely to be left blank; it is difficult to incorporate many conditional sequences of questions (e.g., "If the response to question 12 is "yes," go to question 13; if not, skip to question 18") without causing confusion.

Self-administered questionnaires are likely to be useful in situations where literate respondents are already gathered together in a setting where they can write—for example, in a classroom or an office. Self-administered questionnaires can be especially useful in evaluating population education programs or training programs.

Sometimes blank questionnaires are mailed out to respondents, who are requested to complete them and send them back. This has the virtue of being very inexpensive, but it has all the drawbacks noted above, plus the added problem of high nonresponse rates. It is common for mailed questionnaires to elicit less than a 50 percent return, even after one or two reminders. This drawback affects the representativeness of the sample and may render the validity of quantitative findings so questionable that they are of very little use

4 Secondary Data Sources

Information from recent censuses, vital records, other official statistics, and even previous surveys often can be incorporated with data collected specially for a study to enrich the analysis. However, such information rarely would be sufficient by itself.

B Qualitative Data

The data collection techniques most appropriate for studies whose objectives call for descriptive, qualitative data tend to be different from those most appropriate for quantitative analysis.

Unstructured interviews, focus group discussions, direct observation of operations, and content analysis are among the techniques most commonly employed.

1 Unstructured Interviews

The chief drawback of structured interviews is that the responses obtained tend to be superficial. An alternative approach to interviewing, which permits greater depth of meaning, is to seek detailed, open-ended responses to questions. Such interviews are often called **in-depth interviews.** Instead of reading formal questions from a structured interview schedule, the interviewer has an outline of topics or a set of general questions to serve as a guide to the kind of information required. Details that are not brought out initially are sought through follow-up questions, called **probes.**

The chief drawbacks of unstructured interviews are (1) the responses, being unstandardized, are difficult to quantify; (2) the interviews require highly skilled and experienced interviewers; and (3) the analysis is exceedingly time consuming. The shortage of qualified interviewers and analysts, and the high cost of conducting and processing such interviews usually means a small sample size must be used (sometimes as few as 20–30 respondents). In-depth interviews are usually most useful for exploratory studies that seek to clarify concepts or generate hypotheses prior to developing questionnaires for quantitative surveys. They also are useful for generating supplementary, explanatory data to augment the relatively superficial findings from larger surveys.

2 Focus Group Discussions

A way of reducing the amount of time and personnel required for conducting and analyzing in-depth interviews and yet getting detailed qualitative information from a relatively large number of respondents is to bring the respondents together in groups. The interviewer (or facilitator) follows much the same procedure as in the unstructured interview situation, using a general discussion guide and eliciting details through probes. Participants are usually sampled purposively to reflect population variations that are of particular relevance to the research topic. For instance, cases sampled for focus group discussions on natural family planning (NFP) might involve purposively selected samples of husbands, wives, current users of each type of NFP (calendar rhythm, temperature method, cervical mucus method, and sympto-thermal method), past users no longer using NFP, and never users.

3 Direct Observation of Operations

This technique can generate either quantitative or qualitative data but tends to be used more for small-scale exploratory studies than for large-scale quantitative studies. This is because it usually requires relatively highly skilled observers and analysts and prolonged periods of observation, resulting in a high cost per unit of observation. This technique lends itself particularly well to observation of community responses to program efforts. It is the chief method of ethnographers, who specialize in community studies. It is also useful for organizational studies, such as observation of clinic operations, activities of field-workers, and administrative procedures.

4 Content Analysis of Written Materials

This is usually not employed as the sole method in a family planning operations research study, but it can serve as a useful adjunct to other types of data collection and analysis. For instance, documents related to a training curriculum may be content analyzed to determine just what knowledge and skills the training is supposed to develop. The content analysis can serve to guide the researcher in devising procedures to test the knowledge and skills of the trainees. IEC materials may be content analyzed to indicate whether messages are being overemphasized or underemphasized in relation to program needs. Research reports may be content analyzed to determine the present state of knowledge about a particular research topic so as to guide future research efforts.

What To Do—Data Collection

1 Review your study objectives and hypotheses as well as your list of independent and dependent variables. What types of information do you require? Which data gathering technique(s) would be most appropriate and feasible for gathering the desired information? Is some of the information already available from other sources?

2 If you intend to collect information through a survey, review the list of points under the heading Structured Interviews. Be sure your proposal discusses important steps, such as translation, pretesting, training of interviewers, and rules regarding call-backs and substitutions.

3 Make an outline of the data gathering instrument(s) (e.g., interview schedule, discussion guide, observation guide) that you intend to use. Give examples in the proposal of questions to be asked, especially those designed to elicit information regarding key variables found in the hypotheses. Be sure you incorporate measures (or at least descriptions of measures) of all the variables you intend to study. It is useful to list all the variables of the study and then under each variable record the question or questions that will be used in the questionnaire. For example:

Variable 1: Family planning practice
Q1: Have you ever used any method of family planning?
Q2: Are you currently using any method of family planning?
Variable 2: Education
Q1: What is the highest class in school you have completed?

4 Prepare a description of your data collection procedures and include it with the outline of the data gathering instrument(s) in the proposal.

C *Data Quality Checks*

There are several ways to check the quality of interview data:

1 Sometimes a researcher will deliberately ask two or more questions that give the same type of information. The first question might be asked at the beginning of the interview and the second at the end. The two questions are then examined for consistency of response. This is one way to check the reliability of the data.

2 For difficult questions, sensitive questions, or questions where the researchers want to be sure the information is correct, the interviewers can be instructed to probe. That is to say, the interviewer can repeat the question in a slightly different form or repeat the respondent's answer and then ask if the information is correct. For example, a woman might report that she has two sons and three daughters. The interviewer might then say, "You have a total of five children, two male and three female. Is that correct? Are there any other children whom you may have forgotten to tell me about?"

3 Field supervisors should be used to help the interviewers with difficult situations and to check that they are actually doing their work. Occasionally, interviewers complete their questionnaires in tea stalls! Some studies use a ratio of one supervisor for every five interviewers.

4 Most studies that use an interview procedure make an attempt to reinterview a certain percentage of the respondents. Depending upon the size of the sample, a general rule is to reinterview between 5 and 10 percent of the sample. The data from the first interview are then checked against the data from the second interview for consistency. This is another check on the reliability of the data. Obviously, if there are major inconsistencies, particularly on such questions as age, marital status, and parity, then clearly there is a problem somewhere. The problem might be with the questionnaire, with the interviewers, with the tabulation procedures, or in some other area.

5 Once the data have been collected and tabulated, it is possible to do statistical checks for errors or for consistency of response. For example, a frequency distribution of the parity of women may reveal that several women claim to have 18 or 19 living children. Since this is highly unlikely, the investigator is faced with the choice of either discarding these questionnaires, eliminating the information from the question on parity, or going back and reinterviewing the women who claim to have 18 or 19 children.

What To Do—Data Quality Checks

1 Describe the procedures you will use to check the quality of the data collected. Consider the following procedures:
   a Include repeat questions in your questionnaire that can be used to check for consistency of response.
   b Have supervisors monitor the work of the interviewers in the field.
   c Reinterview a percentage of the respondents, and look for inconsistencies.
   d Recode a percentage of the questionnaires to be sure that there are no coding errors.
   e Examine the frequency distribution on all variables to see if there are odd codes or items that are not logical.

D Confidentiality of Information

It is always important to maintain the confidentiality of the information collected from respondents. Unless absolutely required, do not collect information that is sensitive. Whenever possible, use code numbers instead of names. Assure the respondent that the information he or she gives will be kept confidential. Do not let other people use the information you have collected when there is a chance that the use of the information could be damaging to the respondent. You have an obligation to protect the confidentiality of your respondents. Also, if a respondent does not want to be interviewed, you have an obligation to respect his or her wishes. Use an Informed Consent Form to explain the basic nature of the study and seek the agreement of the respondent to be interviewed.

What To Do—Confidentiality of Information

1 Describe in detail how you plan to maintain the confidentiality of information collected.
2 Include a sample of the Informed Consent Form in your study proposal.

Tabulation and Analysis of Data

8

If the only purpose of your proposed study were to present qualitative, descriptive findings, analysis would consist simply of rearranging the raw information according to an outline and then writing a narrative report incorporating it. Most studies, however, especially in operations research, involve quantitative manipulation of the information collected, in which case the analysis phase is more complex. First, you need to decide whether the tabulation of data will be performed by hand or by machine. Second, you need to decide how the information can be converted into a form that will allow it to be processed efficiently. Third, you must specify the statistical manipulations to be performed. Finally, you need to present the important findings resulting from these manipulations in a report or series of reports.

A Preparing Tabulations

Data can be tabulated by hand, by a mechanical counter-sorter, or by electronic computer. The choice will depend on the facilities and personnel available, the nature of the data, the size of the sample, and the relative costs. In small studies that employ relatively simple statistical procedures, computer tabulation may be unnecessary. The data can be tabulated nearly as easily (and possibly at lower cost) by hand or counter-sorter, and calculators can be used to perform statistical operations, especially if computer facilities are relatively inaccessible. If a computer is readily available, you probably should use it, since it can produce tables and perform statistical operations much more easily (and usually with much less error) than is possible either by hand or by counter-sorter.

Once you have decided how tabulations will be generated, you must consider how the raw information will be transformed for tabulation and analysis. Regardless of whether tabulation will be done by hand or by machine, non-numerical data that are to be analyzed quantitatively must be converted into numerical codes. In addition, it usually must be transferred from its raw form (marks on interview schedules or questionnaires, transcripts of interviews, published census reports, etc.) to a form that is better suited for tabulation. For hand tabulation, it is often useful to transfer coded data to cards specially designed for hand sorting. For counter-sorter tabulation and for initial computer input, the information should be coded on standard (Hollerith) computer cards. For extensive computer processing, the data eventually should be transferred to magnetic tape or disks, since computer time is shorter (and therefore costs less) when data are read from tape or disks.

Data coding instructions must be prepared. For all variables studied, the coding instructions convert each of the possible response categories into a unique numerical code. If the number of categories for a particular variable (including, if relevant, codes for "noninformation," "not applicable," "don't know," and "others") is fewer than 10, numerical codes should be single-digit numerals. If the total possible number of categories is between 10 and 99, two-digit codes should be used instead. For some variables, it may be necessary to use three-digit, four-digit, or even larger code numbers.

Coded information should be spot-checked for errors. Some proportion of the coding should be repeated independently by a person other than the original coder and the results compared. If discrepancies are found, they should be checked and corrected by the coding supervisor. Most often, discrepancies will be due to problems with the coding instructions or poor training of the coders.

After the information has all been coded, it should be **edited** (even if the raw data have already been edited in the field). The coded data can be edited primarily by searching for the following types of errors:

1 **"Illegal" Codes**—values that are not specified in the coding instructions.
2 **Omissions**—for example, a failure by an interviewer to follow correctly the SKIP instructions in a questionnaire.
3 **Logical Inconsistencies**—for example, a respondent whose current age is less than her age at marriage.
4 **Improbabilities**—for example, a 25-year-old woman with 10 living children.

Once found, errors should be checked by reference to the original data forms and the necessary corrections made. Editing of coded data is most easily performed on the computer and is referred to as **machine editing** to distinguish it from **field editing.** Generally, data for computer processing are kept on cards until the computer editing is complete, since this simplifies the job of correcting the errors. Once the cards are "clean" (i.e., once the detected errors have all been corrected), the data should be transferred to magnetic tape or disks if tabulation and statistical analysis will be extensive.

What To Do—Coding and Editing

1 Decide whether the data should be tabulated by hand, counter-sorter, or computer. Take into account:
   a Availability of facilities.
   b Nature of the data (quantitative or qualitative).
   c Type of analysis to be performed.
   d Personnel available.
   e Costs.
   If you decide to use a computer, be sure to check the availability of existing (package) programs and programming assistance.
2 Decide which information is to be coded. Some information may be best left uncoded. Even a structured interview schedule may contain some open-ended items that lend themselves best to descriptive analysis. If so, indicate which items will be analyzed qualitatively.
3 Indicate in your research proposal where the coded data will be stored: on tape, on computer cards, on disks, or on cards designed for hand sorting.
4 Indicate in the proposal what provisions will be made for spot-checking and editing the coded data.

B Plan for Data Analysis

1 Attributes of the Data

One of the most important parts of a research study proposal is the plan for data analysis. The purpose of data analysis is to provide answers to the research questions being studied. The analysis breaks the data into parts. The data are **categorized, arranged,** and **summarized.** Usually, a researcher wants to find from the data one or more of the following attributes:

a Central Tendency or Characteristic of the Data

Suppose a researcher wants to know the **mean** parity of a group of women, the most frequently used method of contraception (the mode), or the **median** level of the income. The mean, the mode, and the median are measures of central tendency. The mean is simply an arithmetic average. It is the sum of individual scores or measures divided by the number of individuals. The mode is the most frequently occurring figure in a set of figures. The median is the midpoint measure of a group of measures. In a slightly different way, each of these statistical measures describes the typical characteristic or tendency of a group.

b Variance in the Data

Often a researcher is interested not only in the average characteristics of a group but also in the variation within the group. The researcher may want to know how individuals within a group differ from the average or central tendency of the group. For example, if the average age of a group of pill acceptors is 28, what is the **range** in ages of the group? What is the age of the youngest pill acceptor and the oldest? A common statistical measure of variation within a group is the **standard deviation.** This measure gives the average distance of individual measurement observations from the group mean.

c Differences within the Data

Most often researchers are interested in knowing whether the findings they observe from their data could have occurred by chance alone. For example, if 43 percent of an experimental group of 1,000 married women are currently using a method of contraception compared with only 21 percent of the control group, is this difference statistically significant or could it have occurred by chance alone? Probably the most commonly used statistical tests for differences are the chi-square (X^2) test and the t test for the difference between means. The use of these test measures is discussed later.

d Relationships within the Data

Most studies try to determine whether a relationship exists between one variable and another. For example, is contraceptive acceptance related to age? Is improved work performance related to training? Are radio family planning messages related to increased family planning knowledge? Is the level of education of a person related to contraceptive use? There are many different statistical measures of relationship, such as the product-moment correlation or the nonparametric correlation coefficient gamma. These and other measures are discussed later in this section.

2 *Types of Analytical Procedures*

In a proposal, the plan of analysis will probably describe a number of different analytical procedures the researcher plans to use in order to determine the central tendencies of the data, the variance, the differences, and the relationships among variables. Among the most commonly used types of **quantitative analytical procedures** in family planning operations research are:

- **a** Variable transformations
- **b** Univariate analysis
- **c** Time series analysis
- **d** Comparisons
- **e** Bivariate relationships
- **f** Multivariate analysis
- **g** Cost-effectiveness analysis
- **h** Use-effectiveness analysis
- **i** Fertility analysis

You need to indicate in your proposal which of these analytical approaches are most appropriate for meeting your study objectives. A prime consideration in making a selection is the extent to which the analytical procedure is appropriate to answer your research questions. Another important factor to consider is the extent to which you or someone associated with the data processing and analysis understands the terms and operations involved in the analytical procedure chosen and can interpret the results correctly.

The purpose of the following discussion is merely to indicate the options for data analysis available to you. This discussion cannot teach you how to analyze data if you are not already familiar with analytical procedures. All statistical tests and measures referred to below are widely used by social scientists. The reader who wishes to learn how to calculate and/or interpret them can find discussions and examples in almost any introductory textbook on statistical methods.

a Variable Transformations

It is often necessary to transform variables for subsequent analysis. For instance, instead of the respondent's age, the interview questionnaire may record only the month and year of birth. If age is a variable to be studied, it can be obtained simply by subtracting the month and year of

birth from the month and year of interview. Since this calculation can be done more easily and accurately by the computer than by hand, it should be done during the analysis phase of the research rather than during coding. This transformed variable might be transformed further still for certain kinds of additional analysis. For instance, if you wish to cross-tabulate age by other variables, it is preferable to limit the age distribution to relatively few age categories (usually 5- or 10-year categories) or even to dichotomize it (e.g., ages 15–29 and ages 30 or more). There are several ways to transform variables; among the most common are the following.

Recodes In recoding, category labels are changed. This technique is used for collapsing large numbers of variable categories into smaller numbers. For example, single years of age can be collapsed and transformed into age categories, such as 15–19, 20–24, 25–29.

Counts If you are collecting information on whether the respondents have ever used each of 10 family planning methods, you might want to count the number of methods ever used by each respondent, thereby generating a new variable that might be called "Number of Methods Ever Used." Counting is useful in studies involving more than one variable.

Scaling. A variety of scaling techniques are used in family planning operations research studies. Probably the most common is the Likert scaling technique, sometimes used for studying attitudes. Likert scales are constructed by presenting respondents with a series of statements that reflect different aspects of an underlying attitude in a variety of ways (e.g., pills cause cancer; pills are highly effective; pills are simple to use; pills cause discomfort). The respondent is then asked about each statement whether he or she agrees strongly, agrees moderately, or disagrees strongly. Code numbers 1–5 are assigned to each response, a code of 1 indicating a strongly favorable attitude and 5 a strongly unfavorable attitude (or vice versa). The scores for all attitude statements are added and then divided by the number of statements, yielding a scale value between 1.0 and 5.0.

Conditional Transformations. Where the nature of the transformation of one variable depends on the second variable, conditional transformations may be useful. For instance, suppose you have asked respondents three questions: (1) How many children do you have? (2) Do you want more? (3) If yes, how many more do you want? With the information from these three questions, you can then create a new variable called "Total Desired Number of Children." This can be done by adding the existing number of children (from question 1) to the additional number desired (from question 3). The number in question 3 would be added to the number from question 1 only if the answer to question 2 is "yes."

Other Mathematical Transformations. The calculation of age from the date of birth and the date of interview is an example of a mathematical transformation. Another example, from community-level analysis, is to divide the number of contraceptive users in a community by the number of eligible couples to obtain a community prevalence rate.

b Univariate Analysis

As the name implies, this form of analysis examines single variables. Often it is important to analyze the distribution of sample cases with regard to a single variable. For example, the age distribution of a sample population might be studied to be sure it is consistent with the age distribution of the larger population from which the sample is drawn. Univariate analysis of sample data can also be used to indicate the composition of the larger population with regard to important study variables for which such information is not otherwise available. Sometimes univariate analysis is sufficient to meet major study objectives. For instance, in a diagnostic survey of clients intended to indicate strengths and weaknesses of a particular project, univariate analysis of responses to questions about perceptions of and reactions to various project operations would constitute an important part of the analysis.

The simplest form of single-variable analysis is to count the number of cases in each category. However, the resulting count, called a **frequency distribution,** usually is not very interesting without additional statistical manipulations. The nature of the statistical manipulations that are possible depends on the type of variable or, more accurately, on the level of measurement. For family planning operations research, it is useful to distinguish three levels of variable measurement: nominal, ordinal, and interval.

Nominal Measurement. In nominal measurement, the categories of variables differ one from the other in name only. In other words, one category of a variable is not necessarily higher or lower, or greater or smaller than another category. It is just different in name, or it is the same. For example, the variable sex has two categories—male and female. A researcher can assign the number 1 to the category male and the number 2 to the category female. The only meaning these numbers have is to distinguish one category from the other. The researcher could just as well assign the number 2 to males and the number 1 to females. Or the number 1,598 could be assigned to males and the number 83 to females. It makes no difference. In another example, a researcher could assign numbers to different contraceptive methods: 1 = female sterilization; 2 = male sterilization; 3 = condom; 4 = IUD; 5 = DMPA; 6 = spermicides. These numbers simply distinguish the categories. A high number does not mean that the contraceptive is better or worse than one with a low number. Any selection of numbers could be used to classify the different contraceptive methods. The only important point is to be consistent and always use the same number for the same category.

Nominal measurement is the weakest of all measurements. A nominal variable is one that has mutually exclusive categories, but these categories differ in name only. The only meaning the numbers assigned to the categories have is to tell us that one thing belongs in the **same** category as another thing (if they both have the same number), or in a **different** category (if they have different numbers). Only very limited statistical manipulations are possible with nominal variables. You can calculate the mode (the most frequently occurring number). You can also calculate a percentage distribution. But you cannot calculate a mean. It makes no sense to speak of the "mean sex" or the "mean contraceptive method" or the "mean religion." Each of these is a nominal variable.

Ordinal Measurement. When there is an ordered relationship among the categories, the variable is said to be an **ordinal variable.** In other words, a category that is assigned the number 1 might be considered higher than a category assigned the number 2, which would be higher than the category assigned the number 3. For example, very often in family planning operations research, respondents are asked about their attitude toward the use of contraception. The response categories might be assigned numbers in the following manner:

1 = Approve very much.
2 = Approve somewhat.
3 = Approve very little.
4 = Do not approve.

The numbers assigned to the categories not only distinguish whether things are in the same category or a different category (as they do with nominal variables), but they also indicate an **ordered ranking** from 1, which equals high (approve very much), to 4, which equals low (do not approve). With ordinal variables, you can use all the statistical manipulations appropriate for nominal variables (such as the mode and the frequency distribution). In addition, because there is a rank order to the numbers, you can also use such statistics as the median and percentile, and various non-parametric statistical tests discussed later. But you cannot use a mean, a standard deviation, or a Pearson product-moment correlation. It makes no sense to speak about the mean family planning attitude of respondents. The reason is because the distance or interval between the categories is not known. In the example above, we do not know if the distance between 1 (approve very much) and 2 (approve somewhat) is the same as the distance between 3 (approve very little) and 4 (do not approve). It might be that respondents who fall into categories 1, 2, and 3 are really very similar to each other, while those who fall into category 4 are very different. We could represent this situation as in Figure 8.1. While there is a rank order in the numbers assigned to the categories of the variable, the distance between the categories is not equal.

Interval Measurement. An **interval variable** is one where the distance (or interval) between any two categories is known. This is the first truly quantitative level of measurement. The numbers assigned to categories of interval variables have all the characteristics of nominal and ordinal variables **plus** the added characteristic of a constant unit of measurement between categories that are equally spaced. Age is an interval variable. Parity is an interval variable. Duration of contraceptive practice and date of family planning acceptance are interval variables. For each of these, the distance between one category and the next is known and constant. A person aged 34 is four years older than a person aged 30. A person at parity 6 has two more

FIGURE 8.1
Ordinal measurement of the attitude toward the use of contraception

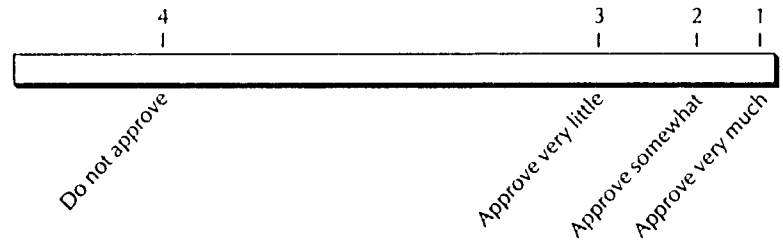
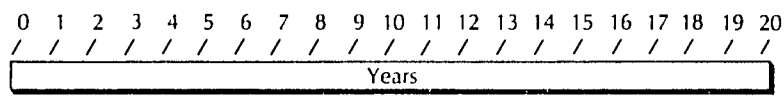


FIGURE 8.2
Interval measurement of age



children than a person at parity 4. Time, temperature, money, and weight a e all interval variables because the distance between the categories is known and constant. We could represent this situation as in Figure 8.2. A child, for example, who is placed in category 5 is not only different (nominal) from one placed in category 6 but also younger (ordinal), and one year younger (interval). All the statistical manipulations possible for nominal and ordinal variables are also possible but sometimes less appropriate for interval variables. In addition, with interval variables you can use such measures as the mean, standard deviation, Pearson product-moment correlation, and many other parametric statistics.

c Time Series Analysis

If a study is designed in such a way as to generate time series data, the analysis can be both very simple and very revealing. Time series analysis consists of looking at a particular measure as it changes over time. The measure can take many forms; most commonly it is a frequency (e.g., monthly number of acceptors), a percentage (e.g., the prevalence rate), or a measure of central tendency (e.g., mean desired family size). If an intervention has occurred during the time period covered (e.g., $O_1\ O_2\ O_3\ O_4\ X\ O_5\ O_6\ O_7\ O_8$), the trend before the intervention can be compared with the trend afterward to provide insights about its effects. Consider Table 8.1.

A quick review of this time series data reveals that something must have happened during the month of May. If what happened was a major nationwide campaign to increase the number of male sterilizations, the simple frequency distribution table above provides evidence that the campaign had a dramatic effect. But this table also reveals that, immediately following the campaign in May, the number of sterilization cases dropped to an average of 55 per month, compared against 116 per month before the campaign.

The table suggests several possible interpretations. One interpretation is that the campaign in May "borrowed cases from the future." In other words, people who might normally have had the operation performed in June, July, August, etc. decided instead to have their sterilization operations done during the campaign month of May. A second possible interpretation might be that, after the major effort in May, the family planning doctors, nurses, and field staff reduced their motivational efforts and simply did not recruit as many cases as they did before May. The frequency distribution table **only suggests** these interpretations, it does not prove them. These suggestions might form the basis for a research study designed to determine the reasons why the number of sterilization cases dropped immediately following the month of May.

d Comparisons

Study objectives often call for comparing variables. For instance, in an experimental design, the most important findings are usually those that are derived from comparison of experimental and control groups. If an experimental study is designed to determine the effect of an intervention on contraceptive practice, the percentage distributions of the experimental and control groups with regard to contraceptive practice should be compared. If it can be assumed safely that the two groups were initially similar and were not subjected to different influences other than the experimental treatment, then a significant difference in the proportion of the experimental group practicing contraception compared with the control group could be used to conclude that the intervention was successful.

TABLE 8.1 Time series data: Number of male sterilizations performed per month in 1982

Month	Number	
January	120	
February	110	
March	130	Average January–April = 116
April	105	
May	1,200	
June	30	
July	45	
August	15	
September	30	Average June–December = 55
October	60	
November	90	
December	115	
Total	2,050	

For testing the difference between pairs of means, the ***t* test** is most often used. The *t* test is used with small samples (*N* equal to or less than 30) and assumes that elements or observations are drawn from normally distributed populations with the same variance. For comparing percentage distributions, the **chi-square test** (X^2) is most commonly used. Sometimes you may want to compare percentages for responses obtained from the same sample at two different points in time. In such cases, the appropriate significance test is the **McNemar test for significance of changes.**

If the plan of analysis calls for comparing more than two means simultaneously, the most appropriate analytical technique is **analysis of variance (ANOVA).** This technique employs the ***F* test** to determine significance.

e Bivariate Relationships

Usually, study objectives in family planning operational research call for examining the **relationships** between pairs of variables, particularly between measures of program operations and measures of expected effects. The types of analytical techniques that are appropriate for studying bivariate relationships depend on the nature of the variables—whether they are nominal, ordinal, or interval.

Relationships between Nominal Variables. When examining the relationship between two nominal variables, the first step is to prepare a cross-tabulation of the two variables. The chi-square test can then be applied to the cross-tabulation to determine whether there is a significant relationship. However, the chi-square statistic does **not** measure the strength of the relationship. For this purpose, a "measure of association" is needed, and one of the best such measures of nominal variables is **Cramer's V,** a statistic derived from the chi-square value.

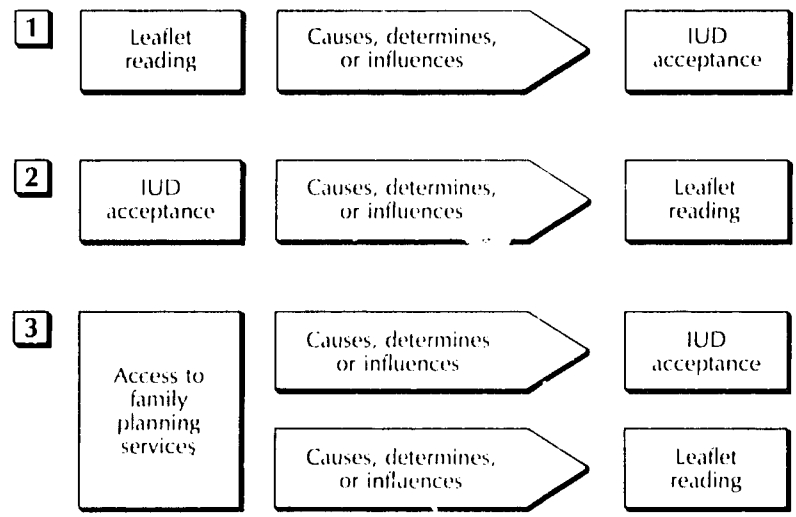
Relationships between Ordinal Variables. There are several different measures of association for cross-tabulations of ordinal variables. One of the simplest to calculate is **gamma.** Unfortunately, there is no simple significance test for gamma. (The chi-square test may be used to assess whether the relationship between variables is significant, but it should be remembered that this test is not sensitive to the ordinality of the data and therefore does not provide a true test of the significance of gamma.)

Relationships between Interval Variables. Relationships between interval variables may be studied with or without cross-tabulations. If interval variables are cross-tabulated, the nature of the relationship will become apparent, and the values of gamma or Cramer's V can be assessed. It is more common to measure the relationship between pairs of interval variables without reference to any cross-tabulations by means of the **Pearson correlation coefficient** (*r*). The statistical significance of *r* can be assessed by a *t* test.

If a distinction is made between independent and dependent variables, linear regression analysis may be used to examine the relationship. In this case, the measure of association is the **zero-order regression coefficient,** which indicates the average amount of change in the dependent variable associated with a unit of change in the independent variable.

Both the correlation and regression coefficients measure linear relationships, where the value of one variable rises or declines in direct proportion to rises or declines in the value of

FIGURE 8.3
Three possible relationships between leaflet reading and use of an IUD



the other variable. These two coefficients are not sensitive to nonlinear relationships where high values of one variable are associated with both high and low (but not middling) values of another variable.

All the measures of association discussed above, except for the regression coefficient and Cramer's V, range in value from −1.00 (a perfect negative relationship) to +1.00 (a perfect positive relationship). When there is no relationship, the coefficient is 0.00.

It should also be noted that none of the measures of association indicate whether a relationship is causal. If it is found that people who have read a leaflet about the IUD are more likely to have used the IUD, this does not necessarily mean that reading the leaflet **caused** them to use the IUD. It may be that people who have access to IUD services are also more likely to have access to the leaflet and that the relationship results entirely from this coincidence. It could even be that the relationship is causal but that the direction of causality is just the opposite. For instance, leaflets may tend to be given to IUD acceptors to provide them with more information about their new device, or IUD users may deliberately seek out information on the method following insertion, especially if they start to experience side effects. Thus, there are at least three possible relationships, as shown in Figure 8.3.

f Multivariate Analysis

Bivariate relationships alone often raise questions that can be answered only by multivariate analysis. For instance, if a relationship is found between religious affiliation and contraceptive practice, it may appear at first that the difference observed indicates a direct, causal relationship. However, it may also be argued that the difference is due to other factors. For instance, it has long been noted in the Philippines that Muslims are less likely than the Christian majority to practice contraception. However, this relationship may not be causal, since Muslims also tend to have much lower socioeconomic status and to live in areas served by proportionately fewer family planning field-workers and service points. To test whether the relationship between religious affiliation and contraceptive practice is **explained** by differences in status or access to family planning information and services, it is necessary to **control for** or **hold constant** measures of status and access. If the relationship between religious affiliation continues undiminished even when these measures are held constant, it can be concluded that status and access have little or no influence on the relationship. If the relationship disappears, it can be concluded that it was **spurious**—caused by the simultaneous relationship of both the independent and dependent variables to the control variables.

Techniques that permit a researcher to study the effect of controlling for one or more variables are called **multivariate,** since they involve multiple (more than two) variables. Most multivariate techniques also permit the measurement of the **degree of relationship** between a dependent variable and two or more independent variables considered simultaneously. There are several different multivariate techniques, each of which is particularly suitable for certain configurations of variables. With an interval dependent variable, three types of multivariate analysis are most commonly used. The one selected is based on the types of independent variables described below.

All Interval. If the dependent and independent variables are all interval variables, the appropriate technique is **multiple regression analysis.** This technique is the best known and most readily

available of the multivariate techniques. It generates **partial regression coefficients,** which indicate the amount of increase in the dependent variable associated with a single unit of increase in each independent variable, controlling for all the other independent variables simultaneously. These coefficients are often converted into **standardized regression coefficients** (beta), which indicate the change in the dependent variable associated with changes in the independent variable in terms of standard deviations rather than units of measurement. It also produces a **multiple correlation coefficient** (R), which measures the simultaneous relationship between the dependent variable and all the independent variables. The square of the multiple correlation coefficient, known as the **coefficient of determination** (R^2), is a measure of the proportion of variance in the dependent variable that is explained by all the independent variables taken together.

Both Interval and Categoric. If the dependent variable is interval but the independent variables include both interval and categoric variables (nominal and ordinal), the most commonly used procedure is **multiple regression analysis with dummy variables.** A dummy variable is a variable consisting of only two categories. A categoric variable can be converted into several dummy variables. For instance, a nominal three-category variable representing religious affiliation (Christian, Muslim, and Hindu) can generate three dummy variables of two categories each. The variable Christian can be created by assigning a code of 1 to all Christians and a code of 0 to all others; the variable Muslim can be created by assigning a code of 1 to Muslims and a code of 0 to all others; and the variable Hindu can be created by assigning a code of 1 to Hindus and a code of 0 to all others. A variable of four categories can generate four dummy variables in the same manner, a variable of five categories can generate five dummy variables, and so forth. For multiple regression analysis, **all but one** of the dummy variables are entered as independent variables for each of the original categoric variables. With dummy variables, the regression coefficients indicate the difference in the dependent variable between the category specified by the dummy variable and the category omitted from analysis. The interpretation of R and R^2 is the same as in ordinary regression analysis.

All Categoric. With an interval dependent variable and all categoric independent variables, the most appropriate technique is **multiple classification analysis** (MCA). This is really an extension of the multiple regression analysis with dummy variables, but it produces a Beta value for each of the original categoric variables, whereas regression analysis produces less useful beta values for each dummy variable included. Like multiple regression analysis, MCA also generates values of R and R^2.

If the dependent variable is categoric, analysis is generally limited to multidimensional cross-tabulations. (If some or all of the independent variables are interval, they can be easily converted to categoric variables for this purpose.) Efforts have been made in recent years to develop more sophisticated techniques for multivariate analysis involving nominal dependent variables, but few researchers know how to use them, and the necessary computer programs are not yet widely available.

Multivariate techniques can be very powerful analytical tools, but they must be used with great care. They are all based on numerous assumptions, some of which usually are not met. As a result, apparent findings often are not valid. Your plan of analysis should not include any multivariate technique unless you are already well acquainted with it or can call on the assistance of a consultant who knows how to use it.

g Cost-Effectiveness Analysis

All family planning programs cost money. The money buys program activities, which are designed to produce desired outcomes or effects. In most family planning programs, it is possible to identify two or more activities that are intended to produce the same effect, such as a new acceptor, a birth averted, or a couple-year of protection. A logical question for administrators to ask is, "Which activity is the least costly?" Cost-effectiveness analysis (CEA) is intended to answer this question. Essentially, CEA is a decision-making tool. It has meaning only when it is possible to make a cost comparison between two or more different approaches, or a comparison of a single approach over time. It is a procedure that analyzes alternative means to reach the same end objective in terms of cost.

CEA involves calculating a ratio with program costs as the numerator and program effects as the denominator. While the calculation is simple, there are a number of difficulties with CEA.

First, there is the problem of determining the exact costs of a program activity. Often this is not easy, particularly if family planning activities are a part of and fully integrated with other health activities. For example, what part of the cost of a health clinic doing IUD insertions should be attributed to family planning? How much of a doctor's or a nurse's time (and thus their salary costs) should be charged against family planning? Should overhead and other types of indirect costs be included? Should inflation or international changes in the value of a currency be considered? Inevitably, these and other similar questions arise in cost-effectiveness analysis.

Besides the difficulty of establishing the costs of a program activity, a second problem arises in trying to collect information to determine the program effect. The unit of effect must be defined clearly. What is a family planning acceptor? What is a couple-year of protection or a birth averted, and how are these terms operationally defined? There are other difficulties as well. For example, the effects of many program activities are distributed over time. A sterilization procedure may avert births over a period of many years. In calculating a cost-effectiveness ratio, assumptions must be made about the distribution of effects over time. Indeed, sometimes CEA involves so many assumptions about costs and effects that its utility as a decision-making tool is questionable. Also, there may be situations where the most cost-effective program approach is not necessarily the most socially acceptable or politically desirable.

Cost-effectiveness analysis is probably most useful when comparing very specific, time-limited experimental interventions where outcomes are well defined and program costs are detailed and can be disaggregated (separated out). For example, suppose an outcome measure is defined to be the recruitment and continuation of a new pill acceptor for at least one year. CEA might then be undertaken to determine whether mobile vans, outreach workers, or clinics are more cost-effective. Similarly, it might be possible to use CEA to determine which of several different types of training is the most cost-effective to obtain a skilled field-worker. CEA could also be used to compare the cost of sterilization operations done in camp settings versus the cost of those done in clinic or hospital settings.

h Use-Effectiveness Analysis

Often in family planning operations research it is important to measure the probability that an acceptor will continue to use a method from month to month and will not become pregnant. The probabilities can be estimated through **life-table analysis of use-effectiveness.** The most important outputs of this type of analysis are **continuation rates** and **pregnancy rates.** Distinctions can be made between monthly and cumulative continuation rates, between single-method and all-method continuation rates, between gross and net pregnancy rates, and between failure rates and overall pregnancy rates:

A **monthly continuation rate** is the probability that a person who starts a month using a method will still be using the method at the end of the month.

A **cumulative rate** is the probability that an acceptor will still be using after a given number of months following acceptance.

A **single-method continuation rate** is the probability of continuing to use the same method.

An **all-method rate** is the probability of continuing to use any method.

A **gross pregnancy rate** is the probability that a person will become pregnant during a given period of time, disregarding competing risks of termination (such as IUD expulsion or voluntary discontinuation due to side effects).

A **net pregnancy rate** is like a gross rate except that it takes into account competing risks.

The **failure rate** is the probability of becoming pregnant during a specified period of time while using a contraceptive method.

The **overall pregnancy rate** is the probability of becoming pregnant during a specified period of time following acceptance, regardless of whether or not a method is being used at the time.

1 Fertility Analysis

Since most family planning programs are designed to help bring about or accelerate a reduction in fertility, it is often necessary to incorporate fertility analysis into family planning operations research. The most useful fertility measures for this purpose are likely to be:

Crude birth rate: the number of births in a year per 1,000 people.

General fertility rate: the number of births in a year per 1,000 women of reproductive age.

Marital general fertility rate: the number of legitimate births in a year per 1,000 married women of reproductive age.

Age-specific fertility rates: the number of births to women in a given age category per 1,000 women in that same category.

Age-specific marital fertility rates: the number of legitimate births to married women in a given age category per 1,000 married women in the same category.

Total fertility rate: the average number of children that would be expected during a woman's reproductive life if a particular pattern of age-specific fertility rates were to prevail throughout.

C *Dummy Tables*

Constructing dummy tables is a very useful way to visualize how data can be organized and summarized. A dummy table contains all the elements of a real table, except, of course, that the cells of the table are empty (see Table 8.2 for an example). In a research proposal, the major relationships between variables should be shown in dummy tables.

Cross-tabulation tables are usually presented with cell frequencies converted into percentages based on either row or column totals. If a dependent variable is cross-tabulated with an independent variable, the percentages should be calculated so that they add up to 100 percent for each category of the independent variable. For example, in the dummy table presented below, the percentage totals are calculated for the categories of the independent variable (e.g., "presence of field-workers in village").

TABLE 8.2 Example of a dummy table: Village contraceptive prevalence by presence of field-workers in village (percentage distribution)

	Presence of health and family planning worker in village		
Village prevalence rate	**No** (N=)	**Yes** (N=)	**All villages** (N=)
50+ percent			
25–49 percent			
0–24 percent			
Total percentage	100	100	100

χ^2 d.f. p= gamma =

What To Do—The Plan of Analysis

1. Describe fully and in detail each of the analytical techniques and statistical measures you plan to use, indicating how each technique and measure will help you meet your study objectives. What variables will be involved? Why have you selected a particular technique for the analysis of the data?
2. Provide examples of important dummy tables.
3. Be sure that your plan of analysis clearly explains how you will meet all your study objectives, utilize all your study variables, and test all your study hypotheses.

9 Reporting Research Findings

There is very little reason to conduct research unless the results are communicated to others in a form that is understandable. Writing a research report provides the investigator with an opportunity to fit all the pieces together. It is important to remember that research reports should be written so that the reader can judge the scientific value of the study, assess the adequacy of the study design used, and, if desired, repeat the study in other areas or with other subjects.

Bear in mind that operations research is conducted primarily to provide useful information to program managers. The report should be written in a style that is appropriate for managers. Since managers usually are not research methodologists, the report should be kept as nontechnical as possible. It is often a good idea to include technical material in an appendix (or in several appendixes) rather than in the body of the report. Sometimes managers have difficulty seeing the relevance of research findings for program management. You should help them by indicating the implications you think the findings have and recommending ways in which managers can make use of them. Since managers are usually very busy and have relatively little time for reading long reports, the report should be made as easy to read as possible. You should provide a concise summary at the beginning of the report, starting with a bare minimum of introductory and background materials and enumerating the major findings and recommendations that you consider most likely to be useful for program managers. Even though it should appear at the beginning of the report, the summary should not be written until after the rest of the report has been completed.

If a study is expected to take a long time to complete (more than about 12 months) and if useful data will be produced early enough to permit, it may be a good idea to schedule more than one report. The final report should be comprehensive, but earlier interim reports can be relatively brief. One of the major complaints of program managers about research reports is that they often take too long to come out and the data are already out of date. This problem can be reduced if interim reports are issued as soon as possible after important data become available.

Once a report is written, it is necessary to make sure that it is disseminated promptly and to the appropriate audience. If you are primarily concerned with reaching policymakers and program administrators, it is helpful to augment formal written reports with oral presentations, using audiovisual aids if possible. If a special seminar or workshop can be arranged for such a presentation, the participants are likely to be especially attentive (and the discussion to be most useful) if the intended output of the meeting, stated explicitly at the outset, is to produce a plan of action based on the findings and recommendations.

If your audience is to be midlevel officers stationed in rural areas, the report should be written in language that will be understood by them. If they speak a language other than the one in which the report is written, the report should be translated. If you are interested in reaching social scientists, it would be a good idea to write a different version of the report for them, focusing more on the technical and theoretical issues that would interest them. If you are interested in reaching a larger, international audience, you should consider submitting a version of the report to an international journal for possible publication.

Example of Major Headings for a Final Research Report

I Title Page (title of the report, authors, institutional affiliation, date)
II Preface (acknowledgments, source of funding)
III Abstract
IV Background (location of study, special circumstances of study)
V Literature Review
VI Study Methodology (objectives, hypotheses, description of program intervention, study design, analytic procedures, limitations of study)
VII Findings
VIII Discussion of Findings
IX Conclusions and Recommendations
X References and Bibliography
XI Appendixes

What To Do—Reporting Research Findings

1 State how you plan to disseminate the findings and recommendations from your study.
2 Indicate the target audience of the report(s) to be produced.
3 Indicate the language(s) in which the report is to be written.
4 Indicate whether a workshop or seminar is planned in which the implications of the findings are to be discussed.

10 Limitations of the Study

A Design and Analysis Limitations

There is probably no such thing as a perfect study. Every research study has some problems with the reliability and validity of the data, with the size of the sample, with the questionnaire design, or with the analysis plan. Good research attempts to limit these problems to the minimum, but some problems will remain. In a research study proposal, it is best to recognize the limitations of your design rather than try to pretend that these limitations do not exist. For example, if you have purposively selected a study sample, then do not try to pretend that this sample is representative of a much larger group. If you are forced because of limited time and funds to use data that may be of questionable quality (such as clinic records or service statistics), do not try to pretend that the data are completely reliable and valid. If your comparison group is not equivalent with your experimental group, do not try to pretend that it is. Instead, recognize the various problems in your study design and try to deal with them statistically in the analysis, if that is possible, or by providing a note of caution to the reader that, because of weaknesses in the study design, the findings must be considered as preliminary.

B Special Situational Factors

In your proposal, mention any special situations that might influence the study. In Asia, for example, it is often difficult to conduct field studies during the monsoon period. If this is a problem, say that the study can be conducted only in nonmonsoon months. If computer facilities are unavailable or if they are unreliable, mention this fact. If conducting your study is dependent first on obtaining approval from senior government officials (which often can cause long delays), say so.

What To Do—Limitations of the Study

1. Be frank about the limitations of your study and the possible situational factors that might influence the study.
2. Explain the limitations and assumptions of your study.

Resources and Facilities

11

A Available Resources and Facilities

Most reviewers of study proposals will want to know the resources and facilities **already available** for conducting the research. For example, are experienced interviewers and coders available? Are computer facilities available? Will other institutions or organizations contribute to the cost of the study? How much time will the principal investigator be able to devote to the study—100 percent, 50 percent, 10 percent? These and similar questions relating to the existing availability of resources and facilities should be answered in the proposal.

B Study Budget

The budget for the research study should be realistic. In general, most funders of research will not provide money for expensive equipment, for building construction, or for vehicles. Also, most funders will not provide salary payments to principal investigators that are above what he or she has received in other jobs. **Any large and unusual item in the budget should be explained and justified.** Be very clear about each cost item in the budget. Show the components of the item. For example, if you plan to employ interviewers, their salary costs might be shown as:

Field Interviewers (20 at Rs. 500 per day × 20 days)

Arrange the budget under major cost categories (see Figure 11.1). If your study extends beyond a year, show first-year costs separate from second-year costs. Also, for long studies, you might wish to include a line item to cover expected inflation.

What To Do—Resources and Facilities

1 Describe the resource and facilities already available for the study. Be sure to indicate:
   - **a** Whether other institutions or organizations will contribute resources.
   - **b** The availability of computers, trained interviewers and coders, secretarial help, vehicles, office space, etc.

2 For the study budget, arrange the cost items under headings. Major headings would include:
   - **a** Salaries and Benefits
   - **b** Materials, Supplies, Computer Services
   - **c** Travel
   - **d** Miscellaneous

3 At the end of the budget, explain and justify any large or unusual cost items.

FIGURE 11.1
Example of a study budget

		Local ฿	US$
A	Salaries and benefits		
	1 Principal investigator, ฿2,000/month x 12 months	24,000	1,200
	2 Associate investigator, ฿1,700 x 12 months	20,400	1,020
	3 Field interviewers, 20 @ ฿100 per day x 20 days	40,000	2,000
	4 Coders, 20 person-days @ ฿150/day	3,000	150
	5 Secretarial services (not charged to study)	--	--
	Subtotal	87,400	4,370
B	Materials, supplies, computer services		
	1 Questionnaire printing	3,000	150
	2 Administrative supplies	7,600	380
	3 IBM card punching/verifying	4,000	200
	4 Computer programming	5,000	250
	5 Computer time	15,000	750
	6 Report printing and mailing (200 copies)	25,000	1,250
	Subtotal	59,600	2,980
C	Travel		
	1 Field testing of questionnaire	10,000	500
	2 Per diem and local travel for field interviewers (20) @ ฿100 per day	40,000	2,000
	Subtotal	50,000	2,500
D	Miscellaneous expenses	2,000	100
	Total direct costs	199,000	9,950
	Total project costs	199,000	9,950

Exchange rate: Baht 23.00 = US$1.00 as of 1981.

Study Work Schedule

12

Each research proposal should include a work schedule. The schedule should list the major activities that will be undertaken and the times when they will be undertaken. See Table 12.1 for an example of a study work schedule.

TABLE 12.1 Example of a study work schedule

Activities	First Year												Second Year											
	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
A Planning Phase																								
1 Determine Sample	—																							
2 Design Questionnaire	—	—																						
3 Recruit Interviewers and Supervisors		—																						
4 Train Interviewers and Supervisors			—																					
5 Pretest Questionnaire				—																				
6 Revise Questionnaire					—																			
B Implementation Phase																								
1 Conduct Interviews						—	—	—	—															
2 Conduct Reinterviews							—	—	—															
3 Code Data						—	—	—	—	—														
4 Check for Consistency						—	—	—	—	—														
5 Key Punch Data							—	—	—	—														
6 Tabulate Data											—													
C Analysis Phase																								
1 Analyze Data												—	—	—										
2 Write Research Report														—	—	—	—							
D Reporting Phase																								
1 Print Report																	—	—						
2 Disseminate Report																			—	—				
3 Present Findings at Conference																					—			

What To Do—Work Schedule

Design a work schedule. Consider the realistic timing for each of the following activities (if they apply in your study):

1. Planning phase
2. Pilot studies
3. Drawing of sample
4. Questionnaire preparation
5. Selection and training of interviewers
6. Pretesting of questionnaire
7. Baseline survey
8. Data collection
9. Data processing
10. Data analysis
11. Report writing
12. Dissemination of research findings

Appendixes

13

Include in the appendix of your proposal any additional information you think might be helpful for a proposal reviewer. For example, include bio-data on the principal investigators. Include the study questionnaire if you have it. Include the Informed Consent Form. Include any explanatory material (such as an annual report) about your institution or the organization under whose name the study will be conducted. If you have cited literature in the proposal, include a list of references in the appendixes.

14 Title Page and Abstract

Although they appear as the first section of a research proposal, the title page and abstract are the last to be written. The title page (Figure 14.1) gives the essential information about the proposal. An abstract (Figure 14.2) is a summary of the basic information contained in all the other sections. Do not overload an abstract with unnecessary information. Keep it short (no longer than one or two pages), precise, and to the point. The abstract should tell the reader:

1. The **problem** to be studied.
2. The main **objectives** of the study.
3. The major expected **implications** of the study.
4. **Who** will conduct the study.
5. **When** the study will be conducted.
6. **Where** the study will be conducted.
7. What **methods** will be used to conduct the study.
8. What **resources** are required for the study.

What To Do—Title Page and Abstract

1. From each major section of your completed proposal, write one or two sentences that give the essence of the information in that section.
2. Arrange the sentences into an abstract.
3. Attach a title page to the abstract.

FIGURE 14.1
Example of a title page

Family Planning Research Proposal

1 Title: An Experimental Study to Increase Family Planning Acceptance through the Use of Development Assistance

2 Location: Khon Khen Province, Northeast Thailand

3 Sponsoring institution(s): The Institute for Rural, Social, and Economic Development (IRSED)

4 Principal investigator(s) (name and address): Mr. Prasert
Director
IRSED
Bangkok, Thailand

5 Starting date: February 1983

6 Completion date: July 1984

7 Total cost: US$31,580

Signature _____

Date _____

FIGURE 14.2
Example of an abstract

There is wide variation in the rate of contraceptive prevalence
(Problem) between villages in Bangladesh. Although in theory all villages receive the same level of health and family planning services, some villages have a contraceptive prevalence rate as high as 80 percent of currently married women, while other villages have a rate as low as 6 percent. The immediate objective of this study is to investigate the
(Objective) social, economic, and health factors associated with areal variation in contraceptive prevalence rates. The ultimate objective is to provide policymakers and program administrators with a better understanding of the reasons why the national family planning program is successful in
(Implications) some areas but not in others. This information can be used to develop educational strategies and modify existing service delivery approaches so that the rate of contraceptive practice is uniformly high in all areas of the country.

(Who) The study will be conducted between January 1982 and December 1982
(When) by the Institute of Population Research located in Dhaka, Bangladesh.
(Where) All villages within two districts will be classified in terms of
(Methods) contraceptive prevalence rates. Twenty-five villages will then be selected randomly from within high, medium, and low prevalence rate strata. All currently married women aged 15-44 will be interviewed in each village selected. Information on village-level social, economic, and health characteristics will also be collected. A team of trained
(Resources) interviewers and supervisors will collect the field data. All coding, editing, and data processing will be done in Dhaka. The computer facilities of Dhaka University will be used to tabulate the data. A final research report will be completed by December 1983. Funds of US$34,500 are required for staff salaries, travel, questionnaires, and computer time.

Bibliography

Andrews, Frank M., Laura Klem, Terrence N. Davidson, Patrick M. O'Malley, and Willard L. Rodgers. *A Guide for Selecting Statistical Techniques for Analyzing Social Science Data*. Ann Arbor: Institute for Social Research, 1974.

Andrews, Frank M., James N. Morgan, John A. Sonquist, and Laura Klem. *Multiple Classification Analysis* (2nd ed.). Ann Arbor: Institute for Social Research, 1975.

Barclay, George Watson. *Techniques of Population Analysis*. New York: Wiley, 1958.

Berelson, Bernard. *Content Analysis in Communication Research*. New York: Hafner, 1971.

Blalock, Hubert M., Jr. *An Introduction to Social Research*. Englewood Cliffs, New Jersey: Prentice-Hall General Sociology Series, 1970.

Blalock, Hubert M., Jr. *Causal Models in the Social Sciences*. Chicago: Aldine-Atherton, 1971.

Blalock, Hubert M., Jr. *Social Statistics* (2nd ed.). New York: McGraw-Hill, 1972.

Blalock, Hubert M., Jr., and Ann B. Blalock. *Methodology in Social Research*. New York: McGraw-Hill, 1968.

Bogue, Donald J. *Cost-Effectiveness Analysis of Family Planning Programs*. Manual #11. Chicago: Community and Family Study Center, 1973.

Bourke, G. J., and J. McGiluray. *Interpretation and Uses of Medical Statistics*. Oxford and Edinburgh: Blackwell Scientific Publications, 1969.

Campbell, Donald T., and Julian C. Stanley. *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally, 1963.

Caro, Francis G. (ed.). *Readings in Evaluation Research*. New York: Russell Sage Foundation, 1971.

Chandrasekaran, C., and Albert I. Hermalin (eds.). *Measuring the Effect of Family Planning Programs on Fertility*. Dolhain: Ordina Editions, 1976.

Cochran, W. G., and G. M. Cox. *Experimental Designs* (2nd ed.). New York: Wiley, 1957.

Cook, Thomas D., and Donald T. Campbell. *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally College & Publishing Co., 1979.

Cuca, Roberto, and Catherine S. Pierce. *Experiments in Family Planning: Lessons from the Developing World*. Baltimore: The Johns Hopkins University Press for the World Bank, 1977.

Davis, James A. *Elementary Survey Analysis*. Englewood Cliffs, New Jersey: Prentice-Hall, 1971.

Dixon, Wilfred J., and Frank J. Mossey, Jr. *Introduction to Statistical Analysis*. (3rd ed.). New York: McGraw-Hill, 1969.

Dunn, Olive Jean. *Basic Statistics: A Primer for the Biomedical Sciences*. New York: Wiley, 1964.

Fink, Allene, and Jacqueline Kosecoff. *An Evaluation Primer*. Beverly Hills: Sage Publications, 1978.

Fitz-Gibbon, Carol Taylor, and Lynn Lyons Morris. *How to Calculate Statistics*. Beverly Hills: Sage Publications, 1978.

Fitz-Gibbon, Carol Taylor, and Lynn Lyons Morris. *How to Design a Program Evaluation*. Beverly Hills: Sage Publications, 1978.

Freund, John E. *Modern Elementary Statistics*. (2nd ed.). Englewood Cliffs, New Jersey: Prentice-Hall, 1960.

Green, Lawrence W., Marshall W. Kreuter, Sigrid G. Deeds, and Kay B. Partridge. *Health Education Planning: A Diagnostic Approach*. Palo Alto: Mayfield Publishing Co., 1980.

Grundy, F., and W. A. Reinke. *Health Practice Research and Formalized Managerial Methods*. Geneva: World Health Organization, 1973.

Guilford, J. P. *Psychometric Methods*. New York: McGraw-Hill, 1954.

Guttentag, Marcia, and Elmer L. Struening (eds.). *Handbook of Evaluation Research* (Vols. I and II). Beverly Hills: Sage Publications, 1975.

Hansen, Morris H., William N. Hurwitz, and William G. Madow. *Sample Survey Methods and Theory*. New York: John Wiley & Sons, 1953.

Haupt, Arthur, and Thomas T. Kane. *Population Handbook*. Washington, D.C.: The Population Reference Bureau, 1980.

Henerson, Marlene E., Lynn Lyons Morris, and Carol Taylor Fitz-Gibbon. *How to Measure Attitudes*. Beverly Hills: Sage Publications, 1978.

Hermalin, Albert I., and Barbara Entwisle (eds.). *The Role of Surveys in the Analysis of Family Planning Programs*. Liege: Ordina Editions, 1980.

Hilton, Elizabeth T., and Arthur A. Lumsdaine. Field Trial Designs in Gauging the Impact of Fertility Planning Programs. In Carl A. Bennett and Arthur A. Lumsdaine (eds.), *Evaluation and Experiment: Some Critical Issues in Assessing Social Programs*. New York: Academic Press, 1975.

Hyman, Herbert. *Survey Design and Analysis*. Glencoe: The Free Press, 1955.

Katz, F. M., and R. Snow. *Assessing Health Workers' Performance: A Manual for Training and Supervision*. Geneva: World Health Organization, 1980.

Katzer, Jeffrey, Kenneth H. Cook, and Wayne W. Crouch. *Evaluating Information: A Guide for Users of Social Science Research*. Reading, Massachusetts: Addison-Wesley Publishing Company, 1978.

Kerlinger, Fred N. *Foundations of Behavioral Research*. New York: Holt, Rinehart and Winston, 1964.

Kish, Leslie. *Survey Sampling*. New York: Wiley, 1965.

Knox, E. G., (ed.). *Epidemiology in Health Care Planning: A Guide to the Uses of a Scientific Method.* Oxford: Oxford University Press, 1979.

Laing, John. *Demographic Evaluation of Family Planning Programs.* Canberra: The Australian National University, 1982.

Lewis-Beck, Michael S. *Applied Regression: An Introduction.* Beverly Hills: Sage Publications, 1980.

Lilienfeld, Abraham, and David E. Lilienfeld. *Foundations of Epidemiology.* New York: Oxford University Press, 1980.

Miller, Delbert C. *Handbook of Research Design and Social Measurement.* New York: Longman, 1964.

Morris, Lynn Lyons, and Carol Taylor Fitz-Gibbon. *Evaluator's Handbook.* Beverly Hills: Sage Publications, 1978.

Morris, Lynn Lyons, and Carol Taylor Fitz-Gibbon. *How to Present an Evaluation Report.* Beverly Hills: Sage Publications, 1978.

Morton, Richard F., and J. Richard Hebel. *A Study Guide to Epidemiology and Biostatistics.* Baltimore: University Park Press, 1979.

Nie, Norman H., C. Hadlai Hull, Jean G. Jenkins, Karen Steinbrenner, and Dale H. Bent. *SPSS: Statistical Package for the Social Sciences* (2nd ed.). New York: McGraw-Hill, 1975.

Payne, Stanley L. *The Art of Asking Question.* Princeton, New Jersey: Princeton University Press, 1951.

Polansky, Norman A. (ed.). *Social Work Research.* Chicago: University of Chicago Press, 1970.

Popham, W. James. *Educational Evaluation.* Englewood Cliffs, New Jersey: Prentice-Hall, 1975.

Population Council. *A Manual for Surveys of Fertility and Family Planning: Knowledge, Attitudes, and Practice.* New York: The Population Council, 1970.

Pressat, Roland. *Demographic Analysis.* Chicago: Aldine Atherton, 1972.

Reynolds, Jack. *A Framework for the Selection of Family Planning Program Evaluation Topics.* Manual #1 (rev. ed.). New York: International Institute for the Study of Human Reproduction, 1973.

Reynolds, Jack. *Operational Evaluation of Family Planning Program Through Process Analysis.* Manual #4 (rev. ed.). New York: International Institute for the Study of Human Reproduction, 1973.

Roberto, Eduardo L. *Strategic Decision-Making in a Social Program: The Case of Family Planning Diffusion.* Lexington, Mass.: Lexington Books, 1975.

Robinson, Warren C. Cost Benefit and Cost-Effectiveness Analysis in Family Planning Programs. In *Financial Management of Population/Family Planning Programs.* Kuala Lumpur: Inter-Governmental Coordinating Committee, 1976, pp. 82–91.

Rosenberg, Morris. *The Logic of Survey Analysis.* New York: Basic Books, 1968.

Rossi, Peter H., and Howard E. Freeman. *Evaluation: A Systematic Approach.* Beverly Hills: Sage Publications, 1982.

Rossi, Peter H., and W. Williams. *Evaluating Social Programs: Theory, Practice and Politics.* New York: Seminar Press, 1972.

Rutman, Leonard. *Planning Useful Evaluations: Evaluability Assessment.* Beverly Hills: Sage Publications, 1980.

Rutman, Leonard (ed.). *Evaluation Research Methods: A Basic Guide.* Beverly Hills: Sage Publications, 1977.

Selltiz, Claire, Marie Jahoda, Morton Deutsch, and Stuart W. Cook. *Research Methods in Social Relations* (rev. 1-volume ed.). New York: Holt, Rinehart and Winston, 1959.

Shryock, Henry S., and Jacob S. Siegel and Associates. *The Methods and Materials of Demography.* New York: Academic Press, 1976.

Siegel, Sidney. *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill, 1956.

Simon, Julian L. *Basic Research Methods in Social Science.* New York: Random House, 1969.

Sinquefield, Jeanne E. *Single and Multiple Decrement Life Table Procedures for Analysis of the Use-Effectiveness of Contraception.* Manual 8. Chicago: Community and Family Study Center, 1973.

Snedecor, George W., and William G. Cochran. *Statistical Methods.* Ames: Iowa State University Press, 1972.

Studies in Family Planning. Special Issue: Focus Group Research. Vol. 12, no. 12, Part 1.

Suchman, Edward A. *Evaluative Research: Principles and Practices in Public Services and Social Action Programs.* New York: Russell Sage Foundation, 1967.

Sudman, Seymour. *Applied Sampling.* New York: Academic Press, 1976.

Summers, Gene F. (ed.). *Attitude Measurement.* Chicago: Rand McNally & Co., 1970.

Tanur, Judith M. (ed.). *Statistics: A Guide to the Unknown.* San Francisco: Holden-Day, 1972.

Taylor, Howard C., Jr., and Robert J. Lapham. *General Guidelines for an MCH-Based Family Planning Project.* New York: The Population Council, 1974.

Torgerson, Warren S. *Theory and Methods of Scaling.* New York: Wiley, 1958.

United Nations. *The Methodology of Measuring the Impact of Family Planning Programmes on Fertility.* Manual IX. New York: United Nations, 1979.

United Nations. Statistical Office. *A Short Manual on Sampling.* Studies in Methods, Series F, No. 9. New York: United Nations, 1960.

Warwick, Donald P., and Charles A. Lininger. *The Sample Survey: Theory and Practice.* New York: McGraw-Hill, 1975.

Webb, Eugene J., Donald T. Campbell, Richard D. Schwartz, and Lee Sechrest. *Unobtrusive Measures: Non-Respective Research in the Social Sciences.* Chicago: Rand McNally, 1966.

Weiss, Carol H. (ed.). *Evaluating Action Programs: Readings in Social Action and Education.* Boston: Allyn and Bacon, 1972.

Weiss, Carol H. *Evaluation Research: Methods for Assessing Program Effectiveness.* Englewood Cliffs, New Jersey: Prentice-Hall, 1972.

Wishik, Samuel M., and Kwan-Hwa Chen. *Couple Years of Protection.* Manual 7. New York: International Institute for the Study of Human Reproduction, 1973.

World Fertility Survey. *Basic Documentation Series.* Nos. 1–10, Core Questionnaire and related manuals and documentation. The Hague-Voorburg: International Statistical Institute, 1975–1977.

World Health Organization. *Statistical Indicators of Family Health.* Geneva: World Health Organization, Technical Report Series, No. 587, 1976.

World Health Organization. *Health Programme Evaluation.* Geneva: World Health Organization, 1981.