DEPARTMENT OF
PUBLIC
HEALTH
& POLICY

CLUSTER I

Course Unit Reader : SU144

Design and Management of Research

Summer Term 1991-92

Moses Atkins

READER

Study Unit: **Design and Management of Research**

Organizer: **Trudy Harpham, Health Policy Unit, Department of Public Health and Policy**

CONTENTS

Identifying and prioritizing problems for research

I. PROBLEM IDENTIFICATION

If the answer to the research question is obvious, we are dealing with a **management problem** that may be solved without further research. If, for example, in the sanitation project essential building materials, such as cement, have been unavailable for a large part of the project period, one should try to ensure the supply of cement rather than embark on research to explore the reasons why the project did not reach its targets.

In the previous module, a number of research questions were presented that may be posed at the various levels of the health system.

These questions can be placed in three broad categories, depending on the type of information sought:

1. **Description of health problems required for planning interventions.**

   Planners need to know the magnitude and distribution of health needs as well as of health resources, to formulate adequate policies and plan interventions.

2. **Information required to evaluate ongoing interventions** with respect to:

   - Coverage of health needs
   - Coverage of target groups
   - Quality
   - Cost
   - Effects/impact

   to assess progress and the need for adjustment on a routine basis.

3. **Information required to define problems situations** arising during the implementation of health activities, **to analyze possible causes to find solutions.**

Although research in support of planning and evaluation (categories 1 and 2 mentioned above) is an important focus for HSR, the modules will concentrate on the third category, because mid-level managers are frequently confronted with problems of this type. It is assumed, however, that research skills acquired in the present course will be of use in the broader field of planning and evaluation as well.

Whether a problem situation requires research depends on three conditions:[1]

1. There should be a **perceived difference or discrepancy** between what exists and the ideal or planned situation;

2. The **reason(s)** for this difference should be **unclear** (so that it makes sense to develop a research question); and

3. There should be **more than one possible answer** to the question or solution to the problem.

---

[1] This paragraph has been adapted from Fisher et al. (1983).

For example:

Problem situation

In District X (pop. 145,000), sanitary conditions are poor (5% of households have latrines) and diseases connected with poor sanitation, such as hepatitis, gastroenteritis, and worms, are very common. The Ministry of Health has initiated a sanitation project that aims at increasing the number of households with latrines by 15% each year. The project provides materials and the population should provide labour. Two years later, less than half of the target has been reached.

Discrepancy

35% of the households **should have** latrines, but only 15% **do have** them.

Research question

What factors can explain this difference?

Possible answers

1. **Service-related factors**, such as forgetting to adequately inform and involve the population, bottlenecks in the supply of materials, differences in training, and effectiveness of sanitary staff.

2. **Population-related factors**, such as situations where community members lack an understanding of the relationship between disease and sanitation or have a greater interest in other problems.

II. CRITERIA FOR PRIORITIZING PROBLEMS FOR RESEARCH

Because HSR is intended to provide information for decision-making to improve health care, the selection and analysis of the problem for research should involve those who are responsible for the health status of the community. This would include managers in the health services and in related agencies, health-care workers, and community leaders, as well as researchers.

Each problem that is proposed for research has to be judged according to certain guidelines or criteria. There may be several ideas to choose from. Before deciding on a research topic, each proposed topic must be compared with all other options. The **guidelines or criteria** discussed on the following page can help in this process:

Criteria for selecting a research topic

1. Relevance
2. Avoidance of duplication
3. Feasibility
4. Political acceptability
5. Applicability
6. Urgency of data needed
7. Ethical acceptability

1. Relevance

The topic you choose should be a priority problem. Questions to be asked include:

- How large or widespread is the problem?
- Who is affected?
- How severe is the problem?

Try to think of serious health problems that affect a great number of people or of the most serious problems that are faced by managers in the area of your work.

Also, consider the question of **who** perceives the problem as important. Health managers, health staff, and community members may each look at the same problem from different perspectives. Community members, for example, may give a higher priority to economic concerns than to certain public health problems. To ensure full participation of all parties concerned, it is advisable to define the problem in such a way that all have an interest in solving it.

Note

If you do not consider a topic relevant, it is not worthwhile to continue rating it. In that case, you should drop it from your list.

2. Avoidance of duplication

Before you decide to carry out a study, it is important that you find out whether the suggested topic has been investigated before, either within the proposed study area or in another area with similar conditions. If the topic has been researched, the results should be reviewed to explore whether major questions that deserve further investigation remain unanswered. If not, another topic should be chosen.

Note

Also, consider carefully whether you can find answers to the problem in already available, unpublished information and from common sense. If so, you should drop the topic from your list.

3. Feasibility

Look at the project you are proposing and consider the complexity of the problem and the resources you will require to carry out your study. Thought should be given first to personnel, time, equipment, and money that are locally available.

In situations where the local resources necessary to carry out the project are not sufficient, you might consider resources available at the national level; for example, in research units, research councils, or local universities. Finally, explore the possibility of obtaining technical and financial assistance from external sources.

4. Political acceptability

In general it is advisable to research a topic that has the interest and support of the authorities. This will increase the chance that the results of the study will be implemented. Under certain circumstances, however, you may feel that a study is required to show that the government's policy needs adjustment. If so, you should make an extra effort to involve the policymakers concerned at an early stage, to limit the chances for confrontation later.

5. Applicability of possible results and recommendations

Is it likely that the recommendations from the study will be applied? This will depend not only on the blessing of the authorities but also on the availability of resources for implementing the recommendations. The opinion of the potential clients and of responsible staff will influence the implementation of recommendations as well.

6. Urgency of data needed

How urgently are the results needed for making a decision? Which research should be done first and which can be done later?

7. Ethical acceptability

We should always consider the possibility that we may inflict harm on others while carrying out research. Therefore, review the study you are proposing and consider important ethical issues such as:

- How acceptable is the research to those who will be studied? (Cultural sensitivity must be given careful consideration).
- Can informed consent be obtained from the research subjects?
- Will the condition of the subjects be taken into account? For example, if individuals are identified during the study who require treatment, will this treatment be given? What if such treatment interferes with your study results?

These criteria can be measured by the following rating scales:

What information should be included in the statement of the problem?

1. A **brief** description of socioeconomic and cultural characteristics and an overview of health status and the health-care system in the country or district **in as far as these are relevant to the problem**. Include a few illustrative statistics, if available, to help describe the **context** in which the problem occurs.

2. A concise **description** of the nature of the problem (the discrepancy between what is and what should be) and of its size, distribution, and severity (who is affected, where, since when, and what are the consequences for those affected and for the services?)

3. An **analysis** of the major factors that may influence the problem and a convincing argument that available knowledge is insufficient to solve it.

4. A brief description of any solutions that have been tried in the past, how well they have worked, and why further research is needed.

5. A description of the type of information expected to result from the project and how this information will be used to help solve the problem.

6. If necessary, a short list of definitions of crucial concepts used in the statement of the problem.

   A list of abbreviations may be annexed to the proposal, but each abbreviation also has to be written out in full when introduced in the text for the first time.

GROUP WORK

1. Select a reporter who will present the statement of the problem in plenary.

2. Discuss comments you received in the previous plenary session on the choice of your topic and revise your topic, if necessary.

3. Make an analysis diagram of the most important components of the problem or the most important factors that you think are influencing it. Use a blackboard or a flip chart and, if possible, separate cards for each factor. (See part I of this module for details on the steps in this process.) After making your initial diagram, try to rearrange the factors identified into broader categories.

SCALES FOR RATING RESEARCH TOPICS

Relevance

1. = Not relevant
2. = Relevant
3. = Very relevant

Avoidance of duplication

1. = Sufficient information already available
2. = Some information available but major issues not covered
3. = No sound information available on which to base problem-solving

Feasibility

1. = Study not feasible considering available resources
2. = Study feasible considering available resources
3. = Study very feasible considering available resources

Political acceptability

1. = Topic not acceptable to high level policymakers
2. = Topic more or less acceptable
3. = Topic fully acceptable

Applicability

1. = No chance of recommendations being implemented
2. = Some chance of recommendations being implemented
3. = Good chance of recommendations being implemented

Urgency

1. = Information not urgently needed
2. = Information could be used right away but a delay of some months would be acceptable
3. = Data very urgently needed for decision-making

Ethical acceptability

1. = Major ethical problems
2. = Minor ethical problems
3. = No ethical problems

Note: If you have already analyzed your problem in Module 4, skip the exercise and go straight to group work.

EXERCISE: The Chobe District Health Team, selecting a research project

(to be carried out in plenary, ½ hour, if this is the first discussion of possible research topics)

Introduction to the exercise

The Chobe District Health Team, responsible for the health of a population of 125,000, has to choose between two important study topics:

Possibility 1

The first possibility is a study into methods for motivating communities to provide voluntary labour for the installation of water systems.

In Chobe District, streams are used as latrines as well as sources of domestic water. Morbidity surveys show an extremely high prevalence of diarrhea and chronic infections with intestinal parasites. UNICEF has offered to supply free plastic pipes if the villagers will provide labour to install community water systems from protected springs.

Because the terrain is rocky, it will take a great deal of labour to dig trenches in which to lay the plastic pipes. Burying the pipes would seem necessary as it is not uncommon that villagers cut into exposed pipes to obtain water. The motivation among male villagers to dig the trenches, however, is not high; the belief that water has a purifying power and that anything dissolved in the streams cannot possibly be dangerous appears to be a stumbling block to increasing motivation.

The District Health Team, encouraged by UNICEF to take action and aware that in pilot projects in neighbouring districts the population was successfully motivated, now wants to take action. So far, invitations to village leaders to attend training programs designed to demonstrate how they could develop and maintain their own water systems remain unanswered.

Proposed study: The District Health Team proposes to undertake a rapid assessment in four villages, two in the pilot project located in the neighbouring district and two in Chobe District, to find out:

- What factors have contributed to the involvement of the community in the project in the neighbouring district;
- Whether it would be feasible to increase the population's interest in the project by providing more detailed information on the relationship between contaminated water and disease;
- Whether it would be possible to keep the burying of pipes to a minimum, if the *whole* population (males and females, youngsters and adults) were involved in the project, and representatives of all these groups participated in the village water committee.

The team would plan to interview project authorities in the neighbouring district and conduct three focus group discussions in each village: one with males, one with females, and one with males and females combined, to explore the questions above.

EXERCISE (continued)

Possibility 2

The second possibility is to examine the reasons for the assumingly increasing perinatal mortality among children delivered at the District Hospital. Various community members have expressed their concern over expectant mothers returning home from the District Hospital without babies. They are demanding an explanation from the health workers before they approach the government with the problem.

The District Health Team wishes to prevent the community from approaching the politicians. First of all it wants to assess whether the perinatal mortality among children born at the District Health Centre has indeed gone up over the past 5 years and, if so, how this could be explained.

Proposed Study: The District Health Team would plan to analyze the records of the maternity ward over the past 10 years to investigate whether there indeed has been an upward trend in the proportion of deaths. What is the cause of each recorded death? Could some have been prevented either by more intensive care in the maternity ward or by earlier prenatal care and referral of high risk cases by TBAs and peripheral units? What other reasons may there have been for the deaths?

In addition to the record review, the District Team would plan to interview maternity staff in the District Hospital and in five peripheral health units. Also, TBAs would be interviewed and focus group discussions would be held with women in the age group of 15-45 years in five villages.

Directions

Rate the two proposals in small groups, using the form on the following page, and prepare to defend your first choice in plenary. (When rating the topics on the criteria, you can either refer to the "Scales for Rating Research Topics" presented right before this exercise or use the summary scales at the bottom of the rating sheet.)

EXERCISE (continued)

Proposed topic	1. Relevance	2. Avoidance of duplication	3. Feasibility	4. Political acceptability	5. Applicability	6. Urgency of data needed	7. Ethical acceptability	Total
1. Community water systems	2	2	1	1	2	1	3	13
2. Perinatal mortality	3	3	3	3	3	3	3	21

Rating scale: 1 = low, 2 = medium, 3 = high.

Literature Review

Why is it important to review already available information when preparing a research proposal?

- It prevents you from duplicating work that has been done before.
- It helps you to find out what others have learned and reported on the problem you want to study. This may assist you in refining your statement of the problem.
- It helps you to become more familiar with the various types of methodology that might be used in your study.
- It should provide you with convincing arguments for why your particular research project is needed.

What are the possible sources of information?

- Individuals, groups, and organizations;
- Published information (books, articles, indexes, and abstract journals); and
- Unpublished information (other research proposals in related fields, reports, records, and computer data bases)

Where can we find these different sources?

Different sources of information can be consulted and reviewed at various levels of the administrative system within your country and internationally.

Administrative level	Examples of resources
Community and district or provincial levels	• Clinic and hospital based data from routine statistics, registers; • Opinions, beliefs of key figures (through interviews); • Clinical observations, reports of critical incidents, etc.; • Local surveys, annual reports; • Statistics issued at provincial, and district levels; • Books, articles, newspapers, mimeographed reports, etc.
National level	• Articles from national journals, books identified during literature searches at university and other national libraries, WHO, UNICEF libraries, etc. • Documentation, reports, and raw data from: - The ministry of health - Central statistical offices - Nongovernmental organizations

International level

- Information from:
  - Bilateral and multilateral organizations (e.g., IDRC, USAID, UNICEF, WHO),
  - Computerized searches for international literature (from national library or international institutions).

You need to develop a strategy to gain access to each source and to obtain information in the most productive manner. Your strategy may vary according to where you work and the topic under study. It may include the following steps:

- Identifying a key person (researcher or decision-maker) who is knowledgeable on the topic and asking if he or she can give you a few good references or the names of other people whom you could contact for further information;
- Looking up the names of speakers on your topic at conferences who may be useful to contact;
- Contacting librarians in universities, research institutions, the ministry of health, and newspaper offices and requesting relevant references;
- Examining the bibliographies and reference lists in key papers and books to identify relevant references;
- Looking for references in indexes (e.g., *Index Medicus*, see Annex 5.1) and abstract journals (see Annex 5.2); and
- Requesting a computerized literature search (e.g., Medline, see Annex 5.3).

Some agencies will assist with your literature search if requested by telephone or in writing. The request, however, should be very specific. Otherwise you will receive a long list of references, most of which will not be relevant to your topic. If you are requesting a computerized search it is useful to suggest key words that can be used in locating the relevant references.

Note:

Facilitators should be able to provide specific information regarding national and international facilities to assist you with the search for literature.

References that are identified:

- **Should first be skimmed or read.**
- **Then summaries of the important information in each of the references should be recorded on separate index cards (Annex 5.4) or as computer entries. These should then be classified so that the information can easily be retrieved.**
- **Finally a literature review should be written.**

Information on an index card should be organized in such a way that you can easily find all data you will need for your report.

For an article, the following information should be noted:

Author(s) (surname followed by initials). Title of article. Name of journal, year; volume number: page numbers of article.

Example:
Gwebu ET, Mtero S, Dube N, Tagwireyi JT, Mugwagwa N. Assessment of nutritional status in pregnancy: use of a reference table of weight-for-height. Central African Journal of Medicine, 1985; 31: 193-196.

For a book, the following information should be noted:

Author(s) (surname followed by initials). Title of book. Edition. Place: Publisher, year: number of pages in the book.

Example:
Abramson JH. Survey methods in community medicine. 2nd ed. Edinburgh: Churchill Livingstone, 1979: 229.

For a chapter in a book, the citation can include:

Author(s) of chapter (surname followed by initials). Chapter title. In: Editors of book (surname followed by initials). eds., Title of book. Place: Publisher, year: page numbers of chapter.

Example:
Winikoff B, Castle MA. The influence of maternal employment on infant feeding. In: Winikoff B, Castle MA, Laukaran VH, eds. Feeding infants in four societies: causes and consequences of mother's choices. New York: Greenwood Press, 1988: 121-145.

This information, recorded in a standard format such as that suggested above, can then easily be used as part of your list of references for the proposal. The formats suggested above have been adopted as standard by over 300 biomedical journals and sometimes referred to as "the Vancouver System." For more information, see International Committee of Medical Journal Editors (1988). Other references in this series follow IDRC's house style.

The index card or computer entry (one for each reference) could contain quotations and information such as:

- Key words;
- A summary of the contents of the book or the article, concentrating on information relevant to your study; and
- A brief analysis of the content, with comments such as:

  - Appropriateness of the methodology;
  - Important aspects of the study; and
  - How information from the study can be used in your research.

Note

Index cards or computer entries can also be used to summarize information obtained from other sources, such as informal discussions, reports of local health statistics, and internal reports.

How do you write a review of literature?

There are a number of steps you should take when preparing a review of available literature and information:

- First, organize your index cards in groups of related statements according to which aspect of the problem they touch upon.

- Then, decide in which order you want to discuss the various issues. If you discover you have not yet found literature or information on some aspects of your problem that you suspect are important, make a special effort to find this literature.

- Finally, write a coherent discussion of one or two pages **in your own words**, using all relevant references. You can use consecutive numbers in the text to refer to your references. Then list your references in that order, using the format described in the section above on index cards. Add this list as an annex to your research proposal.

  Alternatively, you can refer to the references more fully in the text, putting the surname of the author, year of publication, and number(s) of page(s) referred to between brackets, e.g., (Shiva 1988: 15-17). If this system of citation is used, the references at the end of the proposal should be listed in alphabetical order.

Possible bias

Bias in the literature or in a review of the literature is a distortion of the available information in such a way that it reflects opinions or conclusions that do not represent the real situation.

It is useful to be aware of various types of bias. This will help you to be critical of the existing literature. If you have reservations about certain references, or if you find conflicting opinions in the literature, discuss these openly and critically. Such a critical attitude may also help you avoid biases in your own study. Common types of bias in literature include:

- **Playing down controversies and differences in one's own study results;**
- **Restricting references to those that support the point of view of the author; and**
- **Drawing far reaching conclusions from preliminary or shaky research results or making sweeping generalizations from just one case or small study.**

Annex 5.4. Example of a reference recorded on an index card.

Hassouna WA. Solving people's problems. World Health, 1980; April: 26-29.

- This article discusses health services research (HSR) as a relatively new area of investigation (1960).
- This method of research permits the health team and the community to study critical problems, while economizing on time and money. Important to try to collaborate with service administrators.
- If HSR is to be effective, must be done so results available in time to solve problems it addresses – change in health status, not publication, most important result of research.
- Example of HSR study in Maruit (Egypt):
  - In 2 days a multidisciplinary team (25 members) was able to identify the critical problems affecting health and health care in area.
  - Various aspects of the study are discussed.
  - The study results are stated clearly and the role of the traditional healer identified.
  - Among major findings was that "the formal providers of health services were not giving the people the service they required at the time they needed it, at a cost they could pay, and in a manner acceptable to the people." (p. 27)

The reverse side of the index card appears below:

- Points that are emphasized in the article:
  - There's little correlation between size and quality of health services available to population and health status of population (p. 28). Problem is present nature of medical technology.
  - Use of med. technology to improve health status would be more successful if became integral part of socio-cultural and ec. behavioural change process. (p. 28)
  - Article lists characteristics and advantages of PHC and role of community in it.
  - Discusses importance of HSR related to PHC - conviction HSR should form core of WHO "Health for All by the Year 2000" strategy.
  - Important to involve WHO staff in field activities so acquire practical understanding of health service realities.
- Observations:

  Good reference article on applied research, PHC, and research training.

Annex Sample references.

1. Taylor CE. The uses of health systems research. Geneva: WHO, 1984. Public Health Papers 78.

2. Illsley R. Introduction to HSR. In: Health systems research in action. Programme on Health Systems Research and Development. Geneva: WHO, 1988.

3. Bryant Y. Health and the developing world. Ithaca: Cornell University Press, 1969.

4. Health Systems Research Advisory Group. First Meeting, Geneva, 7-10 April 1986. Report and Working Document. Geneva: WHO, 1986.

5. Foster GM, Anderson GE. Medical anthropology. New York: John Wiley and Sons, 1978.

6. Kleinman A. Concepts and a model for the comparison of medical systems as cultural systems. Social Science and Medicine, 1978; 12: 85-93.

7. White KL, Henderson MM (eds.). Epidemiology as a fundamental science: its uses in health services planning, administration and evaluation. New York: Oxford University Press, 1976.

8. Knox EG (ed.). Epidemiology in health care planning. New York: Oxford University Press, 1979.

9. Kwofie K. The process of introducing nutrition objectives into rural and agricultural development: lessons from the Baringo experiment. Lusaka, Kenya: National Food and Nutrition Commission, 1979.

10. Yambi O. Nutritional problems and policies in Tanzania. Ithaca, NY: Cornell Institute, 1980. Monograph no. 7.

11. Gish O, Walker G. Mobile health services. London: Tri-Med Bodus Ltd, 1977.

Research objectives

The objectives of a research project summarize what is to be achieved by the study.

Objectives should be closely related to the statement of the problem. For example, if the problem identified is low utilization of child welfare clinics, the general objective of the study could be to identify the reasons for this low utilization, to find solutions.

The general objective of a study states what is expected to be achieved by the study in general terms.

It is possible (and advisable) to break down a general objective into smaller, logically connected parts. These are normally referred to as specific objectives.

Specific objectives should systematically address the various aspects of the problem as defined under "Statement of the problem" (Module 4) and the key factors that are assumed to influence or cause the problem. They should specify **what** you will do in your study, **where,** and **for what purpose**.

The general objective "to identify the reasons for low utilization of child welfare clinics in District X to find solutions," for example, could be broken down into the following specific objectives:

1. Determine the level of utilization of the child welfare clinics in District X, over the years 1988 and 1989, as compared with the target set.

2. Identify whether there are variations in utilization of child welfare clinics, related to the season, type of clinic, and type of children served.

3. Identify factors related to the child welfare services offered that make them either attractive or not attractive to mothers. This objective may be divided into smaller subobjectives focusing on distance between the home and clinic, acceptability of the services to mothers, quality of the services, etc.

4. Identify socioeconomic and cultural factors that may influence the mothers' utilization of services. (Again, this objective may be broken down into several subobjectives.)

5. Make recommendations to all parties concerned (managers, health staff, and mothers) concerning what changes should be made, and how, to improve the use of child welfare clinics.

6. Work with all parties concerned to develop a plan for implementing the recommendations.

The first objective focuses on quantifying the problem. This is necessary in many studies. Often use can be made of available statistics or of the health information system.

Objective 2 further specifies the problem, looking at its distribution. Objectives 3 and 4 examine possible factors that may influence the problem, and objectives 5 and 6 indicate how the results will be used.

> **Note:**
>
> **An objective focusing on how the results will be used should be included in every applied research study.**

Why should research objectives be developed?

The formulation of objectives will help you to:

- **Focus** the study (narrowing it down to essentials);
- **Avoid** collection of data that are not strictly necessary for understanding and solving the problem you have identified; and
- **Organize** the study in clearly defined parts or phases.

Properly formulated, specific objectives will facilitate the development of your research methodology and will help to orient the collection, analysis, interpretation, and utilization of data.

How should you state your objectives?

Take care that the objectives of your study:

- Cover the different aspects of the problem and its contributing factors in a **coherent** way and in a **logical sequence**;
- Are **clearly phrased** in **operational terms,** specifying exactly what you are going to do, where, and for what purpose;
- Are **realistic** considering local conditions; and
- Use **action verbs** that are specific enough to be evaluated.

  Examples of action verbs are: to determine, to compare, to verify, to calculate, to describe, and to establish.

  Avoid the use of vague nonaction verbs such as: to appreciate, to understand, or to study.

Keep in mind that when the project is evaluated, the results will be compared to the objectives. If the objectives have not been spelled out clearly, the project cannot be evaluated.

Using the previous example on utilization of child welfare clinics, we may develop more specific objectives such as:

- **To compare** the level of utilization of the child welfare clinic services among various socioeconomic groups;
- **To establish** the pattern of utilization of child welfare clinic services in various seasons of the year;
- **To verify** whether increasing distance between the home and the health facility reduces the level of utilization of the child welfare clinic services;
- **To describe** mothers' perceptions of the quality of services provided at the child welfare clinics.

Hypotheses

Based on your experience with the study problem, it might be possible to develop explanations for the problem that can then be tested. If so, you can formulate hypotheses in addition to the study objectives.

> A HYPOTHESIS is a **prediction of a relationship between one or more factors** and the problem under study, which can be **tested.**

n our example concerning the low utilization of child welfare clinics, it would be possible to formulate and test the following hypotheses:

1. Utilization of child welfare clinics is lowest in the rainy season due to the high workload of mothers during that period.
2. Utilization of child welfare clinics is lowest in those clinics in which staff are poorly motivated to provide preventive services.

> **Note:**
>
> **Policymakers and field staff usually feel the need for research because they do NOT have enough insight into the causes of a certain problem. Therefore, most HSR proposals present the specific objectives in the form of open statements (as given in the examples earlier) instead of focusing the study on a limited number of hypotheses.**

Design of interview schedules or questionnaires

I. INTRODUCTION

Interviews and self-administered questionnaires are probably the most commonly used research techniques. Therefore, designing good "questioning tools" forms an important and time-consuming phase in the development of most research proposals.

Once the decision has been made to use these techniques, the following questions should be considered before designing our tools:

- What exactly do we want to know, according to the objectives and variables we identified earlier? Is questioning the right technique to obtain all answers, or do we need additional techniques, such as observations or analysis of records?
- Of whom will we ask questions and what techniques will we use? Do we understand the topic sufficiently to design a questionnaire, or do we need some loosely structured interviews with key informants or a FGD first to orientate ourselves?
- Are our informants mainly literate or illiterate? If illiterate, the use of self-administered questionnaires is not an option.
- How large is the sample that will be interviewed? Studies with many respondents often use shorter, highly structured questionnaires, whereas smaller studies allow more flexibility and may use questionnaires with a number of open-ended questions.

II. TYPES OF QUESTIONS

Before examining the steps in designing a questionnaire, we need to review the types of questions used in questionnaires. Depending on how questions are asked and recorded we can distinguish two major possibilities:

- open-ended questions, and
- closed questions.

Open-ended questions

OPEN-ENDED QUESTIONS permit free responses that should be recorded in the respondent's own words. The respondent is not given any possible answers to choose from.

Such questions are useful to obtain information on:

- Facts with which the researcher is not very familiar,
- Opinions, attitudes, and suggestions of informants, or
- Sensitive issues.

For example

"Can you describe exactly what the traditional birth attendant did when your labour started?"

"What do you think are the reasons for a high drop-out rate of village health committee members?

"What would you do if you noticed that your daughter (school girl) had a relationship with a teacher?

Closed questions

CLOSED QUESTIONS offer a list of possible options or answers from which the respondents must choose.

When designing closed questions one should try to:

- Offer a list of options that are exhaustive and mutually exclusive, and
- Keep the number of options as few as possible.

Closed questions are useful if the range of possible responses is known.

For example

"What is your marital status?"

1. Single ○
2. Married/living together ○
3. Separated/divorced/widowed ○

"Have your ever gone to the local village health worker for treatment?

1. Yes ○
2. No ○

Closed questions may also be used if one is only interested in certain aspects of an issue and does not want to waste the time of the respondent and interviewer by obtaining more information than one needs.

For example, a researcher who is only interested in the protein content of a family diet may ask:

"Did you eat any of the following foods yesterday?" (circle yes or no for each set of items)

- Peas, bean, lentils Yes No
- Fish or meat Yes No
- Eggs Yes No
- Milk or cheese Yes No

Closed questions may be used as well to get the respondents to express their opinions by choosing rating points on a scale.

For example

"How useful would you say the activities of the Village Health Committee have been in the development of this village?"

1. Extremely useful ○
2. Very useful ○
3. Useful ○
4. Not very useful ○
5. Not useful at all ○

Using attitudes scales is advisable only in face-to-face interviews with literates if the various options for each answer are provided for the respondents on a card they can look at while making their choice. If the researcher only reads the options, the respondents might not consider all options equally and the scale will not accurately measure the attitudes.

Table 10B.1. Advantages and disadvantages of open-ended and closed questions and conditions for optimal use.

Open-ended questions	Closed questions
Advantages Issues not previously thought of when planning the study may be explored, thus providing valuable new insights into the problem. Information provided spontaneously is likely to be more valid than answers suggested in options from which the informant must choose. Information provided in the respondents' own words may be useful as examples or illustrations that add interest to the final report.	**Advantages** Answers can be recorded quickly. Analysis is easy.
Disadvantages Skilled interviewers are needed to get the discussion started and focused on relevant issues and to record all important information. Analysis is time-consuming and requires experience.	**Disadvantages** Closed questions are less suitable for face-to-face interviews with nonliterates. Respondents may choose options they would not have thought of themselves (leading questions → bias). Important information may be missed if it is not asked. The respondent and interviewer may lose interest after a number of closed questions.

Open-ended questions	Closed questions
Suggestions	**Suggestions**
Thoroughly train and supervise the interviewers or select experienced people.	Use closed questions only on issues that are simple.
Prepare a list of further questions to keep at hand to use to "probe" for answer(s) in a systematic way.	Pretest closed questions first as open-ended questions to see if your categories cover all possibilities.
Pretest open-ended questions and, if possible, pre-categorize the most common responses, leaving enough space for other answers.	Use closed questions in combination with open-ended questions.

In practice, **a questionnaire usually has a combination of open-ended and closed questions**, arranged in such a way that the discussion flows as naturally as possible.

In interviews questions are often asked as **open-ended questions**, but to facilitate recording and analysis, possible answers are to a large extent **pre-categorized**.

For example

"How did you become a member of the Village Health Committee?"

1. Volunteered ○
2. Elected at a community meeting ○
3. Nominated by community leaders ○
4. Nominated by the health staff ○
5. Other (specify): ○

_____

_____

With this type of half open-ended, half closed question strict guidelines have to be provided and followed!

- In general, such a question should be asked as an OPEN question: NO OPTIONS should be provided. Sometimes it may be useful to probe for an answer: then all interviewers should follow the same guidelines (for example, using the same types of probes).

  (If the question is asked in different ways by different interviewers, you get BIAS.)

- The interview guide or questionnaire should indicate whether the informant can give more than one answer to a question.

For open-ended questions, more than one answer is usually allowed. The interviewers will have to be trained to wait for additional answers. They should also be instructed not merely to tick the options mentioned, but to record any additional information a respondent may provide.

Note

Sometimes it is useful, especially in small-scale studies, to use pictures or drawings when asking certain questions to get the discussion going. In the case of illiterates, a questionnaire may even consist exclusively of pictures. (See Annex 10B.1.)

III. STEPS IN DESIGNING A QUESTIONNAIRE[1] [2]

Designing a good questionnaire always takes several drafts. In the first draft we should concentrate on the **content**. In the second, we should look critically at the **formulation and sequencing of the questions.** Then we should scrutinize the **format** of the questionnaire. Finally, we should do a **test-run** to check whether the questionnaire gives us the information we require and whether both we and the respondents feel at ease with it. Usually the questionnaire will need some further adaptation before we can use it for actual data collection.

Step 1: Content

Take your objectives and variables as your starting point.

Decide what questions will be needed to measure or to define your variables and reach your objectives.

When developing the questionnaire, you should reconsider the variables you have chosen, and, if necessary, add, drop or change some. You may even change some of your objectives at this stage.

Step 2: Formulating questions

Formulate one or more questions that will provide the information needed for each variable.

Take care that questions are specific and precise enough that different respondents do not interpret them differently. For example, a question such as: "Where do community members usually seek treatment when they are sick?" cannot be asked in such a general way because each respondent may have something different in mind when answering the question:

- One informant may think of measles with complications and say he goes to the hospital, another of cough and say he goes to the private pharmacy;
- Even if both think of the same disease, they may have different degrees of seriousness in mind and thus answer differently;
- In all cases, self-care may be overlooked.

---

[1] For the sake of simplicity we take questionnaires as an example. The same steps apply to designing more loosely structured interview schedules and checklists.

[2] This section is largely adapted from Sudman Bradman (1983).

The question, therefore, as a rule has to be broken up into different parts and made so specific that all informants focus on the same thing. For example, one could:

- Concentrate on illness that has occurred in the family over the past 14 days and ask what has been done to treat it from the onset; or

- Concentrate on a number of diseases, ask whether they have occurred in the family over the past X months (chronic or serious diseases have a longer recall period than minor ailments) and what has been done to treat each of them from the onset.

Check whether each question measures one thing at a time.

For example, the question, "How large an interval would you and your husband prefer between two successive births?" would better be divided into two questions because husband and wife may have different opinions on the preferred interval.

Avoid leading questions.

A question is leading if it suggests a certain answer. **For example**, the question, "Do you agree that the district health team should visit each health centre monthly?" hardly leaves room for "no" or for other options. Better would be: "Do you think that district health teams should visit each health centre? If yes, how often?"

Sometimes, a question is leading because it presupposes a certain condition. **For example**: "What action did you take when your child had diarrhea the last time?" presupposes the child has had diarrhea. A better set of questions would be: "Has your child had diarrhea? If yes, when was the last time?" "Did you do anything to treat it? If yes, what?"

Formulate control questions to cross-check responses on "difficult" questions (sensitive questions or questions for which it is difficult to get a precise answer).

Avoid words with double or vaguely defined meanings and emotionally laden words. Concepts such as nasty (health staff), lazy (patients), or unhealthy (food), for example, should be omitted.

Step 3: Sequencing of questions

Design your interview schedule or questionnaire to be "consumer friendly."

- The sequence of questions must be logical for the respondent and allow as much as possible for a "natural" discussion, even in more structured interviews.

- At the beginning of the interview, keep questions concerning "background variables" (e.g., age, religion, education, marital status, or occupation) to a minimum. If possible, pose most or all of these questions later in the interview. (Respondents may be reluctant to provide "personal" information early in an interview and, if they become worried about confidentiality, be wary about giving their true opinions.)

- Start with an interesting but noncontroversial question (preferably open) that is directly related to the subject of the study. This type of beginning should help to raise the informants' interest and lessen suspicions concerning the purpose of the interview (e.g., that it will be used to provide information to use in levying taxes).

- Pose more sensitive questions as late as possible in the interview (e.g., questions pertaining to income, political matters, sexual behaviour, or diseases with stigma attached to them).

- Use **simple, everyday language.**

Make the questionnaire as short as possible. Conduct the interview ~~in~~ two parts if the nature of the topic requires a long questionnaire (more than 1 hour).

Step 4: Formatting the questionnaire

When you finalize your questionnaire, be sure that:

- Each questionnaire has a heading and space to insert the number, date, and location of the interview, and, if required, the name of the informant. You may add the name of the interviewer to facilitate quality control.

- Layout is such that questions belonging together appear together visually. If the questionnaire is long, you may use subheadings for groups of questions.

- Sufficient space is provided for answers to open-ended questions.

- Boxes for pre-categorized answers are placed in a consistent manner (e.g., on the right half of the page). (See examples in this module.)

- If you use a computer, the right margin of the page should be reserved for boxes intended for computer codes. (See Module 13 and consult an experienced facilitator when designing your questionnaire.)

Your questionnaire should not only be *consumer* but also *user friendly*!

Step 5: Translation

If interviews will be conducted in one or more local languages, the questionnaire has to be translated to standardize the way questions will be asked.

After having it translated you should have it retranslated into the original language. You can then compare the two versions for differences and make a decision concerning the final phrasing of difficult concepts.

Work plan

I. INTRODUCTION

What is a work plan?

A WORK PLAN is a schedule, chart, or graph that summarizes, in a clear fashion, various components of a research project and how they fit together.

It may include:

- The tasks to be performed;
- When the tasks will be performed; and
- Who will perform the tasks and the time each person will spend on them.

II. VARIOUS WORK SCHEDULING AND PLANNING TECHNIQUES

1. The work schedule

A WORK SCHEDULE is a table that summarizes the tasks to be performed in a research project, the duration of each activity, and the staff responsible.

The version of a work schedule given on the following page includes:

- The tasks to be performed;
- The dates each task should begin and be completed;
- Research team, research assistants, and support staff (drivers and typists) assigned to the tasks; and
- Person-days required by research team members, research assistants, and support staff (the number of person-days equals the number of working days per person).

Note:

The period for field research for the course project should not exceed 6 months. Week 1 is the first week after completion of the present workshop.

This work schedule was developed for a study of factors contributing to low utilization of child-spacing (C/S) services in a certain region. The research team consisted of four persons (mainly regional health team members). The study consisted of two main parts: (1) analysis of the child-spacing records to assess the percentage of C/S users and the regularity with which they use the services, and interviews with staff responsible for the services; and (2) interviews with female users of C/S services (sampled from the records) and nonusers, and interviews with husbands of female users of C/S and of nonusers.

EXAMPLE OF WORK A SCHEDULE: CHILD-SPACING STUDY (C/S)

Tasks to be performed	Dates	Personnel assigned to task	Person days required
1. Finalize research proposal and. literature review	week 1-3 4-24 Apr.	Research team (4)	4 × 3 = 12 days
2. Clearance from national and funding authorities	week 1-5 4 Apr.-8 May	Research unit - ministry of health	
3. Clearance and orientation of local authorities	week 6 9-15 May	PI (Regional Health Officer) Driver	2 days 2 days
4. Compilation of child spacing records and interviews of C/S staff	week 6-9 9 May-5 June	Public health nurse Driver	10 days 10 days
5. Analysis of C/S records and sampling study units	week 10 6-12 June	Research team Secretary	4 × 2 = 8 days 1 day
6. Training of research assistants and field testing questionnaire	week 11 13-19 June	Research team Research assistant(s) Facilitator	4 × 3 = 12 days 5 × 3 = 15 days 1 × 4 = 4 days
7. Interviews in community	week 12-13 20 June-3 July	Research team Reasearch assistants	4 × 10 = 40 days 5 × 10 = 50 days
8. Preliminary data analysis	week 19-22 8-28 Aug.	Research team Research assistants Facilitator	4 × 7 = 28 days 5 × 1 = 5 days 1 × 2 = 2 days
9. Feedback to local authorities and district health teams	week 27 3-9 Oct.	Research team Driver	4 × 1 = 4 days 2 days
10. Feedback to communities	week 28 10-16 Oct.	Research team Driver	4 × 1 = 4 days 1 day
11. Data analysis and reporting workshop	week 29-30 17-30 Oct.	Research team Facilitator	4 × 10 = 40 days 1 × 10 = 10 days
12. Report finalization	week 31-34 31 Oct.-28 Nov.	Research team Secretary	4 × 2 = 8 days 1 × 5 = 5 days
13. Discussion of recommendations/plan of action with local authorities and districy health teams	week 36-37 12-25 Dec.	Research team Secretary Driver	4 × 3 = 12 days 3 days 3 days
14. Monitoring research project	continuous	Research team	4 × 1 = 4 days

You will notice that, if the workshops are excluded, each team member roughly spent 30 working days on the research, except the regional public health nurse. She visited all centres with C/S services in the region to analyze the records and interview staff. Although she integrated these tasks with her normal supervisory duties, she spent about 10 working days more than the other team members. Five research assistants (two community health nurses and three district health Inspectors) were recruited to assist with the interviewing. The number of working days required was multiplied by four (for the research team) and five (for the research assistants) to arrive at the number of person-days.

How to develop a work schedule

- Review and revise, if necessary, the list of tasks you prepared for your plan for data collection (Module 12). Add to the list other tasks you must complete not related to data collection (such as clearance of proposal; data analysis and report writing; and feedback to authorities and target group). Number all tasks.

- Now review the staffing for the different tasks, taking into account your experience during the pretest. Consider:

  - Who will carry out which tasks;
  - The amount of time needed per research unit (interview/observation/record) including travel time; and
  - The number of staff needed to complete each task in the planned period of time.

  Make revisions, if required. Complete the staffing for the tasks you have just added.

- Consider whether the use of short-term consultants is necessary for certain tasks. Always consider using local consultants. If consultants are used, involve them in the planning stage of the project so you can incorporate any useful suggestions they may have concerning the design of the methodology.

In reviewing your tentative staffing plan you should ask:

- Are the types of personnel and levels of expertise you require likely to be available for the project? For example, is there a sufficient range of disciplines available including, where appropriate, personnel from outside the health field?

- If special staff have to be recruited or reassigned from other ministries or agencies, what regulations or procedures will have to be followed?

- Is the staffing plan realistic, taking into account the project budget that is likely to be available?

- To what extent can community members, traditional healers, students, or other nonprofessionals be involved in the study?

- What training would the research assistants or data collectors require? How long would the training last? Who would do the training? How do you intend to supervise the assistants and data collectors? Review what you have tentatively planned in Module 12 and revise it, as necessary.

Then fix the dates (in weeks) indicating the period in which each task will have to be carried out and calculate the number of working days per person required to complete each task.

2. The GANTT chart

The GANTT chart is a planning tool which depicts graphically the order in which various tasks must be completed and the duration of each activity.

The GANTT chart shown on the following page indicates:

- the tasks to be performed;
- who is responsible for each task; and
- the time each task is expected to take.

The length of each task is shown by a bar that extends over the number of days, weeks or months the task is expected to take.

How can a work plan be used?

A work plan can serve as:

- A tool in planning the details of the project activities and later in budgeting funds.
- A visual outline or illustration of the sequence of project operations. It can facilitate presentations and negotiations concerning the project with government authorities and other funding agencies.
- A management tool for the principal investigator and members of his or her team, showing what tasks and activities are planned, their timing, and when various staff members will be involved in various tasks.
- A tool for monitoring and evaluation, when the current status of the project is compared to what had been foreseen in the work plan.

When should the work plan be prepared and when should it be revised?

- The first draft of the work plan should be prepared when the project proposal is being developed, so the schedule can be discussed easily with the relevant authorities.
- A more detailed work plan should be prepared after the pretest in the study area.

Example of a GANTT chart for the child spacing study.

Tasks to be performed	Responsible person	April	May	June	July	Aug	Sept	Oct	Nov	Dec
1. Finalize research proposal	Research team									
2. Clear national authorities	Research unit MOH									
3. Clear and orient local authorities	PI									
4. Compile CS records and interview CS staff	Reg. PH nurse									
5. Analyze CS records and sample study units	Research team									
6. Train research assistants and field-test questionnaire	Research team, facilitator									
7. Interviews in community	Research team, research assistants									
8. Preliminary data analysis	Research team, research assistants, facilitator									
9. Feedback to local authorities and district health teams	Research team									
10. Feedback to communities	Research team									
11. Data analysis and report-writing workshop	Research team, facilitator									
12. Finalize report	Research team									
13. Discuss recommendations/plan of action with local authorities and district health teams	Research team									
14. Monitor research projections	Research team									

- There should be no hesitation in revising work plans or preparing new ones after the project is underway based on a reassessment of what can be realistically accomplished in the coming months.

What factors should be kept in mind when preparing a work plan?

- It should be simple, realistic, and easily understood by those directly involved.
- It should cover the preparatory and the implementation phases of the project, as well as data analysis, reporting, and dissemination/utilization of results.
- The activities covered should include technical or research tasks; administrative, secretarial, and other support tasks; and training tasks.
- The realities of local customs (local holidays, festivals) and working hours should be considered when preparing the work plan.
- Also seasonal changes and their effect on travel, work habits, and on the topic you are studying (such as incidence of disease or nutritional status) should be kept in mind as the schedule is planned.

Plan for data collection

I. INTRODUCTION

Where are we now in the development of our research proposal?

Look again at the diagram in Module 7 that introduces the research methodology. We have just finished four crucial theoretical sessions, in which we have defined:

- what information we want to collect to answer the research questions implied in our objectives **(Module 8: Variables)**
- what **approach** we will follow to collect this information **(Module 9: Study type)**
- what **techniques** and **tools** we will use to collect it **(Module 10: Data-collection techniques)**
- **where** we want to collect the data, **how** we will select our sample, and **how many** subjects we will include in our study **(Module 11: Sampling)**

Now we enter a new phase in the development of our research methodology: planning our fieldwork. We have to plan concretely **how** we will **collect** the data we need **(Modules 12 and 15)**, how we will **analyze** it **(Module 13)**, and how we can **test** the most crucial parts of our methodology **(Module 14)**. Finally, we will have to develop a **plan** for project administration and monitoring **(Module 16)** and to budget the **resources** necessary to carry out the study **(Module 17)**.

A PLAN FOR DATA COLLECTION can be made in two steps:

1. Listing the tasks that have to be carried out and who should be involved, making a rough estimate of the time needed for the different parts of the study, and identifying the most appropriate period in which to carry out the research.

2. Actually scheduling the different activities that have to be carried out each week in a workplan.

Before the workshop is finished, a pretest of the data collection and data analysis procedures should be made. The advantages of conducting the pretest **before** we finalize our proposal is that we can draft the workplan and budget based on realistic estimates, as well as revise the data collection tools before we submit the proposal for approval.

However, if this is not possible (for example, because the proposal is drafted far from the field, and there are no similar research settings available close to the workshop site), the field test may be done after finishing the proposal, but long enough before the actual fieldwork to allow for a thorough revision of data collection tools and procedures.

Why should you develop a plan for data collection?

A plan for data collection should be developed so that:

- you will have a clear overview of what tasks have to be carried out, who should perform them, and the duration of these tasks;

- you can organize both human and material resources for data collection in the most efficient way; and

- you can minimize errors and delays that may result from lack of planning (for example, the population not being available or data forms being misplaced).

It is likely that while developing a plan for data collection you will identify problems (such as limited manpower) that will require modifications to the proposal. Such modifications might include adjustment of the sample size or extension of the period for data collection.

II. STAGES IN THE DATA-COLLECTION PROCESS

What are the main stages in the data-collection process?

Three main stages can be distinguished in the data-collection process:

Stage 1: PERMISSION TO PROCEED
Stage 2: DATA COLLECTION
Stage 3: DATA HANDLING

Stage I: Permission to proceed

Consent must be obtained from the relevant authorities, individuals, and the community in which the project is to be carried out. This may involve organizing meetings at national or provincial level, at district, and at village level. For clinical studies this may also involve obtaining written informed consent.

Most likely the principal investigator will be responsible for obtaining permission to proceed at the various levels. The health research unit in the ministry of health or the institution organizing the course may assist in obtaining permission from the national level.

Note:

In many countries research proposals have to be screened for scientific and ethical integrity by national research councils. However, proposals developed during workshops may be exempted from this procedure if the research is considered as a training exercise and the research council assumes that the course facilitators have screened the methodology during the workshop.

Stage II: Data collection

When collecting our data, we have to consider:

- Logistics: who will collect what, when, and with what resources; and
- Quality control.

1. Logistics of data collection

WHO will collect WHAT data?

When allocating tasks for data collection, it is recommended that you first list them. Then you may identify who could best implement each of the tasks. If it is clear beforehand that your research team will not be able to carry out the entire study by itself, you might look for research assistants to assist in relatively simple but time-consuming tasks.

For example, in a study into the effects of improvements in delivery care on utilization of these services the following task division could be proposed:

Task	To be carried out by
Record study	Research team
Focus group discussions with health staff before and after individual staff interviews	Research team
Individual health staff interviews	Research team
Shadowing MCH nurses	Principal investigator
Interviews with mothers (community based) before and after delivery	Research assistants, under supervision of research team

HOW LONG will it take to collect the data for each component of the study?

Step 1: Consider:

- The time required to reach the study area(s).
- The time required to locate the study units (persons, groups, records). If you have to search for specific informants (e.g., users or defaulters of a specific service), it might take more time to locate informants than to interview them.
- The number of visits required per study unit. For some studies it may be necessary to visit informants a number of times, for example, if the information needed is sensitive and can be collected only after informants are comfortable with the investigator or if observations have to be made more than once (follow-up of pregnant mothers or malnourished children). Allowing time for follow-up of nonrespondents should also be considered.

Step 2: Calculate the number of interviews that can be carried out per day (e.g., 4).

Step 3: Calculate the number of days needed to carry out the interviews. For example:

- you need to do 200 interviews,
- your research team of 5 people can do $5 \times 4 = 20$ interviews per day,
- you will need $200 \div 20 = 10$ days for the interviews.

Step 4: Calculate the time needed for the other parts of the study, (for example, 10 days)

Step 5: Determine how much time you can devote to the study. Because the research team usually consists of very busy people, it is unlikely that team members can spend more than 30 working days on the entire study.

- 5 days for preparation (including pretesting and finalizing questionnaires)
- 20 days actual fieldwork
- 5 days data processing + preliminary analysis.

If the team has 20 days for fieldwork, as in the example above, it could do the study without extra assistance. However, it the research team has only five days available for the interviews, they would need an additional five research assistants to help complete this part of the study.

Note:

Recruiting research assistants for data collection may, on one hand, relieve the research team, but, on the other hand, the training and supervision of research assistants require time (see Annex 12.1). The team has to carefully weigh advantages and disadvantages. If none of the team members has previous research experience, they might prefer designing a study that they can carry out themselves, without or with only minimal assistance.

If research assistants are required, consider to what extent local health workers can be used. They have the advantage of knowing the local situation. They should never be involved, however, in conducting interviews to evaluate the performance of their own health facility. Local staff from related services (teachers, community development) or students might help out. Sometimes village health workers or community members can collect certain parts of the data.

Note:

It is always advisable to slightly overestimate the period needed for data collection to allow for unforeseen delays.

In WHAT SEQUENCE should data be collected?

In general, it is advisable to start with analysis of data already available. This is essential if the sample of respondents is to be selected from the records. Another rule of thumb is that qualitative research techniques (such as focus group discussions) that are devised to focus the content of questionnaires should be carried out before finalization of the questionnaires. If the FGDs are to provide feedback on issues raised in larger surveys, however, they should be conducted after preliminary analysis of the questionnaires.

To use time and transport efficiently, data to be drawn from different sources in one locality should be collected at the same time. (For example, interviews with staff in a health centre, observations of equipment available in the centre, and interviews with mothers living nearby should be scheduled together.)

WHEN should the data be collected?

The actual time that the data will be collected will be determined by the type of data to be collected and the demands of the project. Consideration should be given to:

- availability of research team members and research assistants,
- the appropriate season(s) to conduct the fieldwork (if the problem is season-related or if data collection would be difficult during certain periods),
- accessibility and availability of the sampled population, and
- public holidays and vacation periods.

Note:

The field visit to obtain consent from local authorities for the research may also be used to obtain necessary details about the best period for data collection and availability of local resources (research assistants, transport), if required.

2. Ensuring quality

It is extremely important that the data we collect are of good quality, that is, reliable and valid. Otherwise we may come up with false or misleading conclusions.

In the previous modules **possible sources of data distortion** (bias) have been discussed. Biases we should try to prevent include:

- **Deviations from the sampling procedures** set out in the proposal.
- **Variability or bias in observations or measurements** made because:
  - Our **study subject changes his or her behaviour** as a consequence of the research. For example, a subject may act more positively while being observed; blood pressure and pulse may increase when the subject is apprehensive.
  - We use **unstandardized measuring instruments.** For example, we may use unstandardized weighing scales or imprecise or no guidelines for interviewing.
  - Researchers themselves vary in what they observe or measure (**observer variability**). For example, researchers may be selective in their observations (observer bias); measure, question, or note down answers with varying accuracy or follow different approaches (one being more open, friendly, probing than the other).

- **Variations in criteria for measurement or for categorizing answers** because we changed them during the study.

There are a number of measures that can be taken to prevent and partly correct such distortions, but remember: prevention is FAR better than cure! Cure is usually surgery: you may have to cut out the bad parts of your data or, at best, devise crutches.

There are several other aspects of the data-collection process that will help ensure data quality. You should:

- **Prepare a fieldwork manual for the research team as a whole,** including:
  - guidelines on **sampling procedures** and what to do if respondents are not available or refuse to cooperate (see Module 11, p. 7),
  - a clear **explanation** of the purpose and procedures of the study, which should be used to introduce each interview, and
  - **instruction sheets** on how to ask certain questions and how to record the answers.

- **Select your research assistants, if required, with care.** Choose assistants that are:
  - from the same educational level;
  - knowledgeable concerning the topic and local conditions;
  - not the object of study themselves; and
  - not biased concerning the topic (for example, health staff are usually not the best interviewers for a study on alternative health practices).

- **Train research assistants carefully in all topics covered in the fieldwork manual as well as in interview techniques** (see **Annex 12.1**) and make sure that all members of the research team master interview techniques such as:
  - asking questions in a neutral manner;
  - not showing by words or expression what answers one expects;
  - not showing agreement, disagreement, or surprise; and
  - recording answers precisely as they are provided, without sifting or interpreting them.

- **Pretest research instruments and research procedures** with the whole research team, including research assistants (see **Module 14**).

- **Take care that research assistants are not placed under too much stress** (requiring too many interviews a day; paying per interview instead of per day).

- **Arrange for on-going supervision** of research assistants. If, in case of a larger survey, special supervisors have to be appointed, supervisory guidelines should be developed for their use.

- **Devise methods to assure the quality** of data collected by all members of the research team. For example, quality can be assured by:
  - requiring interviewers to check whether the questionnaire is filled in completely before finishing each interview;

- asking the supervisor to check at the end of each day during the data collection period whether the questionnaires are filled in completely and whether the recorded information makes sense;
- having the researchers review the data during the data analysis stage to check whether data are complete and consistent.

Stage III: Data handling

Once the data have been collected, a clear procedure should be developed for handling and storing the data:

- First, it is necessary to check that the data gathered are complete and accurate (see section on quality control above).
- At some stage questionnaires will have to be numbered. Decide if this should be done at the time of the interview or at the time the questionnaires are stored.
- Identify the person responsible for storing data and the place where they will be stored.
- Decide how data should be stored. Record forms should be kept in the sequence in which they have been numbered.

Annex 12.1. Training interviewers

1. Interviewers' tasks

During the fieldwork, interviewers (or research assistants) may work independently or together with one of the researchers. If they go out independently, they may have to carry out the following tasks:

- Do the **sampling in the field** (for example sampling of households within a village and/or sampling of individuals to be interviewed within households).
- Give a clear **introduction** to the interviewee concerning the purpose and procedures of the interview.
- **Perform the interviews.** Obviously it is best to give interviewers standard questionnaires to administer. It is not wise to assign the more difficult tasks of performing highly flexible interviews or focus group discussions to interviewers.

It is imperative that interviewers be trained by the researchers so they can carry out their tasks accurately and correctly, according to the procedures developed by the researchers. Interviewers should not be left to develop their own procedures. If each interviewer is allowed to develop his own approach, bias is almost certain to result.

The training of interviewers may take 2 to 3 days. The first day may be devoted to theory, followed by 1 or 2 days of practical training, depending on the local circumstances and the nature of the study.

2. Theoretical training

Interviewers must be thoroughly familiar with the objectives of the research project and the methodology. Therefore, it is recommended that they be provided with a copy of the research protocol and that the most relevant sections be discussed thoroughly, including:

- statement of the problem,
- objectives,
- data-collection tools to be used (an overview),
- sampling procedures (if sampling has to be done in the field),
- plan for data collection, and
- plan for data analysis.

It is important at this stage that the interviewer trainees get ample opportunity to ask questions.

Then a more in-depth discussion should follow concerning the data-collection tools (questionnaires and possibly checklists) that are to be used by the interviewers. For each and every question they should know WHY the information is required.

Pre-testing

What is a pretest or pilot study of the methodology?

A PRETEST usually refers to a small-scale trial of a particular research component.

A PILOT STUDY is the process of carrying out a preliminary study, going through the entire research procedure with a small sample.

WHY do we carry out a pretest or pilot study?

A pretest or pilot study serves as a trial run that allows us to identify potential problems in the proposed tudy. Although this means extra effort at the beginning of a research project, the pretest or pilot study nables us, if necessary, to revise the methods and logistics of data collection before starting the actual ieldwork. As a result, a good deal of time, effort, and money can be saved in the long run. Pretesting is impler and less time consuming and costly than conducting an entire pilot study. Therefore, we will oncentrate on pretesting as an essential step in the development of the research projects.

'hat aspects of your research methodology can be evaluated during pretesting?

Reactions of the respondents to the research procedures can be observed in the pretest to determine:

- availability of the study population and how respondents' daily work schedules can best be respected;
- acceptability of the methods used to establish contact with the study population;
- acceptability of the questions asked; and
- willingness of the respondents to answer the questions and collaborate with the study.

The **data-collection tools** can be pretested to determine:

- Whether the tools you use allow you to collect the information you need and whether those tools are reliable. You may find that some of the data collected are not relevant to the problem or are not in a form suitable for analysis. This is the time to decide not to collect these data or to consider using alternative techniques that will produce data in a more usable form.

- How much time is needed to administer the questionnaire, to conduct observations or group interviews, and to make measurements.

- Whether there is any need to revise the format or presentation of questionnaires or interview schedules, including whether:

  - The sequence of questions is logical,
  - The wording of the questions is clear,
  - Translations are accurate,
  - Space for answers is sufficient,

- There is a need to precategorize some answers or to change closed questions into open-ended questions,
- There is a need to adjust the coding system, or
- There is a need for additional instructions for interviewers (e.g., guidelines for "probing" certain open questions).

3. **Sampling procedures** can be checked to determine:

   - Whether the instructions to obtain the sample are followed in the same way by all staff involved.
   - How much time is needed to locate individuals to be included in the study.

4. **Staffing and activities of the research team** can be checked, while all are participating in the pretest, to determine:

   - How successful the training of the research team has been.
   - What the work output of each member of the staff is.
   - How well the research team works together.
   - Whether logistical support is adequate.
   - The reliability of the results when instruments or tests are administered by different members of the research team.
   - Whether staff supervision is adequate.

   The pretest can be seen as a period of extra training for the research team in which sensitivity to the needs and wishes of the study population can be developed.

5. **Procedures for data processing and analysis** can be evaluated during the pretest. Items that can be assessed include:

   - Appropriateness of data master sheets and dummy tables and ease of use.
   - Effectiveness of the system for quality control of data collection.
   - Appropriateness of statistical procedures (if used).
   - Clarity and ease with which the collected data can be interpreted.

6. The proposed **work plan and budget for research activities** can be assessed during the pretest. Issues that can be evaluated include:

   - Appropriateness of the amount of time allowed for the different activities of planning, implementation, supervision, coordination, and administration.
   - Accuracy of the scheduling of the various activities.

When do we carry out a pretest?

You might consider:

- **Pretesting at least your data-collection tools, either during the workshop, or, if that is impossible, immediately thereafter, in the actual field situation.**

- Pretesting the data-collection and data-analysis process 1-2 weeks before starting the fieldwork with the whole research team (including research assistants) to allow time for revisions.

Which components should be assessed during the pretest?

1. Pretest during the workshop

Depending on how closely the pretest situation resembles the area in which the actual fieldwork will be carried out, it may be possible to pretest:

- The reactions of respondents to the research procedures and to questions related to sensitive issues.
- The appropriateness of study type(s) and research tools selected for the purpose of the study (e.g., validity: Do they collect the information you need?; and reliability: Do they collect the data in a precise way?).
- The appropriateness of format and wording of questionnaires and interview schedules and the accuracy of the translations.
- The time needed to carry out interviews, observations or measurements.
- The feasibility of the designed sampling procedures.
- The feasibility of the designed procedures for data processing and analysis.

Even if you cannot assess all these components fully, the field experience will provide information that will be quite valuable to you when reviewing the methodological aspects of your proposal and when developing your work plan and budget.

2. Pretest in the actual research area

All the issues mentioned above may have to be reviewed again during a pretest in the actual field situation. Other issues, such as the functioning of the research team, including newly recruited and trained research assistants, and the feasibility of the work plan, can only be tested in the research area. An important output of the pretest should be a fully developed work plan.

If choices have to be made as to what to include in the pretest, the following considerations may be helpful:

- **What difficulties do you expect in the implementation of your proposal?** Think of possible sources of bias in data-collection techniques and sampling and ethical issues you considered during the preparation of your plan for data collection (**Module 12**). Can some of these potential problems be overcome by adapting the research design?

- **If you feel you have little experience with a certain data-collection technique** you may want to do some extra practice during the pretest.

- **Which parts of your study will be most costly and time consuming?** Questionnaires used in large surveys, for example, should always be tested. If many changes are made the instruments should be pretested again. If a questionnaire or interview schedule has been translated into a local language, the translated version should be pretested as well.

Note

It is highly recommended that you analyze the data collected during the pretest right away. Then finalize and adjust the master sheets, if necessary. Make totals for each variable included in the master sheets. Fill in some dummy tables and prepare all the dummy tables you need, considering your research objectives.

Do all this even if you plan to analyze the data by computer. You will detect shortcomings in your questionnaires that you can still correct!

Who should be involved in the pretest or pilot study?

- The research team, headed by the principal investigator.
- Any additional research assistants or data collectors that have been recruited.

How long should the pretest or pilot study last?

The time required for a pretest or pilot study will be determined by a number of factors:

- The size and duration of the research project. (The longer the study will take, the more time you might reserve for the test run.)
- The complexity of the methodology used in the research project.

Keep in mind that this is the last chance you will have to make adjustments that will help to ensure the quality of your fieldwork. If you have a 20-day fieldwork period, you might reserve at least 3-5 days for pretesting your data-collection tools, analyzing the results of the pretest, finalizing your tools, and elaborating the work plan.

GROUP WORK I: To prepare the pretest during the workshop (1-1½ hours)

Only half a day will be available for conducting a pretest of your methodology during the course:

1. **Determine what parts of the methodology you would like to test. Include all data-collection tools, if possible.**
2. **Decide with your facilitator and course manager where in the local area you could best carry out the pretest.**
3. **Decide which members of your team will conduct various aspects of the pretest. You are advised to work in pairs, so that you can discuss observations during the pretest.**
4. **Prepare a short list of questions you wish to answer during the pretest. (See Annex 14.1 for suggestions.)**

Annex 14.1. Summary of points to assess during a pretest or pilot study

1. **Reactions of respondents to your research procedures**	Acceptable	Not acceptable	Suggestions
Availability of sample needed for full study			
Work schedules of population that may affect their availability			
Desire of population to participate			
Acceptability of questions			
Clarity of the language used			

2. **The data-collection tools**	Acceptable	Not acceptable	Suggestions
Whether the tools provide the information you need and are reliable			
Time needed for administering each of the data-collection tools			
Presentation of questions and format of questionnaire			
Accuracy of translation			
Precategorizing of questions			
Coding system and coding guidelines			
Handling and administering the tools			

3. **Sampling procedures**	Acceptable	Not acceptable	Suggestions
Whether the instruction to obtain the sample are used uniformly by all staff			
Time needed to locate the individuals to be included in the study			

4. **Preparation and effectiveness of research team**	Acceptable	Not acceptable	Suggestions
Adequacy of staff training			
Output of each team member			
Team dynamics			
Reliability of tools when administered by different team members			
Accuracy of interpretation			
Appropriateness of plan for supervision			

5. **Procedures for data processing and analysis**	Acceptable	Not acceptable	Suggestions
Use of data master sheets			
Effectiveness of data quality control			
Appropriateness of statistical procedures			
Ease of data interpretation			

6. **Schedule for research activities**	Acceptable	Not acceptable	Suggestions
Amount of time allowed for: • field trips for data collection			
• supervision			
• administration			
• analysis of data			
Sequence of activities			

Annex 14.2. Summary of possible fallacies in the design and implementation of studies

As we have now gone through all steps of the study design, including the planning of data processing and analysis, it may be useful to summarize the critical points at which a researcher can go wrong:

- In the SELECTION of RESPONDENTS or study elements, and
- In the COLLECTION of data.

These potential errors should be reviewed while you are pretesting your research methodology.

Errors in selection of respondents or study elements

In the selection of respondents we may distinguish several major possibilities for error.

Too limited (or inappropriate) definition of the study population or use of incorrect sampling procedures, for example by:

- Studying registered patients only;
- Obtaining responses from male opinion leaders only (if one needs the opinion of the whole community);
- Choosing a sample because it is close to a road or in some other way easier to access (tarmac bias); or
- Conducting the study during only one season of the year (when results may be biased by not including other seasons or because access is difficult).

Errors in the assignment of research subjects to study groups in analytic and experimental studies:

- Defective matching in case-control studies;
- The inclusion of volunteers for study groups in cohort studies;
- Nonrandomization in experimental studies; or
- If randomization is impossible, failure to develop a quasiexperimental design that corrects as much as possible for "rival explanations."

Selective dropouts or nonresponse

Dropouts or subjects who do not respond to selected questions may represent a special category of respondents. If attrition is high or the rate of nonresponse excessive, results may be biased.

In cohort studies, follow-up of individuals can pose problems. Bias in follow-up results if there is a differential dropout between those exposed to the risk and those without exposure.

Errors in data collection

We may obtain:

Invalid data, by applying indicators and measuring techniques or instruments that do not adequately measure what we want to measure.

Unreliable data due to:

- Variation in the characteristics of the research subject measured, as a consequence of the research;
- The use of unstandardized measuring instruments; or
- Differences between observers and interviewers.

Reliability of data collected is always required, but it is of crucial importance if we want to measure changes over time. If we find changes we must be sure that these are not caused by errors in our research methods that could have been prevented.

All the above-mentioned shortcomings may threaten the **validity** of your findings and conclusions. The shortcomings can be prevented to some degree by being alert to them when designing and implementing the study; otherwise they have to be mentioned in the study design.

Budget

Why do we need a budget?

- A detailed budget will help you to identify which resources are already locally available and which additional resources may be required.
- The process of budget design will encourage you to consider aspects of the work plan you have not thought about before and will serve as a useful reminder of activities planned, as your research gets underway.

When should budget preparation begin?

A complete budget is normally not prepared until the final stage of project planning. However, cost is usually a major limiting factor and, therefore, must always be kept in mind during planning so that your proposals will not have an unrealistically high budget. (See Module 4, Analysis and statement of the problem.) Remember that both ministries and donor agencies usually set limits for research project budgets.

The use of locally available resources increases the feasibility of the project from a financial point of view.

How should a budget be prepared?

It is convenient to use the work plan as a starting point. Specify, for each activity in the work plan, what resources are required. Determine for each resource needed the **unit cost** and the **total cost**.

Example:

In the work plan of a study to determine the utilization of family planning methods in a certain district, it is specified that 5 interviewers will each visit 20 households in clusters of 4 over a time period of 5 working days. A supervisor will accompany one of the interviewers each day using a car. The other 4 interviewers will use motor cycles. The clusters of households are scattered over the district but are on average 50 kilometres from the district hospital from where the study is conducted.

The budget for the field work component of the work plan will include funds for personnel, transport and supplies.

Note that UNIT COST (e.g., per diem or cost of petrol per km), the MULTIPLYING FACTOR (number of days), and TOTAL COST should be clearly indicated for all budget categories.

Table 17.1. Costs involved in fieldwork for a family-planning study.

Budget category	Unit cost	Multiplying factor	Total cost
1. Personnel	**Daily wage (including per diem)**	**Number of staff-days (no. staff × no. of working days)**	**Total**
Interviewers	$10	5 × 5 = 25	$250
Supervisor	$20	1 × 5 = 5	$100
		Personnel TOTAL	**$350**
2. Transport	**Cost per km**	**Number of km (no. vehicles × no. days × no. km/day)**	**Total**
Motorcycles	$0.10	4 × 5 × 100 = 2000	$200
Car	$0.40	1 × 5 × 100 = 500	$200
		Transport TOTAL	**$400**
3. Supplies	**Cost per item**	**Number**	**Total**
Pens	$1.00	12	$12
Questionnaires	$0.20	120	$24
		Supplies TOTAL	**$36**
		GRAND TOTAL	**$786**

If more than one budget source will be used (e.g., the ministry of health and a donor), it would be useful to indicate in the budget which source will pay for each cost. Usually a separate column is used for each funding source. (See Annex 17.1.)

Advice on budget format

An example of a project budget is provided in Annex 17.1. This budget includes the major categories that are usually needed for small projects: personnel, transport, and supplies and equipment.

The type of budget format to be used may vary depending upon whether the budget will be supported by your own organization or the ministry of health or submitted to a donor organization for funding. Most donor organizations have their own special project forms, which include a budget format.

If you intend to seek donor support it is advisable to write to the potential funding organization as early as possible during the period of project development.

Advice on budget preparation

- Keep in mind the tendency to underestimate the time needed to complete project tasks in "the real world." Include a 5% contingency fund if you fear that you might have budgeted for the activities rather conservatively. (If inclusion of a contingency fund is not allowed, an alternative is to slightly over-budget in major categories.)

- Do not box yourself in too tightly with very detailed categories and amounts, especially if regulations do not allow adjustments afterward. Ask the supervising agency to agree that there may be some transfer between "line items" in the budget, if needed.

- If your government or department has agreed to contribute a certain amount for the project, try to arrange that the contribution be administered separately, so that the administrators remain aware of the commitment. This may also ensure easier access to the funds.

- If the budget is for a period longer than a year, build in allowances for inflation before the project begins and in subsequent years by increasing costs by a set percentage. (If inflation is high in the local economy, you may have to build in allowances for even shorter projects.)

Budget justification

It is not sufficient to present a budget without explanation.

The budget justification follows the budget as an explanatory note justifying briefly, in the context of the proposal, why the various items in the budget are required. Make sure you give clear explanations concerning why items that may seem questionable or are particularly costly are needed and discuss how complicated expenses have been calculated. If a strong budget justification has been prepared, it is less likely that essential items will be cut during proposal review.

How can budgets be reduced?

- Explore whether other health-related institutions are willing to temporarily allocate personnel to the project.

- When possible, use local rather than outside personnel. If consultants are needed at the beginning, train local personnel as soon as possible to take over their work.

- Explore the use of students or community volunteers, where appropriate.

- Plan for strict control of project expenditures, such as those for vehicle use, supplies, etc.

Obtaining funding for projects

To conduct research, it is usually necessary to obtain additional funding for the research project. Suc funding may be available from local, national, or international agencies. In addition to preparing a goo research proposal, the following strategies are useful for researchers who need to obtain their own fundinç

1. Familiarize yourself with the policies and priorities of funding agencies. Such policies and priorities ma be:

   - explicit, i.e., available from policy documents issued by the agency;
   - implicit, i.e., known to officials in the agency and to other local researchers who have previousl been funded by that agency.

   Obtain the names of such persons and make direct contact with them.

   The funding policies of many agencies may emphasize:

   - priority for research aimed at strengthening a particular program (e.g., MCH, PHC);
   - institution building (i.e., building the capacity of an institution to do research);
   - research credibility.

   Annex 17.2 gives a list of some prominent research funding agencies.

2. Identify the procedures, deadlines, and formats that are relevant to each agency.

3. Obtain written approval and support from relevant local and national health authorities and submit thi together with your proposal.

4. If you are a beginning researcher, associate yourself with an established researcher. Host agencie scrutinize the "credibility" of the researcher to whom funds are allocated. Such credibility is based or previous projects that have been successfully completed.

5. Build up your own list of successfully completed projects (i.e., your own reports, publications, etc.).

ANNEX 17.1. Example of budget for a child-spacing study (in kwachas)

1. **Personnel costs** (excluding workshops)	**Ministry of health**	**Donor**	**Total**
Research team			
88 person-days in provincial capital	Salary		
56 person-days in field			
per diem 56 × K 45	"	2,520	2,520
Research assistants			
20 person-days in provincial capital			
per diem 20 × K 45	"	900	900
50 person-days in field			
per diem 50 × K 45	"	1,750	1,750
Facilitator			
6 person-days in provincial capital			
per diem 6 × K 120		720	720
per diem driver 6 × K 35		210	210
Drivers of project			
18 person-days			
per diem 10 × K 35	"	630	630
Secretary			
8 person-days	"		
2 seniors of each of the 5 district hospitals			
11 person-days in provincial capital	"		
per diem 11 × K 70		770	770
2 senior officials MOH	"		
4 person-days in provincial capital			
per diem 4 x K 70		280	280
driver per diem 2 x k 35		70	70
SUBTOTAL	4630	7,850	12,480

2. Transport costs	MOH	Donor	Total
Clearance local leaders (340 km) Compilation CS records staff interviews (21 clinics) (2100 km)			
Training research assistants and field test (100 km)			
Data collection in 2 districts (1400 km)			
Discussion District Health Teams and HQ authorities (1540 km)			
Facilitators' visits (2880 km)			
TOTAL MILEAGE (8360 km)			
8360 × K 0.35/km for petrol		2,926	2,926
8360 × K 1/km for operating costs	8,360		8,360
Public transport for research assistants		210	210
2 return air tickets for senior MOH staff		450	450
SUBTOTAL	8360	3,586	11,946

3. Supplies	MOH	Donor	Total
12 reams duplicating paper × K 37.50	450		
1 ream writing paper	50		
1 ream photocopy paper	70		
20 folders × K 5	100		
5 writing pads × K 8	40		
Pens, rubbers, etc.	60		
4 boxes stencils × 4.50	200		
5 tubes duplicating ink	110		
SUBTOTAL	1,080		1,080

SUMMARY

Personnel costs	4,630	7,850	12,480
Transport costs	8,360	3,586	11,946
Stationery	—	1,080	1,080
TOTAL (kwachas)	12,990	12,516	25,506
5% contingency	650	626	1,275
GRAND TOTAL (kwachas)	13,640	13,142	26,781
(US$)	5,683	5,476	11,159

(Exchange rate 1 US$ = K 2.40)

nnex 17.2. International sources of funding for research

International multilateral agencies

WHO and associated special programs:

- WHO Regional Offices
- WHO Headquarters
- TDR (Tropical Disease Research)
- CDD (Control of Diarrheal Disease)
- HRP (Human Reproduction Programme)

UNICEF (United Nations Children's Fund)
World Bank
IARC (International Agency for Research on Cancer)

Bilateral agencies

USAID (United States Agency for International Development)
IDRC (International Development Research Centre)
SAREC (Swedish Agency for Research Cooperation with Developing Countries)
GTZ (Deutsche Gensellschaft Fur Technische Zusammenardeit)
JICA (Japanese International Cooperation Agency)
BOSTID (Board on Science and Technology for International Development)
CIDA (Canadian International Development Agency)
SIDA (Swedish International Development Agency)
ODA (Overseas Development Agency)
ADAB (The Australian Development Assistance Board)

Private foundations

Rockefeller Foundation
Carnegie Corporation
Ford Foundation (Child Health)
Kellogg Foundation (Health Services; primary interest in Latin America)

National sources

This will vary from country to country.

Addresses of some funding agencies

1. Rockefeller Foundation
   1133 Avenue of Americas
   New York, NY 10036
   U.S.A.

2. Carnegie Corporation of New York
   437 Madison Avenue
   New York, NY 10022
   U.S.A

3. Director, International Health Policy Program,
   S-6133, 1818 "H" Street, NW
   Washington, DC 20433
   U.S.A.

4. Health Sciences Division,
   International Development Research Centre
   P.O. Box 8500
   Ottawa, Canada K1G 3H9

5. The Asia-Pacific Academic Consortium for Public Health
   420/1 Rajvidhi Road, Pyathai,
   Bangkok 10400, Thailand.

6. Primary Health Care Operations Research,
   Center for Human Services,
   5530 Wisconsin Avenue
   Chevy Chase, MD 20815
   U.S.A.

Report writing

I. STEPS IN PREPARING A REPORT: PRELIMINARY CONSIDERATIONS

The Audience

The purpose of a research report is to convey information to the reader. Therefore, it is important to begin by clarifying in your mind:

WHO is the reader?
WHY does he or she want to read the research report?

In health systems research, it is particularly important to remember the needs of the audience because the audience is not only the research community, but also health managers and community leaders. Many research papers that are meant for scientific people are not suitable for managers and lay people. Therefore, special attention should be devoted to preparing reports that are simply worded and are explicit regarding findings.

Furthermore, it is important to present not only the scientific findings, but also specific recommendations that take into consideration the local characteristics of the health system, constraints, feasibility, and usefulness of the proposed solutions. The community and the manager are more interested in learning "what to do about a problem" than in being told "there is a problem."

> Reports should meet the **NEEDS OF THE AUDIENCE** of community leaders, health managers, and researchers.

How the Reader Reads a Research Report

Recognizing the "reading strategies" of people who read research reports will help you write a good report. The research was done to provide new information. Therefore, this should be the highlight and focus of the report. This "new information" should be summarized as the **conclusions** of the study. Most readers will begin by reading the conclusions. If this section is interesting, useful, and attractively presented, the reader will look at the other sections. The other sections of the report are intended to support the conclusions by helping the reader clarify two basic questions in his or her mind:

- **How will this "new information" help improve the health of the community?** (i.e., What is the problem and the health system in which the problem occurs and how will this information help solve or reduce the problem?)

- **Can I "believe" these findings?** (i.e., Are the findings valid and reliable?) The research design, sampling, methods of data collection, and the data analysis will substantiate the validity and reliability.

Note that a report that highlights the methodology sections rather than the conclusions might interest a researcher audience, but will not interest the manager audience.

Completing the Data Analysis

Before you begin the outline and first draft of your report, you need to review your analysis of the data asking several of the following questions:

- **Are conclusions appropriate to the specific objectives? Are they comprehensive?**

  The earlier steps in data analysis should have produced:

  - one or more conclusions stated as simple sentences; and
  - one or more analytic tables together with the relevant descriptive statistics or statistical tests to support the conclusions.

  Review these conclusions and check whether:

  - every specific objective has been dealt with;
  - all aspects of each objective have been dealt with; and
  - the conclusions are relevant and appropriate to the objectives.

- **Are further analytic tables needed?**

  If the conclusions are not comprehensive, prepare further dummy analytic tables and analyze the data as described in **Modules 22-30.**

- **Have all qualitative data been used to support and specify conclusions drawn from tables?**

Once you have completed this review, you need to complete a couple of additional tasks:

- **State the final conclusions in relation to each objective.**

  During earlier stages of analysis, every analytic table would have had a conclusion. These conclusions should now be reviewed, combined whenever possible, and stated in such a way that the main findings of the study are easily identifiable by a reader who is "scanning" the report. Very often the most important numerical information (%, means etc.) can be included in these statements.

- **Select supportive tables to appear in the text of the report.**

  The number of tables in the body of the report should be very limited. A table should be included only if it illustrates an important conclusion or provides evidence to support it. When possible, combine information from several analytic tables into one or more and present a summary table in the body of the report. (If necessary, more detailed tables can be placed in annexes.) The title of each table should tell the reader in as few words as possible exactly what the table contains. Column and row headings should be brief, but self-explanatory.

Compile the conclusions and tables relating to each specific objective. You are now ready to draft the report.

II. WRITING THE REPORT

The aim of the report is to tell the reader the facts in a simple, logical, sequential fashion. Avoid confusion and distracting the reader.

In writing the report, it is important to consider:

- the CONTENT,
- the STYLE of writing,
- the LAYOUT of the report,
- FIRST DRAFT,
- SECOND DRAFT,
- finalizing the report.

Each of these aspects of report preparation will be discussed in turn.

Content: Main Components of a Research Report

The research report should contain the following components:

Title or cover page,
Summary of findings and recommendations,
Acknowledgments (optional),
Table of contents,
List of tables, figures (optional),
List of abbreviations (optional),
1. INTRODUCTION,
2. OBJECTIVES,
3. METHODOLOGY,
4. FINDINGS AND CONCLUSIONS,
5. DISCUSSION,
6. RECOMMENDATIONS,

References,
Annexes (data collection tools, tables).

The findings and conclusions, discussion of findings, and recommendations will form the most substantial part of your report, which has to be written from scratch. For the introduction you can rely to a large extent on your research proposal, although you may summarize, revise, and sometimes expand certain sections.

We, therefore, strongly advise that you **start with the findings and conclusions.** Nevertheless we will briefly elaborate on each component in the sequence in which they will finally appear in your report.

- **Cover Page**

  The cover page should contain the title, the names of the authors with their titles and positions, the institution that publishes the report, and the month and year of publication. The institution that publishes the report will most likely be the one that administered the project, for example, the Research Unit of the Ministry of Health or a research institute.

- **Summary**

  The summary can only be written after the first or even the second draft of the report has been completed. It should contain:

  a very brief description of the problem (WHAT),

- the main objectives (WHY),
- the place of study (WHERE),
- the type of study and methods used (HOW),
- the main findings and conclusions, followed by
- the major, or all, recommendations.

The summary will be the first (and for busy health decision-makers most likely the only) part of your study that will be read. Therefore, its writing demands thorough reflection and is time consuming. Several drafts may have to be made, each discussed by the research team as a whole.

As you will have collaborated with various groups during the drafting and implementation of your research proposal, you may consider writing different summaries for each of these groups. For example, you may prepare different summaries for policymakers and health managers, for health staff of lower levels, for community members or the public at large (newspaper, TV), and for professionals (articles in scientific journals). (See **Module 32**.)

- **Acknowledgments**

You may wish to thank those who supported you technically or financially in the design and implementation of your study. Also your employer, who has allowed you to invest time in the study, and the respondents may be acknowledged. Acknowledgments are usually placed right after the cover page or at the end of the report, before the references.

- **Table of Contents**

A table of contents is essential, as it gives the reader a quick overview of the major sections of your report, and page references, if he wishes to go through the report in a different order or skip certain sections.

- **List of Tables, Figures (optional)**

If you have many tables or figures it is helpful to list these also, in a "table of contents" type format with page numbers.

- **List of Abbreviations (optional)**

If there are many abbreviations or acronyms in the report, these could be listed in addition.

The latter three sections should be prepared last, as you have to include the page numbers of all chapters and subsections in the table of contents, and be sure there are no mistakes in the final numbering of figures and tables.

1. INTRODUCTION

The introduction is a relatively easy part of the report which may be written after a first draft of the findings has been made. It should certainly contain some background data about the country, the health status of the population, and health service data related to the problem that has been studied. You may slightly revise or make additions to the corresponding section in your research proposal and use it here.

Then the statement of the problem should follow, again revised from your research proposal with comments or additional data based on your research experience added, if useful. It should contain a paragraph on what you hope to achieve with the results of the study.

A brief review of the literature pertaining to your topic of study should then be given. (Consult **Module 5** and your research proposal.) This section should include relevant points to help the reader:

- understand the problem providing a review of available information on it, and
- understand methods of investigating or resolving the problem.

NOTE: This section should NOT be a summary of all the papers and books on the topic. Be selective, remembering that this section serves to lend support for your study, not to display your ability to read literature.

2. OBJECTIVES

The general and specific objectives should be included. If necessary, you can adjust them slightly for style and sequence. However, you should not change their basic nature. If you have not been able to meet some of the objectives, this should be stated in the methodology section and in the discussion of the findings.

3. METHODOLOGY

The methodology you followed for the collection of your data should be described in detail. It should include :

- the study type,
- the variables on which data was collected,
- the population from which the sample was selected,
- the size of the sample and method of sampling,
- the data collection techniques:
  sources of data (cards, households, clinic registers, etc.),
  how the data was collected and by whom,
  procedures for data analysis, including statistical tests (if applicable).

If you have deviated from the original study design presented in your research proposal, you have to explain to what extent and why. The consequences of this deviation for meeting certain objectives of your study should be indicated. If the quality of some of the data is weak, resulting in possible biases in a certain direction, this should be described.

4. FINDINGS AND CONCLUSIONS

The systematic presentation of your findings and conclusions in relation to the research objectives is the crucial part of your report.

A description of the findings may be complemented by a limited number of tables or graphs that summarize the findings. The text will become more lively if you illustrate some of the findings with examples using the respondents' own words, or with observations and case-studies that you recorded during the fieldwork.

5. DISCUSSION

The findings can be discussed by objective or by cluster of related variables. The discussion should also mention findings from other related studies that support or contradict your own. It is important, as well, to present and discuss the limitations of the study. In the discussion of findings some general conclusions may be included as well.

Note:

The text and annexes should include sufficient details for professionals to enable them to follow how you substantiate your findings and conclusions. The report should be so self-explanatory that it should be possible to repeat the study, if desired.

6. RECOMMENDATIONS

The recommendations should follow logically from the discussion of the findings. They may be summarized according to the groups toward which they are directed, for example:

- policymakers,
- health and health-related managers at district or lower level,
- health and health-related staff who could implement the activities,
- potential clients, and
- the community at large.

Remember that action-oriented groups are most interested in this section.

In making recommendations, use not only the findings of your study, but also supportive information from other sources and available information on other related factors. The recommendations should be discussed with all concerned before they are finalized.

If your recommendations are short, you might include them all in your summary of findings and recommendations and omit them as a separate section.

- **References**

The references in your text can be numbered in the sequence in which they appear, then listed in this order in the reference section. Another possibility is to list the author's names in the text followed by the date of the publication in brackets, for example (Shan 1990). In the list of references, the publications are then arranged in alphabetical order by the principal author's last name (see Module 5).

You can choose either method, but if you wish to publish an article you must follow the method used in the journal to which you wish to submit your article.

- **Annexes or Appendices**

The annexes should contain any additional information needed to enable professionals to follow your research procedures and data analysis.

Information that would be useful to special categories of readers but is not of interest to the average reader could be included in annexes, as well.

Examples of information in annexes are:

- tables referred to in the text but omitted to keep the report short;
- lists of criteria, definitions, and flow-charts;
- lists of hospitals, districts, villages, etc., that have participated; and
- all data collection tools.

Style of Writing

Remember that your reader:

- Is short of time,
- Has many other urgent matters demanding his or her interest and attention, and
- Is probably not knowledgable concerning "research jargon."

Therefore the rules are:

- **Simplify**. Keep to the essentials.
- **Justify**. Make no statement that is not based on facts.
- **Quantify**. Avoid "large," "small"; instead, say "almost 75%," "one in three," etc.
- **Be precise and specific**.
- **Inform, not impress.** Avoid exaggeration.
- **Use short sentences**.
- **Use adverbs and adjectives sparingly**. Be consistent in the use of tenses (past, present tense). Avoid the passive voice, if possible.
- **Aim to be clear, logical, and systematic in your presentation.**

Layout of the Report

A good physical layout is important as it will help your report:

- Make a good initial impression,
- Encourage the reader, and
- Give an idea of how the material has been organized so the reader can make a quick determination of what he or she will read first.

Particular attention should be paid to make sure there is:

- An attractive layout for the title page; a clear table of contents,

- Consistency in margins and spacing,

- Consistency in headings and subheadings, (e.g., **capitals, underlined**, for headings of chapters; **capitals** for headings of major sections; **lower case, underlined**, for headings of subsections, etc.),

- Good quality typing and photocopying. Correct drafts carefully. (For more detailed information, see Keithly and Schreiner 1971).

- Numbering of figures and tables, provision of clear titles for them, and labels for columns and rows, etc.,

- Accuracy and consistency in quotations and references.

Preparing the First Draft of the Report

Prepare a written OUTLINE of the report. An outline will help to organize your thoughts and is an essential step in producing a logical, sequential report.

An outline should contain:

- Headings of the main sections of the report,
- Headings of subsections,
- Points to be made in each section, and
- A list of tables and figures (if relevant) to illustrate each section.

The outline for the chapter on findings and conclusions is the most difficult. A discussion with your team members concerning the main findings and conclusions of your data in relation to objectives and variables should help you structure your findings in a logical and coherent way.

- The first section under findings and conclusions is usually a description of the sample, for example in terms of location, age, sex, and other relevant background variables.

- Then, depending on the study design, you may provide more information on the problem or dependent variable(s) of your study.

- Next an analysis of the different independent variables in relation to the problem may follow.

You might start by listing headings and subheadings, with ample space between them so that you can scribble key words related to what you intend to write under each heading. It is advisable to number sections and subsections as you list them.

For example, in a study on malnutrition, the chapter "Findings and Conclusions" may look like this:

CHAPTER 4: FINDINGS AND CONCLUSIONS

4.1 DESCRIPTION OF THE SAMPLE

4.2 EXTENT AND SEASONAL VARIATION OF MALNUTRITION IN DISTRICT X

4.3 POSSIBLE CAUSES OF MALNUTRITION

4.3.1 Limited availability of food

4.3.2 Non-optimal utilization of food

4.3.3 High prevalence of communicable diseases

4.3.4 Limited access to MCH and curative services

4.3.5 Conclusions

This system of numbering is flexible and can be extended according to need with further headings or subheadings. It allows you to keep an overview of the process when different group members work on different parts at the same time. If your findings are very elaborate so that you get sub-sub-subheadings with 4 or 5 numbers, you might decide to split up the findings into several chapters. (In addition, you may consider leaving off some of the numbering on subsections, if it's clear under what major heading they belong. However, keep all the numbering until the final draft, as it helps you keep your report in order when various members of the group are working on different sections.)

TABLES and FIGURES in the text need numbers and clear titles. It is advisable to first use the number of the section to which the table belongs. In the last draft you may decide to number tables and figures in sequence.

Include only those tables and figures that present main findings and need more elaborate discussion in the text. Others may be put in annexes, or, if they don't reveal interesting points, be omitted.

Note that it is unnecessary to describe in detail a table that you include in the report. Only present the main conclusions.

The first draft is never final. Therefore you might concentrate primarily on content rather than on style. Nevertheless, it is advisable to structure the text straight from the beginning in paragraphs, and to attempt to phrase each sentence clearly and precisely.

Notes:

Never start writing without an outline. Make sure that all sections written carry headings and numbers consistent with the outline before they go for typing. Have the outline visible on the wall so everyone will be immediately aware of any additions or changes.

Type the first draft double-spaced with large margins so that you can easily make comments and corrections in the text.

Have several copies made of the first draft, so you will have one or more copies to work on and one copy on which to insert the final revisions you will hand in for retyping.

Preparing the Second Draft

When a first draft of the findings and conclusions has been completed, all working-group members and facilitators should read it critically and make comments.

The following questions should be kept in mind when reading the draft:

- Have all important findings been included?
- Do the conclusions follow logically from the findings? If some of the findings contradict each other, has this been discussed and possibly explained? Have weaknesses in the methodology, if any, been revealed?
- Are there any overlaps in the draft that have to be removed?
- Is it possible to condense the content? In general a text gains by shortening. Some parts less relevant for action may be included in annexes. Check if descriptive paragraphs may be shortened and introduced by a concluding sentence.
- Do data in the text agree with data in the tables? Are all tables consistent (with the same number of informants per category), are they numbered in sequence, and do they have clear titles?
- Is the sequence of paragraphs and subsections logical and coherent? Is there a smooth connection between successive paragraphs and sections? Is the phrasing of findings and conclusions precise and clear?

The original authors of each section may prepare a second draft, taking into consideration all comments that have been made. However, you might consider the appointment of two editors amongst yourselves, to draft the complete version.

In the meantime, other group members may (re)write the introductory sections. The INTRODUCTION, OBJECTIVES, and METHODOLOGY sections of your original research proposal may be used, in many cases, after being reviewed and adjusted (see page 8).

Now a first draft of the summary can be written (see page 7).

Finalizing the Report

It is advisable to have one of the other groups and facilitators read the second draft and judge it on the points mentioned in the previous section. Then a final version of the report should be prepared. This time you should give extra care to the presentation: structure, style, and consistency of spelling.

Use verb tenses consistently. Descriptions of the field situation may be stated in the past tense (e.g., "Five households owned less than one acre of land.") Conclusions on data are usually in the present tense (e.g., "Food taboos hardly have any impact on the nutritional status of young children. Those species of fish and meat that are forbidden for certain clans rarely appear in the daily diet.")

For a final check on readability, you might skim through the pages and read the first sentences of each paragraph. If this gives you a clear impression of the organization and results of your study, you may conclude that you did the best you could.

SECOND EDITION

PROPOSALS THAT WORK

A Guide for Planning Dissertations and Grant Proposals

LAWRENCE F. LOCKE
WANEEN WYRICK SPIRDUSO
STEPHEN J. SILVERMAN

1987

SAGE PUBLICATIONS
The International Professional Publishers
Newbury Park London New Delhi

Posttesting. One day after the last training session, and three weeks after pretesting, the subjects will be administered the TORCPR, the TORCSS, the Posttest Questioning task (POSQUES), and the Posttest Free Recall task (POSCALL). After posttesting, all subjects in the experimental group will be asked if they had used the strategy, about advantages and disadvantages of the technique, and about whether or not they would use the technique again. The nature of the research will again be explained to the subjects and any questions they have about the study will be answered to the best of the experimenter's ability at that time. The following week, treatment conditions will be reversed, with the control group receiving the experimental training, and vice versa.

In the final section of the body of the proposal the focus is on data analysis. The author reports all the techniques to be used for each part of the analysis. More detail would help many readers. It would be particularly helpful if the author had discussed the analysis in light of each hypothesis stated earlier. Including sample tables for the analyses to be performed is a good idea for all studies, particularly when there are as many variables as in this study. An illustration of the path analytic model to be tested should be included here so the model to be tested is apparent.

Data analysis. Reliability estimates for PREQUES, POSQUES, PRECALL, and POSCALL were calculated using a hand calculator with programmable statistics functions. These estimates were doublechecked by the experimenter. All other analyses will be performed using the Statistical Package for the Social Sciences (SPSS) statistics package (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975) edition 8.3, on the CDC 6000/Cyber 700 computers at the university. Data checking activities will include running subprogram FREQUENCIES to check on variance distributions and plotting distributions on probability paper. Subprograms FREQUENCIES and CROSSTABS will be used to generate the characteristics of treatment groups. Subprograms REGRESSION and PLOT will be used to run preliminary path analyses and to check that regression assumptions have been met, and subprogram REGRESSION will be used to run the final, restricted model path analysis. Subprogram PLOT will be used to generate figures which illustrate significant aptitude-treatment interactions.

PROPOSAL 4: FUNDED GRANT

A Field Test of a Health-Based Educational Intervention to Increase Adolescent Fertility Control

OVERVIEW AND OBJECTIVES

Adolescent premarital pregnancy is a major social and mental health problem in the U.S. and in Texas. More than one-half of all teenagers 15-19 years of age and almost one-quarter of those under 15 are sexually active. Significantly, the age at first intercourse continues to decline; younger adolescents (under 15) are less likely to use effective contraception and wait substantially longer to begin contraceptive usage initially than do older adolescents. As a result of the preceding factors and the larger proportion of teenagers (19 and under) presently in the general population, Texas has the *second* highest number in the nation of pregnancies among women under 15 and the *fifth* highest pregnancy rate (pregnancies per 1000 women in this age group) among 15-19 year olds. In most cases, these pregnancies create serious negative consequences for the adolescents, their families and babies (if they deliver).

The author begins by immediately establishing that the topic of the proposed research is not only of importance, but that it is of particular importance to the state in which the funding foundation resides. That

The original of this proposal was prepared by Marvin Eisen, Ph.D., formerly of the University of Texas at Austin, who is currently a Research Psychologist with Sociometrics Corporation of Palo Alto, California. The project was cooperatively funded for the first year by a number of sources, including two regional foundations, an agency of the state government, and two campus research institutes. The second and third years have been approved for funding by an agency of the federal government.

the problem is one of considerable regional significance is highlighted by the effective use of national rankings. The opening paragraph is easy to read, devoid of social science jargon, and sustains the reader's interest.

The foundation to which this proposal was submitted has as its major goal the funding of research devoted to mental health. Thus the author spends the next paragraph showing how the problem to be addressed by the research has direct implications for the mental health of the principals of the study and all others directly or indirectly involved with the problem. Words in the next paragraph such as health risks, clinical depression, suicide, stress, and mental health are guaranteed to catch the eye of members of the board of directors.

The negative consequences involve physical and mental health problems as well as economic and financial burdens for their families and communities. Prenatal, neonatal and maternal mortality rates are higher than those of older mothers. Young teenage mothers are more likely to suffer from medical complications of pregnancy and childbirth. Low maternal age is associated with a higher incidence of anemia, toxemia, surgical delivery, prematurity, low birth weight, birth defects and neurosurgical deficits. Poor nutrition, inadequate or late prenatal care and physical immaturity contribute to the health risks for young teenaged mothers and their children. In addition, teenage mothers are more than twice as likely to be suffering from clinical depression and seven to ten times more likely to attempt suicide than older mothers. Adolescent mothers may be more stressed by caring for their infants, especially if those infants were born prematurely or with developmental delays, as is more likely to occur with teenagers than older mothers. This increased stress and the young mother's general lack of maturity often leads to child abuse and neglect. Overall, both adolescent mothers and their offspring are at greater risk for physical and mental health related problems than older mothers and their children.

The demonstration project proposed is intended to field test preventive services on a community basis for adolescents (13-16 years) who are likely to be at risk for premarital sexual activity and pregnancy. In the initial twelve months of the demonstration we developed and pilot tested an educational intervention program designed to strengthen teenagers' beliefs about the value of individual sexual responsibility and to enhance their motivation

for self-discipline and fertility control. The content of the intervention was drawn from the *Health Belief Model* (HBM), a conceptual framework used successfully to predict and understand individual decisions to seek and use preventive health services. The intervention successfully modified teenagers' beliefs about their own *susceptibility* to pregnancy, the *seriousness* of a personal premarital pregnancy and reduced perceived *barriers* to personal abstinence or fertility control.

In the foregoing overview, note particularly the third paragraph, in which the author does two critical things. He introduces the crux of the intervention (the HBM), and he does so by showing that it was used successfully in previously completed research. Readers now have reason to believe that the intervention technique proposed will work, and that the author is well qualified because he already has completed work in this area. This introduction is immediately followed in the next paragraph by a concise description of the proposed project.

In the controlled field test phase presently proposed, experimental educational services will be organized and coordinated by a number of family planning services providers located throughout the State. These agencies will be selected in collaboration with the Texas Department of Human Resources (TDHR) to reflect a statewide mix of provider types and characteristics of interest. During several six-month cycles of the three-year experimental phase of the intervention, groups of unmarried male and female teens will be randomly assigned to the HBM intervention or a "control" educational program by the provider agency that recruited their participation. The impact of the educational interventions on beliefs, motivation, and sexual and fertility control behaviors will be compared twice over a 12-month follow-up period for each provider agency's clients by the University's project staff.

The actual intervention sessions will be delivered by specially trained volunteer health professionals and graduate students who will be recruited from among persons working or training in each specific community or area. The volunteers will present the HBM materials in a series of small group discussions covering approximately 12-15 hours. As suggested by our pilot work, this

mode of delivery should address the inconsistent levels of cognitive development and abstract thinking that characterize adolescents in this age range.

Once again, the author reminds the reviewers that pilot work has been completed on this subject, which reinforces the reader's impression that the proposed project will be successful.

The objectives of this controlled field test of the preventive services demonstration project are:

(1) To apply the Health Belief Model (HBM) to the problems associated with prevention and reduction of premarital sexuality and pregnancy among adolescents;
(2) To apply an educational intervention based on HBM concepts and components both to the training of adolescents who are not (yet) sexually active and to those who are (already) sexually active;

Three more objectives, relating to impact of the study, replication plans, and the value of this approach were included in the proposal but have been deleted here. This section on the objectives of the study comes early in the document and alerts the reader to the purposes of the study. The reviewer now has been provided with all of the information needed to read and appreciate the review of literature contained in the expanded Introduction below. This following section artfully combines aspects of the introduction, rationale, and literature review. Notice that it moves from general aspects of pregnancy in teenagers to more specific information important for the proposed study.

INTRODUCTION

Incidence of Adolescent Pregnancy

National and State of Texas projections indicate that approximately 50 percent of older adolescents (ages 15 to 19) are sexually active, and that about 10 to 15 percent of younger adolescents (aged 14 and under) are sexually active (Guttmacher Institute, 1980, 1981; TDHR, 1982). When these estimates are applied to the Texas adolescent population, the figures suggest

that about 650,000 older adolescents and 120,000 to 180,000 younger adolescents engage in sexual activity.

Because regular contraceptive use is not common among sexually active adolescents (Zelnik, Kantner, & Ford, 1981), the fertility rate is rather high and is increasing relative to other segments of the sexually active population. In Texas, about one in nine females aged 15 to 19 became pregnant in 1980 (TDH, 1982). This pregnancy rate (about 133 per 1,000 females aged 15 to 19) was the fifth highest in the nation. Of the 92,300 pregnancies of older adolescents in 1980, about one-third (approximately 31,000) occurred to unmarried women (Guttmacher Institute, 1981).

More than half of the pregnancies to older adolescents resulted in live births. These approximately 49,000 births resulted in a fertility rate in this age group of 70.5 per 1,000. Approximately 38 percent (about 1,000) of the pregnancies to younger adolescents resulted in live births, producing a fertility rate of 2.3 per 1,000. Out-of-wedlock births comprised about one-third of the births to older adolescents and nearly three-fourths of the births to younger adolescents (up 10 and 2 percent, respectively, for older and younger teenagers between 1970 and 1980). Combining data from both age groups, it is estimated that teenagers represent about 18 percent of the sexually active women in Texas who are capable of becoming pregnant, but account for almost one-half (46 percent) of all out-of-wedlock births (TDH, 1982).

These figures suggest strongly that the problem of adolescent sexual activity, pregnancy, and birth is even more severe in Texas than in the nation as a whole. When regional and ethnic trends are examined within the State, it can be seen that the problem is especially acute in certain cases. For example, older Hispanic and Black adolescents each account for a disproportional number of out-of-wedlock births in Texas (TDH, 1982). However, if national trends can be applied to Texas, such births have increased disproportionally for Anglo teens during the 1970's.

When fertility rates for older adolescents are examined on a regional basis within Texas, considerable variation is found. This variability is related in a rather complex manner with differences in the ethnic composition of the various regions. In Table 1, for each of the 12 health service regions in the State, (a) proportions of older adolescents falling into the three major ethnic groups, (b) the fertility rates for each ethnic group, and (c) the overall fertility

TABLE 1

Texas Fertility Rates[a] (age 15-19) by Health Service Area and Ethnicity/Race (1980)[b]

Health Service Areas	ANGLO		BLACK		HISPANIC		OVERALL
	% of Population	Fertility Rate	% of Population	Fertility Rate	% of Population	Fertility Rate	Regional Fertility Rate
1. Panhandle Amarillo	80.7	66.0	4.0	129.0	15.0	157.0	82.0
2. South Plains Lubbock	64.0	45.0	6.8	177.0	29.0	152.0	85.0
3. West Texas El Paso	29.5	52.0	3.6	86.0	66.9	62.0	59.6
4. Tri-Region Abilene	79.8	61.0	5.3	129.0	14.8	145.0	77.0
5. Area 5 Irving	75.0	49.0	15.0	135.0	9.6	122.0	69.4
6. Central Texas Austin	70.0	45.0	14.0	118.0	15.0	109.0	65.0
7. Northeast Texas Marshall	76.0	55.0	21.0	111.0	2.5	162.0	69.5
8. South Texas Kingsville	29.0	53.0	2.4	97.0	68.5	84.0	75.0
9. Camino Real San Antonio	42.6	42.0	5.8	92.0	51.7	86.0	66.9
10. Greater East Texas Beaumont	73.0	56.0	23.0	108.0	3.9	138.0	70.0
11. Houston - Galveston	62.9	47.0	19.9	105.0	17.0	110.0	69.1
12. Permian Basin	66.0	65.0	5.0	131.0	29.0	120.0	90.7

a - Live births/1,000 females 15-19
b - Source: Texas Department of Health (1982)

rate for each region are presented. It can be seen from this table that the regional fertility rates range from a low of almost 60 live births per 1,000 in Region 3 to a high of about 91 per 1,000 in Region 12.

While ethnic composition may be partially responsible for this variation, there is considerable region-to-region variation even within ethnic groups. For example, among Black adolescents, the fertility rate ranges from a low of 82 in Region 9 to a high of 177 in Region 2. This latter figure is about 2.5 times higher than the statewide fertility rate of older adolescents. Figures such as those presented in Table 1 may be useful for identifying target client groups or populations and specific regions for educational services in the proposed demonstration project (see Appendix A for a more detailed discussion of the adolescent pregnancy difference by regions).

Inclusion of Table 1 at this point allows the reviewer to continue reading without having minute detail that clutters the text. The table is comprehensive and makes a number of important points that are intended to convince the reviewer of the magnitude of the problem. This type of information is more effectively presented as a table than summarized in a paragraph. The author does go on, however, to focus the reviewer's attention on particular data in the table that have special relevance to the proposed project. Supplemental information on adolescent pregnancy is appropriately placed in an appendix.

Existing Preventive Services

Publicly funded family planning services are provided to adolescents in Texas through Titles X, XIX, and XX. These funds are used to provide both outreach and educational services, as well as contraceptive services.

Deleted here, this section continues, giving details of funding over the years and the numbers of teenagers receiving various forms of service.

Unmet Needs

The Alan Guttmacher Institute (1980) estimates that in 1979 about 315,570 Texas females aged 15-19 were in need of family planning services. The best estimates of the number of ad-

olescents receiving family planning services in Texas indicate that only about one-fourth of those in need are currently being served.

Deleted here, this section continues, providing detailed statistics supporting the fact that a large number of females at the project's target age are not receiving family planning services.

The General and the Specific Research Problem

The general research (and practical) problem is how to get teenagers who are sexually active to use effective contraception consistently and those who are not yet active to see the potential value of contracepting effectively, so that both groups will take action to reduce substantially their risk of having unintended premarital pregnancies. The specific problem addressed by this project is to field test an educational intervention mechanism that will help promote personal responsibility for each participant's sexual behavior and will enhance motivation for effective and consistent contraceptive usage by demonstrating that it produces significantly greater contraceptive usage and leads to less premarital pregnancy than presently used "control" educational services to 13-16 year-olds of both genders and various racial or ethnic groups throughout Texas.

This section is effective in presenting both the general problem and the specific focus of this research. Where the author discusses the "specific problem" the term is used synonymous with "purpose" as used in Chapter 2.

A Public Health Approach to the Problem

We believe that a preventive public health approach to the problems associated with adolescent sexuality and pregnancy that is fashioned within a salient theoretical framework has the best prospects for significant impact. Our approach, based on the Health Belief Model (HBM), seems particularly appropriate because it has been used with good success in various preventive programs for predicting both the initiation and continuing compliance of older children and adolescents; because it has been used to understand fertility control decision making and to predict

completed family size for married women; and because it suggests salient intervention points and modes to modify prevention-related motivations, beliefs, and behavior pertaining to fertility control through contraceptive usage (see Appendix B for a detailed review of these studies).

This is a complex, but exceptionally powerful review paragraph: Following a clear statement of commitment, the several situations are recalled in which the intervention plan (HBM) has been shown to work. In addition, the reader is given the choice of reading not only more detailed information about these situations, but critical reviews of research using the HBM. Reviewers will be impressed with the care with which the author has studied the relevant literature, and with the thorough, but thoughtful strategy of providing as much supporting material as possible without interrupting the flow of major ideas in the proposal.

The Model asserts that the probability that an individual will undertake a particular health measure is linked to a number of *personal* perceptions, including his/her *perceived susceptibility* to the disease or condition, the *perceived seriousness* of contracting the disease or developing the condition, and the *cost-benefit ratio* of available preventive health actions (see Figure 1 and Nathanson & Becker, 1983).

The paragraph above repeats information already provided, but serves to introduce Figure 1, which is an elegant demonstration of how the theoretical model (HBM) will be applied to this particular health problem. In the section below, the reviewer is provided with more detail about the evolution of the HBM model.

Development of a Health-Based Educational Intervention Model

Because the HBM-based approach to combatting adolescent sexuality and pregnancy was a new and novel one, the actual intervention content and structure required some development and shaping through pilot work and testing prior to the proposed large scale field testing around the State. The pilot phase took place in the Austin area during the first year of the project. The

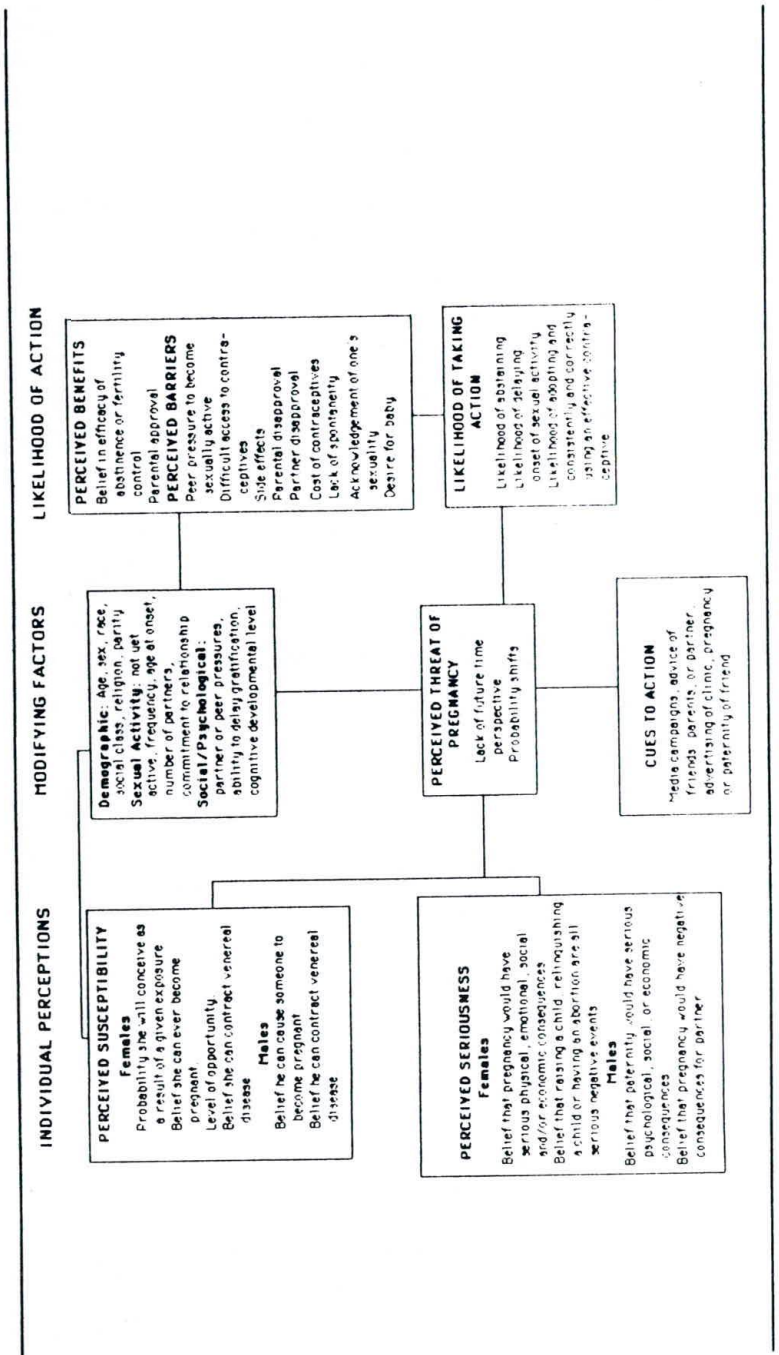


Figure 1

HBM conceptual framework and components to guide and focus the structure and content of the educational intervention stemmed from components of the HBM.

Deleted here, this section continues to explain the development of the HBM model, with frequent reference to specific components in Figure 1. The author guides the reader through the figure, so that the use of HBM in this particular project will be perfectly clear. The next sections explain in general terms the basis for evaluating the intervention and the relationship between that formative process and the planned statewide dissemination of the model. Although details of evaluation methodology are provided later in the section dealing with field testing, the matter of evaluation is touched on immediately after description of the proposed project. This serves to reassure reviewers about provisions for this critical feature.

Evaluation of the Proposed Experimental Intervention Model

The HBM-based preventive services project will be evaluated and its impact compared with each agency's regular program in terms of its *ability to develop, maintain, or increase individual sexual responsibility and fertility control* (i.e., self-discipline, abstinence, or consistent and effective contraceptive usage behavior patterns). Across provider agencies these comparisons will focus upon intervention effects for a wide range of client population subgroups: younger and older adolescents of both genders; all racial or ethnic groups represented in the State; adolescents differing in socio-economic status, income, and family characteristics; and those differing in preintervention sexual and reproductive knowledge, health beliefs, and sexual experience. The Department of Human Resources evaluation staff and The University of Texas at Austin staff will carry out the evaluation plan.

Dissemination of the Intervention Model

Over the projected three-year demonstration period modification and improvements will be made in the intervention and procedures on the basis of formal and informal evaluations. Thus on an iterative basis, the prevention model will be assessed in a

variety of settings, with a wide range of client populations, and against a reasonably representative cross-section of preventive services approaches employed by family planning service agencies around the State. If the HBM proves to meet the stated objective more effectively and/or more efficiently than other programs tested, it's our hope that it will be implemented on a programmatic basis statewide following the demonstration period.

Granting agencies are interested in how the proposed project is innovative or unique. It is important to include this as a separate section within the proposal so that reviewers do not have to infer the uniqueness of the study from other parts of the introduction. This proposal has a very strong section that stands out nicely and impresses the reader that here indeed is a fresh approach to a serious health problem. We have included below only the opening descriptive sentence of the first three innovative features described in this section.

Innovative Features of the Demonstration Project

The demonstration project has several innovative features:

- The educational services approach is built upon a well-developed conceptual framework about preventive health beliefs and behavior.
- The HBM approach is combined with a service delivery mechanism—small group discussions—that is especially suitable for use with adolescent audiences.
- The HBM framework allows focus on adolescents who are not (yet) sexually active, as well as those who are.

Expected Results and Benefits of the Field Test Phase of the Project

It is probably premature to generate a large number of specific hypotheses pertaining to empirical results in the Field Test Phase of the study. However, a set of working hypotheses regarding the results and benefits to be expected are:

Here the author presents several hypotheses which, due to the nature of this project, were not written as testable hypotheses. We have not included them because, as noted in Chapter 1, we believe that in most

cases it is preferable to write hypotheses in a form that can be directly accepted or rejected.

In the section that follows the author describes the various groups of subjects that will participate in the study. Since a number of types of subjects are needed, it was important to describe each. As examples, we have included here only two of the populations described in the proposal.

PROCEDURE

Participants in the Controlled Field Test

Family planning services providers. We propose that in collaboration with TDHC we select and then contract with various family planning services providers around the state who receive Title XIX and XX funds for participant recruitment, coordination, and selection of community facilities for the intervention sessions. Major selection criteria include:

(1) receipt of Title XIX and/or XX funds;
(2) representation of an important family planning service provider segment and contribution to the overall provider mix being established;
(3) having some type of educational services program in place at the time the study commences;
(4) having in place or being able to start-up easily a community outreach/recruitment campaign aimed at adolescents between 13 and 16 years of age; and
(5) relatively close proximity to relevant professional health organizations and graduate and professional training programs so that community and student volunteers who will be serving as the actual HBM instructors and discussion leaders can be recruited relatively easily.

Health professionals and discussion group leaders. Delivery of services to adolescents through family planning agency auspices will be provided on a volunteer basis by community professionals and by graduate students to fulfill course, experience, or internship requirements in their training or graduate programs. We propose that volunteers be recruited by the TDHR's Office of

Volunteer Services personnel in each region in association with the individual family planning service agencies with whom they ultimately work. These volunteers should include physicians, nurses, clinical and community psychologists, social workers, health educators, as well as graduate or professional students in these disciplines. Every effort should be made to match volunteer ethnic and cultural characteristics with their clients' characteristics when possible. These volunteer trainers and discussion leaders will be specially trained in the HBM approach and appropriate small group discussion techniques by The University of Texas at Austin project staff.

A lengthy section, not reproduced here, was next used to describe the adolescent client populations. In this section the author again employed lists of population characteristics followed by discussion designed to help the reader identify how the proposed project matched the nature of the clients.

As indicated in Chapter 3 it often is beneficial to use instruments that previously have been shown to be reliable and valid. At the beginning of the next section the author integrates already validated measures into the proposed project.

Educational Materials and Research Instruments

The general orientation to designing education materials and research instruments for this project has been to use those previously developed and readily available whenever possible. Therefore, most of the material and models relating to biological, reproductive, and birth control facts and information were selected from existing programs or research projects geared to adolescents of similar ages, backgrounds, and cultures as those being served in this project. Some materials relating to the use of small group methods to address adolescent sexuality and pregnancy issues were selected from available materials developed in other research projects and were modified to meet particular needs and project goals.

Much of the instructional material to train the small group discussion leaders and to distribute to adolescents during the intervention was developed by project consultants and staff to meet specific requirements and objectives. Thus, the materials described below continually will be developed and produced for

eventual dissemination to family planning agencies within the State (and perhaps in other areas):

In the full proposal, each of the following items was now presented with a short description of the material's intended use.

(1) Training designs and syllabus guides
(2) Discussion guides for small group leaders
(3) Materials to be provided to the discussion participants
(4) Materials to be provided to the discussion leaders

In the next paragraph, the author presents information on data collection instruments that had been developed in pilot studies. Since much of this material was still in development, draft copies were placed in appendices.

Data collection and research instruments were selected or developed over the course of an earlier pilot study. These instruments assessed areas such as the following: sexual and birth control knowledge, pregnancy and contraception health beliefs and perceptions, sexual activity and contraceptive history, Health Locus of Control, sociodemographic variables, and social relationships (see Appendix D and E for draft materials).

Content of the Educational Intervention

The anticipated content and the general format/structure for the HBM-based educational intervention program is discussed in detail in Appendix F.

In the deleted section above, the author gave a short description of the educational intervention and then referred the reader to an appendix where more detail could be found. The section below summarizes the time frame for the study. Much of the data collection was scheduled to take place over an extended period of time. The use of tables makes it easy to follow which measures will be employed at which points in the field test.

Procedure for the Field Test Phase

Following participant recruitment within each agency, we expect adolescents to be *randomly assigned* to that agency's

HBM-based educational program or to their regular (ongoing) educational services program. Once the teenagers have completed the preventive services intervention (HBM or regular), they will be followed over at least a 12-month period with process and outcome data collected one week after intervention, and six months and 12 months post-intervention by paid student project interviewers who are coordinated by each agency participating in the study. More specifically, the procedure will involve four data collection points during a 12-month period: a pretest session and group educational (or control) small group discussion to be conducted over 12-15 hours, (Time 1); a posttest follow-up one week after the interventions (which will be individually scheduled, Time 2); a follow-up to collect dependent variables at six months post-intervention (individually scheduled, Time 3); and a 12-month follow-up to collect dependent variable data and debrief the participants (individually scheduled, Time 4). (See Tables 2 and 3 for more details.)

A second paragraph, omitted here, contained a short overview of Tables 2 and 3. The variables were introduced in a manner that allows a reader to follow the text without having to spend a substantial amount of time discovering how to use the tables.

Note that on Table 3, a new variable (#1—"Pregnant or not") is added to the sequence of dependent variables in columns 2 and 3. It might help the reader to leave a blank space in the first column where "Consistent use of contraceptives in reporting period" is and move numbers 1 through 7 down so each variable is aligned across the page. If the reviewers are likely to be unfamiliar with how each variable will be measured, a short note after the variable name (e.g., "yes" or "no" after "Contraceptives used at last intercourse") would help to clarify that point.

In the first part of this next section the author reminds the reviewer of the study goals. This excellent strategy serves to emphasize the relationship between the purpose of the study and the data analysis.

PLAN OF DATA ANALYSIS

Overview

The analyses are designed to answer the following two evaluation questions:

TABLE 2
Experimental Design for the Proposed Study

	Time 1	Time 2 (1 week)	Time 3 (6 months)	Time 4 (12 months)
Experimental Groups	1) Pretest: Sex Knowledge Contraceptive Knowledge Health Beliefs Personality Scales Sociodemographic Data Sex and Contraception History 2) Intervention	1) Posttest: Sex Knowledge Contraceptive Knowledge Health Beliefs Manipulation Checks 2) Posttest (same as above)	1) Health Beliefs Dependent Variables	1) Health Beliefs 2) Sex and Contraception
Control Groups	1) Pretest (same as above) 2) Control Discussion	1) Posttest (same as above) 2) No Sex and Contraception Dependent Variables	(same as above)	(same as above)

TABLE 3
Dependent Variables for the Proposed Study

Time 2 (1 week post)	Time 3 (6 months post)	Time 4 (12 months post)
1) Consistent use of contraceptives in reporting period	1) Pregnant or not (Male: responsible for pregnancy)	1) Pregnant or not (Male: responsible for pregnancy)
2) Contraceptives used at last intercourse	2) Consistent use of contraceptives in reporting period	2) Consistent use of contraceptives in reporting period
3) Family planning advice sought or fertility control program enrollment	3) Contraceptives used at last intercourse	3) Contraceptives used at last intercourse
4) More information on sex and contraception sought	4) Family planning advice sought or fertility control program enrollment	4) Family planning advice sought or fertility control program enrollment
5) Ceased or reduced level of sexual activity	5) More information on sex and contraception sought	5) More information on sex and contraception sought
6) Became sexually active for first time; use of contraception	6) Ceased or reduced level of sexual activity	6) Ceased or reduced level of sexual activity
7) Consistent abstinence	7) Became sexually active for first time; use of contraception	7) Became sexually active for first time; use of contraception
	8) Consistent abstinence	8) Consistent abstinence

(1) As a result of the HBM intervention, do subjects in the experimental groups exhibit more positive sexual behaviors on follow-up than subjects in the control groups?
(2) Did the intervention have a more positive effect on sexual behaviors for subjects with specific characteristics in the experimental groups (i.e., subjects of certain ages, genders, sexual experience, ethnic groups, personality types)?

These questions seek to determine whether the intervention produced "more positive sexual and fertility control behaviors," where more positive sexual behaviors are defined as:

(1) Lower rates of pregnancy (for males: responsible for fewer pregnancies);
(2) Higher rates of reported contraceptive use;
(3) Greater consistency of reported contraceptive use (contraceptives used for a higher proportion of instances of intercourse);
(4) Higher rates of reported contraceptive use at most recent intercourse;
(5) Higher rates of enrollment in family planning programs;
(6) Advice on sex and contraception sought more often (e.g., from family planning service providers and other sources);
(7) Lower rates of reported sexual activity; and
(8) Longer delays before first intercourse (among those who were previously not sexually active).

The evaluation plan is designed to determine, for each of these outcomes, whether the intervention was effective and for which groups of subjects it was most effective.

In many instances information on study design and sample size would have been presented much earlier in the proposal. In this instance, the delay seems justified by the other important tasks that took logical precedence. As illustrated here, study design and statistical analysis techniques often are closely related. Grouping the two sections together, as below, may improve continuity in some proposals.

In the section on sample size, results from pilot work again are used to support the decisions made. As indicated in Chapter 3, every proposal should provide a rationale for the number of subjects selected. Pilot data are particularly useful in establishing the number of subjects required to obtain the desired inferential power. The author also uses the pilot data

on subject attrition to justify the cost of the project. This is an excellent way to convince the foundation that the projected costs for paying subjects are legitimate and needed.

Design

Subjects, recruited by family planning services providers through outreach programs, will be randomly assigned to experimental HBM and control groups. Groups will be further segregated according to age: younger teenagers (13 to 14 years of age) and older teenagers (15 to 16 years of age). Thus, subjects will be assigned to groups according to a two-by-two design in which one between-groups factor is treatment (Experimental vs. Control) and the second is age (Younger vs. Older), for a total of four cells or groups.

Half of the subjects will be randomly assigned to experimental groups, while the other half will be assigned to control groups. Because subjects will be assigned to age-dependent groups according to their age ranges, the sample sizes of these groups will reflect the proportions with which younger and older teens are recruited. It is anticipated that approximately two-thirds of the teens will be older (15 to 16 years of age).

Subjects assigned to the experimental treatment will be placed in HBM-based discussion groups. Control subjects will received the educational programs currently offered by providers. The HBM-based discussion groups will meet over a three-week period. Data will be collected from these subjects before the first discussion session (the pretest) as well as after the last session (the one week posttest). Thus, it should be possible to process one cohort of experimental subjects through the pretest, the discussion groups, and the posttest within a month.

Sample Size

Discussions with Austin area family planning service providers indicate that a provider of average size should be able to process approximately 50 subjects per month. If each provider processes 50 participants per month for six months, it will process about 300 subjects in a six-month period. Four providers delivering the intervention during each six-month period, will result in a total of

1,200 subjects receiving services (2,400/year). Half of these (600) will receive the experimental treatment and half (600) will receive the control treatment. Within each of these treatments, about two-thirds will be older teens and one-third young teens. Thus, among older teens, there will be about 400 experimental and 400 control subjects. Among younger teens there will be about 200 experimental subjects and 200 control subjects per six-month period.

Based on our pilot study attrition data for paid six-month participants, it is conservatively estimated that about half of the unpaid subjects will be available for data collection by the time of the second follow-up (one year). It does not seem unreasonable to anticipate this level of cooperation in a study that deals with a mobile population, depends upon voluntary cooperation, and seeks to collect sensitive information without monetary compensation.

If this estimate is correct, about half of the 2,400 subjects served will be available for follow-up. These 1,200 subjects per 12 months will be sufficient for the data analyses planned. It is considered sufficient to obtain complete data from about 1,200 subjects (300 subjects from each of the four cells in the two-by-two) each 12 months.

Thus, only a sample of the population served will receive testing. Testing a sample, rather than the entire population, has the salutory effect of reducing the scope and expense of the data collection effort. A sample of the size described above will provide quite adequate statistical power for the analyses planned. In order to ensure a sufficient number of subjects with complete data, data will be collected from a larger number during the earlier data collection periods. These sample sizes will be determined while taking into account the expected rate of attrition and the number of subjects served at each site. Subjects will be selected for testing using a proportional selection procedure, so that equal proportions of subjects served will be selected from each provider site.

The data analysis portion of the proposal is subdivided into two categories: major and supporting data analyses. Here the author reacquaints the reader with the purposes and independent variables for the study. In addition, where a little known technique is used (logistical

regression), the author adds a few sentences to help the reader understand why the technique is appropriate. The author does not, however, overwhelm the reader with unneeded statistical jargon.

Major Data Analyses

Independent variables will be of two types. The first type consists of between-group variables, namely treatment (Experimental vs. Control) and age (Young Teens vs. Older Teens). The second type of independent variable will be individual differences variables such ethnicity, amount of previous sexual experience, pretest health beliefs, and various personality variables. These individual differences variables are characteristics of the subjects that will vary within groups (e.g., a particular discussion group may have subjects of various ethnic groups). However, the groups will be pure with regard to the between-groups variables (e.g., all subjects in a particular discussion group may be young teens receiving the experimental treatment).

Data will be analyzed with two goals in mind (corresponding to the two evaluation questions discussed previously). First, it will be determined whether the experimental subjects have more successful outcomes than control subjects. Second, it will be determined whether the experimental treatment was more successful for certain groups or types of subjects than for others.

Analysis-of-variance (ANOVA) is the most powerful statistical procedure available for this type of analysis and will be used when appropriate. In terms of ANOVA, the first goal involves testing for a main effect for the treatment variable, while the second goal involves testing for interactions between the treatment variable and the other between-group and within-group (individual differences) variables. In practice, these tests are performed and interpreted simultaneously.

Where the dependent variable is continuous (e.g., amount of sexual activity), a repeated-measures ANOVA can be used, employing the follow-up periods as repeated measures. It may prove useful to include certain pretest measures or personality measures as covariates (i.e., to hold these individual differences constant) in these analyses or to construct linear regression (i.e., multivariate) models to control for potential individual difference variables while testing the necessary main effects and interactions.

However, several of the dependent variables will be binary variables (pregnant vs. not pregnant, used contraception vs. did not use contraception, etc.). Such binary dependent variables make the use of ANOVA or linear regression inappropriate. For the analysis of these variables, logistic regression will be used. Logistic regression is a multivariate technique analogous to ordinary linear regression but is designed for the analysis of a binary dependent variable when the independent variables are categorical (e.g., Anglo, Black, Hispanic) and/or continuous (e.g., scores on personality measure). This method involves estimating the probability of "success" (e.g., no pregnancy, used contraception, etc.) for each combination of the predictor variables.

Thus, for the dependent variables, which are binary in nature, logistical regression will be used in a manner analogous to linear regression to assess the main effects and interactions described above. A computer program to perform logistical regression is available in the Statistical Analysis System (SAS) statistical package. This package is available on the IBM computer operated by The University of Texas at Austin.

In summary, ANOVA, linear regression, and logistical regression will be used, as appropriate, to test the effectiveness of the treatment and to determine whether it is more effective for certain types of subjects (subjects of certain ages, ethnicity, sexual experience, personality types, etc.).

In the paragraph above, the author summarizes and reminds the reader of the techniques that will be used. The use of a summary at the end of an extensive section involving technical details helps the reader pick out and retain the essential elements.

In the next section, the supporting data analyses are presented. Note that in the second paragraph the author describes plans for analyzing the effect of attrition on the study outcome. This type of analysis often is included in social scientific research because it allows the author to know if attrition within and among the groups might have changed the results of the study.

Supporting Data Analyses

The major data analyses are designed to answer the evaluation questions. However, the data collected in this study will be used

for several additional purposes: (1) to assess the effects of attrition, (2) to determine the degree to which the treatment was implemented as intended, and (3) to test certain hypotheses concerning the relationship between the HBM and sexual knowledge, beliefs, and behavior.

It is expected that a number of subjects will be lost through attrition. Some subjects who begin the group discussions will not complete them, and others will not be available for the first or second follow-up. In an investigation of this type, it is important to determine the attrition rates for the groups of interest and to determine the impact of attrition on the outcome variables. This is necessary to avoid attributing outcome effects to treatment variables when they are actually a result of differential attrition. Analyses of the attrition rates for the various treatment groups will, therefore, be conducted to detect differential rates of attrition. Comparisons will also be made on the pretest measures between those who completed and did not complete the treatment in order to determine whether systematic differences exist between these two groups.

Several omitted paragraphs dealt here with a discussion of variables that might have changed from pretest to posttest. In addition, the author identified the correlations that would be of primary interest.

The next section presents some limitations of the study. It always is valuable to confront the limitations in any study design, and to describe the ways in which you will try to lessen their effects. The treatment of the first limitation, attrition, is a model for how this can be accomplished.

Limitations

Four limitations inherent in the sample and in the longitudinal design of the demonstration project may reduce the power of statistical analyses and limit the internal validity and generalizability of the findings. First, attrition of participants is a potential problem because it poses a threat to the power of the statistical analyses for testing experimental hypotheses. Since many of the dependent variables in the study are dichotomous (e.g., yes/no) losing participants directly affects the power of the tests. Moreover, attrition may not be distributed randomly among conditions. Interpretation of differences between experimental and control

subjects (at 6 and 12 months) could be confounded by the characteristics of those who drop out versus those who remain. Realistically, it is expected that it will not be possible to collect follow-up data for a relatively large proportion of the subjects. Attempts will be made to determine whether there is differential attrition for certain types of subjects and whether those who drop out differ on various pretest measures from those who do not. However, it will not be possible to correct for these differences, if they occur. Thus, inferences will be generalizable only to those subjects who complete the study.

A number of actions will be taken to reduce the attrition rate:

(1) We will impress upon subjects that this is an important study which depends heavily on the follow-up component.
(2) Appointments for succeeding follow-ups will be made at each interview to underscore the fact that follow-up interviews will occur.
(3) Pretest and follow-up interviews will be on a one-to-one basis, and follow-ups will be scheduled to occur at times and places most convenient for subjects.
(4) Some participants will probably come from TDHR client groups; thus their locations are potentially available for follow-up reminders and phone calls.
(5) Where appropriate, letters will be sent to participants to remind them of their 6 and 12 month follow-up appointments.
(6) Follow-up phone calls will be made as necessary. These and other follow-up procedures will be conducted by the same person who interviewed the subject at Time 1. In this way, subjects may develop a more personal relationship with the interviewer which will reduce the likelihood of attrition.

The author follows by discussing other limitations in a similar fashion. This material has been omitted. Below, in presenting the fourth and final limitation, the author discusses the problem presented by self-reports involving sensitive personal data. This potentially serious source of bias is treated frankly and carefully. Information concerning measurement strategies that appear to have lessened the impact of self-report bias is used to assure reviewers that the author is fully sensitive to this potential limitation.

Finally, the potential underreporting bias in self-reports of sexual and contraceptive behavior is a problem with no really

satisfactory solution. The extent of underreporting of sexual activities and contraceptive usage is not known in a general teenage population, or, for that matter, in more selected populations (e.g., family planning clinic users). Typically, efforts to validate these reports depend on validation criteria that in themselves are self-report data (e.g., number of previous pregnancies gathered from a medical history).

Some efforts have been made to employ measurement techniques designed to reduce self-report bias. Zelnik, Kantner and Ford (1981) found that there was little difference in reported incidence of sexual intercourse when interviewers asked the question directly or asked the question within a *randomized response technique* format in their 1976 sample of 15-19 year old females. No data for males are available on this point. Others have examined potential response biases in sex surveys and found little underreporting for either sex among 18-22 year-olds (Delameter, 1974; Delameter and MacCorquodale, 1979).

The self-reporting bias is not necessarily a major problem in the present study unless participants' underreporting of sexual activity or overreporting of contraceptive usage interacts with the experimental intervention. Thus, if participants exposed to the HBM intervention are more likely to underreport sexual activity or to underreport contraceptive usage, the group means will be affected and interpretations of causality could be clouded. Thus, it might not be clear whether the treatment was effective or whether the treatment simply led to a higher level of response bias. Again, pretest (i.e., baseline) data on personality and sociodemographic variables may provide some help in eliminating alternative interpretations of treatment effects.

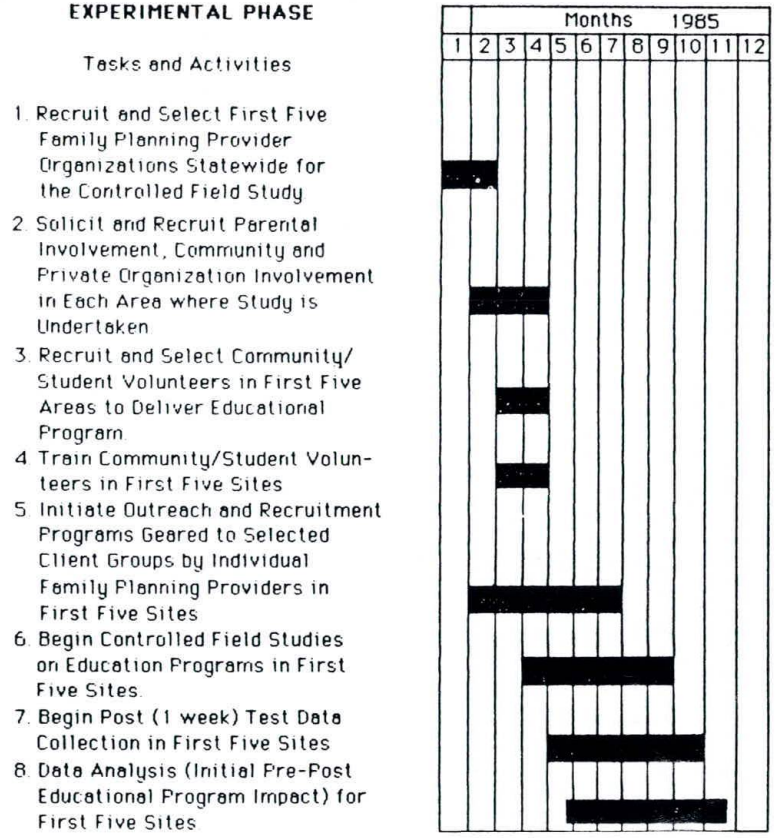
The use of a timetable for indicating when each part of the study will be completed is valuable for both the reviewers and investigators (Table 4). The format of this table is particularly valuable because it progressively shows each step for completing the study. In the interest of space, only the first eight tasks and activities are presented here as examples. The use of both flow diagrams (Chapter 1) and projected time tables (Chapter 7) have been discussed.

The reference section, which is not presented here, contained the 72 references cited in the text of the proposal. Since a specific format for references was not required by the foundation, references were listed in a

TABLE 4

DRAFT WORK PLAN AND TIMETABLE (EXAMPLE: 1ST CYCLE)

EXPERIMENTAL PHASE



format style common to sociological/public health research journals.

The final section of the proposal is the budget. Since this project was funded by a number of different cooperating sources, we have combined and edited the budget for this example. Note that the budget is subdivided by time periods and categories of funding and that each subdivision and category has a separate heading. We have included only

the first year of the budget and the summary. The budget for the second and third years has similar categories and format. The author shows in short phrases the method by which he arrived at the dollar figure for each category (see Chapter 7). In the text of the proposal, the author already has discussed the number of subjects needed and other factors that will help reviewers understand the need for particular expenditures before they read the budget.

Note that we have not included actual salary or fringe benefit figures with this proposal. The proposal submitted to funding agencies includes this information in the format of the budget presented here.

BUDGET

YEAR ONE: June 1, 1985-May 31, 1986

SALARIES:	Salary	Fringe	
Project Director (50% time for 12 months)	$xx,xxx.	$ x,xxx.	
Research Associate (50% time for 12 months)	x,xxx.	x,xxx.	
Programmer I (50% time for 6 months)	x,xxx.	x,xxx.	
Data Entry Operator (100% time for 6 months)	x,xxx.	x,xxx.	
Interviewers (3000 hours × $6.47/hour)	19,410.		
Trainers (Graduate Students) ($100/day × 3 days × 10 trips)	3,000.		
TOTAL SALARY AND FRINGE BENEFITS:	xx,xxx.	x,xxx.	xx,xxx.
ADMINISTRATIVE FEES:			
Agency fee for administrative details (10 sites × $1,200/site)	12,000.		
TOTAL ADMINISTRATIVE FEES:			12,000.
TRAVEL:			
Training Travel:			
Three people × 3 days × $70/day × 10 trips	6,300.		
Three people × $100 travel × 10 trips	3,000.		
Total training travel	9,300.		

BUDGET--Continued

Project Director Travel:		
Ten trips/year, one to each site		
One day at $70 per diem × 10 trips	700.	
Travel at $100/trip	1,000.	
Total Project Director Travel:	1,700	
TOTAL PROJECT TRAVEL:		11,000.
OTHER EXPENSES:		
Expendable Supplies ($100/month × 12 months)	1,200	
Telephone Charges ($150/month × 12 months)	1,800.	
Duplication Charges ($50/month × 12 months)	600.	
TOTAL OTHER EXPENSES:		3,600.
DATA PROCESSING:		
Computer Connect Time (7500 hours × $0.20/hour)	1,500.	
Computer CPU Time (5 hours × $230/hour)	1,150	
TOTAL DATA PROCESSING:		2,650.
GRAND TOTAL FOR YEAR ONE:		$xxx,xxx.

Omitted here were similar budgets for the second and third year of the study.

SUMMARY BUDGET: June 1, 1985-May 31, 1988

	YEAR 1	YEAR 2	YEAR 3
SALARY AND FRINGE BENEFITS	xx,xxx.	xxx,xxx.	xx,xxx.
ADMINISTRATIVE FEES	12,000.	12,000.	—
TRAVEL	11,000.	11,000.	—
OTHER EXPENSES	3,600.	3,600.	3,600.
DATA PROCESSING	2,650.	4,300.	3,300.
TOTALS:	xxx,xxx.	xxx,xxx.	xx,xxx.

TOTAL DIRECT COSTS FOR THREE YEAR PROJECT: $xxx,xxx.

In the full proposal the budget was followed by a comprehensive justification of each line item. Funding agencies require an explanation of how the investigator plans to use the funding if the proposal is approved. Inasmuch as the request for personnel was a large item in this proposal, the author listed the job responsibilities for each position. It usually is helpful if the investigator has particular personnel in mind for each position, and includes evidence of their experiences and expertise. It is even more impressive if the persons to be appointed already have been working

with the pilot project. In that sense, funding the proposal simply would serve to retain previously trained, experienced persons in the proposed project, an obvious advantage to the grantor.

The second area of the budget that merited careful attention was the request for travel funds. Such requests must be given in detail and justified in terms of project demands. Reviewers are alert to detect potential abuses of their resources, and frivolous travel definitely fits that category. The guiding rule to remember when building the budget is that the more carefully the budget is calculated and explained, the more convinced the reviewers will be that the investigator understands what must be done, has been conscientious in planning each step, and not only will accomplish the goals of the study but will do so at minimum cost.

APPENDIX A

Some General Standards for Judging the Acceptability of a Thesis or Dissertation Proposal

	Desirable	*Undesirable*
I. Topic		
A. Importance		
1. Basic Research	A clear relationship exists between the topic and existing information in related areas of knowledge. Topic is recognized as substantial by people who are knowledgeable in the area. Topic is articulated to a body of knowledge recognized as broadly relevant to the discipline.	Proposal does not support the importance of the study. Topic seems unrelated to existing facts and theoretical constructs. Proposed study is not inserted into a line of inquiry.
2. Applied Research	Topic is relevant to professional needs, and recognized as substantial by competent individuals engaged in professional practice. There is a clear relation between the topic and existing problems in practice.	Topic seems unrelated to realistic professional concerns and divorced from matters of practice.
B. Scope	The extent of the proposed study is reasonable in terms of the time and resources available to the candidate. A clear indication exists that the student has considered and made provision for each of the demands implicit within the study.	Projected study is grandiose and unreasonable in terms of time and resources. Or, the study is so small or limited in its concern that it may (a) provide little useful information, and (b) involve less than a reasonable exposure to scholarly inquiry for the candidate.

(continued)

Special Article

Guidelines for reading literature reviews

Andrew D. Oxman, MD
Gordon H. Guyatt, MD

One strategy for dealing with the burgeoning medical literature is to rely on reviews of the literature. Although this strategy is efficient, readers may be misled if the review does not meet scientific standards. Therefore, guidelines that will help readers assess the scientific quality of the review are proposed. The guidelines focus on the definition of the question, the comprehensiveness of the search strategy, the methods of choosing and assessing the primary studies, and the methods of combining the results and reaching appropriate conclusions. Application of the guidelines will allow clinicians to spend their valuable reading time on high-quality material and to judge the validity of an author's conclusions.

Une façon efficace de se tenir au courant de la littérature médicale toujours plus abondante c'est de se rabattre sur les revues générales. Mais si celles-ci ne se conforment pas aux normes scientifiques, elles risquent d'induire en erreur. Il est proposé ici des lignes directrices afin d'aider le lecteur à apprécier la qualité scientifique d'une revue générale. Elles s'attachent à déterminer si la question y est bien énoncée, la recherche bibliographique est complète, les travaux retenus sont bien choisis et bien analysés, et les divers résultats sont mis en regard de façon à cerner des conclusions valables. En suivant ces lignes directrices le clinicien utilisera son temps précieux à bon escient.

From the departments of Clinical Epidemiology and Biostatistics and of Medicine, McMaster University, Hamilton, Ont.

Dr. Guyatt is a career scientist of the Ontario Ministry of Health.

Reprint requests to: Dr. Gordon H. Guyatt, McMaster University Health Sciences Centre, 3H7-1200 Main St. W, Hamilton, Ont. L8N 3Z5

Clinicians who are attempting to keep abreast of developments must find ways to deal with the exponentially expanding literature. Efficient strategies for finding and storing relevant studies[1-6] and for discarding invalid or inapplicable studies[7-12] are available. However, processing the literature for an answer to a clinical question remains time consuming, and it is not feasible for clinicians to read all the primary literature for each of the myriad clinical issues that confront them daily.

One solution to this problem is the literature review or overview in which the primary research relevant to a clinical question is examined and summarized. However, reviews, as well as primary studies, must be read selectively and critically. Just as flawed methods in a study of diagnosis or therapy may invalidate the results, an unscientific literature review may come to incorrect conclusions. Authors of reviews do collect and analyse data from primary research, although this is sometimes done subjectively and subconsciously. The fundamental difference between a review and a primary study is the unit of analysis, not the scientific principles that apply.

Five conflicting recommendations for managing mild hypertension, quoted from the literature, are shown below.

- The available data . . . lead this reviewer to conclude that treatment of mild hypertension [90 to 104 mm Hg] to achieve diastolic pressures below 90 mm Hg is the appropriate public health policy based on current evidence.[13]
- Most patients with diastolic blood pressure in the 90 to 104 mm Hg range should be treated unless contraindications to drug therapy exist. . . . In certain patients, vigorous dietary and behavioral modifications may be attempted before instituting or as an adjunct to pharmacologic therapy.[14]
- Non-drug measures are often effective for mild hypertension. The initial choice between thiazides and beta-adrenoceptor blocking drugs often depends on the physician's personal preference. . . . With care, the risks

← For prescribing information see page 713

of antihypertensive therapy are considerably less than the benefits.[15]

- The benefits of drug treatment for patients with mild hypertension [diastolic blood pressure between 90 and 105 mm Hg] remain unproven. Non-drug therapy has also been insufficiently investigated.[16]
- At present, therefore, with the diuretic-based treatments principally studied in the previous trials, treatment of mild-to-moderate hypertension [diastolic blood pressure below 115 mm Hg] is of directly demonstrated value only if the stroke rate is high enough (perhaps due to age or cerebrovascular disease) for halving it to justify the costs and trouble of therapy. . . . Lipid-sparing antihypertensives might have more important effects on MI [myocardial infarction] than on stroke. But, in the trials reviewed, the size of the MI reduction remains uncertain [Rory Collins: unpublished observations, 1987].

If one doesn't have some guidelines for assessing the reviews from which these recommendations are taken, deciding which review to believe is like deciding which toothpaste to use. It is a question of taste rather than a question of science.

One does not have to look far to find other examples of important clinical questions for which recent reviews have come to different conclusions: Should clinicians avoid administering corticosteroids because of concern about clinically important osteoporosis?[17,18] What are the benefits to critically ill patients of catheterizing the right side of the heart?[19,20] Should mild hypokalemia be treated aggressively?[21,22]

Clearly, the expertise of the author is not a sufficient criterion of a review's credibility, since experts reviewing the same topic often come to different conclusions. Nor is the prestige of the journal or textbook in which the review is published a sufficient criterion. Recent surveys of the medical literature have found that the scientific quality of most published reviews, including those in the most highly regarded journals, is poor.[23-27]

In this article we present a reader's guide to assessing research reviews. Similar guidelines have been suggested before, particularly in the psychology and social science literature.[28-30] We focus on how readers of the medical literature can decide whether a review is worth reading and whether its conclusions are to be believed. Our guidelines may also be of use to those planning to write a research review.

Guidelines

We have framed our guidelines as a series of questions (Table I). Before we address each item in detail some general comments are warranted. First, the questions are intended to be used to assess overviews of primary studies on pragmatic questions. Second, the term "primary studies" refers to research reports that contain original information on which the review is based. Third, the intention of the guidelines is to encourage efficient use of the medical literature and a healthy scepticism, not to promote nihilism. Readers who apply these guidelines will find that most published reviews have major scientific flaws.[23-27] Indeed, surveys on the scientific adequacy of medical research reports have found that most primary studies also have major scientific flaws.[25]

There is a need for improvement in the design, implementation and reporting of both reviews and primary studies. None the less, vast amounts of valuable information exist, and to make informed decisions clinicians must use the research available. Although most published reviews do not provide strong support for their conclusions, critical readers can discern useful information and make their own inferences, which may or may not be the same as those of the authors.

Were the questions and methods clearly stated?

When examining a review article readers must decide whether the review addresses a question that is relevant to their clinical practice or interests. They therefore require a clear statement of the questions being addressed.

Table I — Guidelines for assessing research reviews

Were the questions and methods clearly stated?
Were comprehensive search methods used to locate relevant studies?
Were explicit methods used to determine which articles to include in the review?
Was the validity of the primary studies assessed?
Was the assessment of the primary studies reproducible and free from bias?
Was variation in the findings of the relevant studies analysed?
Were the findings of the primary studies combined appropriately?
Were the reviewers' conclusions supported by the data cited?

Table II — Examples of the elements of a causal question

Nature of the question	Population	Exposure/ intervention	Outcome
Etiology	Homosexual men	Human immunodeficiency virus	Acquired immune deficiency syndrome
Diagnosis	Patients with head trauma	Computerized tomography	Hemorrhage
Prognosis	Patients with ulcerative colitis	Ulcerative colitis	Cancer of the colon
Therapy	Patients with Alzheimer's disease	Cholinomimetic agents	Functional status
Prevention	Postmenopausal women	Calcium supplementation	Hip fracture

Any causal question has three key elements: the population, the exposure or intervention and the outcome. Examples of these elements in five key areas of clinical inquiry are presented in Table II. A clear statement of the question requires explicit specification of all three elements if the reader is to quickly decide whether the review is relevant. If there is no clear statement of the questions being addressed at the beginning of the review the reader might as well stop. Fuzzy questions tend to lead to fuzzy answers.

Many reviews address several questions; for example, an article or a chapter in a textbook about acquired immune deficiency syndrome may review what is known about the cause, diagnosis, prognosis, treatment and prevention of the disease. Such reviews may be extremely helpful for readers seeking a broad overview. However, they tend to provide little, if any, support for most of the inferences they make. Typically, an inference is presented as a fact followed by one or more citations. In this case the reader has no basis upon which to judge the strength or validity of the inferences without reading the articles that are cited. Readers seeking answers to specific clinical questions should not rely on reviews that address broad topics and encompass many questions.

In addition, an explicit statement of the methods used for the research review is necessary for the reader to make an informed assessment of the scientific rigour of the review and the strength of the support for the review's inferences. Unfortunately, this information is often lacking. In general, when a review does not state how something was done — for example, how it was decided which primary studies would be included — it is reasonable to assume that it was not done rigorously and that a threat to the validity of the review exists. Readers looking for answers to specific clinical questions should seek reviews that clearly report the methods used. Without knowing the authors' methods the reader cannot distinguish statements based on evidence from those based on the opinions of the authors.

Were comprehensive search methods used to locate relevant studies?

It is surprisingly difficult to locate all the published research in a particular area, even when the area is relatively circumscribed.[31-33] For example, Dickersin and associates[33] found that a MEDLINE search yielded only 29% of the relevant trials on the prevention and treatment of perinatal hyperbilirubinemia.

This problem is exacerbated by the fact that some of the relevant material may not even be published. Furthermore, the unpublished studies may be systematically different from those that have appeared in peer-reviewed journals, not in that their methods are flawed but in that their results are "negative". Research has suggested that of two articles that use the same methods to investigate a question the study yielding positive results is more likely to be published than the one yielding negative results.[33-37] Research conducted by an agency that has an investment in the treatment being studied (such as a pharmaceutical company with a new drug) may not even be submitted for publication if its results are negative. It thus behoves an author to try to determine the extent of the "publication bias" in the area being reviewed.

Authors' search strategies vary widely, and experts are no more likely than nonexperts to be systematic in their search.[38] The more selective or haphazard the authors' search for papers the more likely it is that there will be bias in the review. For example, authors are likely to attend to papers that support their preconceptions.

The reader needs assurance that all the pertinent and important literature has been included in the review. The more comprehensive the authors' search the more likely it is that all the important articles have been found. The reader should look for an explicit statement of the search strategies used. Ideally, such strategies include the use of one or more bibliographic databases (including a specification of the key words and other aspects of the search strategies[39]), a search for reports that cite the important papers found through a database such as the Science Citation Index, perusal of the references of all the relevant papers found and personal communication with investigators or organizations active in the area being reviewed (to make sure important published papers have not been missed and particularly to look for methodologically adequate studies that have not been published).

Were explicit methods used to determine which articles to include in the review?

A comprehensive literature search will yield many articles that may not be directly relevant to the question under investigation or that may be so methodologically weak that they do not contribute valid information. The authors must therefore select those that are appropriate for inclusion in the review. When, as is often the case, this process is unsystematic, opportunities for bias develop. Thus, it is common to find two reviews of the same question in which different primary studies are included and for the choice of studies to contribute to different conclusions. For example, in two methodologically sophisticated and carefully conducted reviews on whether corticosteroids are associated with peptic ulcer the two teams of authors used different criteria for choosing which studies would be included in the review.[40,41] This difference was the main reason for the remarkable result of the two reviews: diametrically opposed conclusions about whether or not the association exists.

The authors should specify how the articles were chosen by referring to the three basic ele-

ments of primary studies: the population, the exposure or intervention and the outcome. For example, in assessing the effect of cholinomimetic agents in patients with dementia the authors could specify the criteria as follows.

- Population: patients with senile dementia in whom causes other than Alzheimer's disease were excluded.
- Intervention: oral administration of cholinomimetic agents.
- Outcome: indicated by measurements of both memory and functional status.

Other methodologic criteria may be used to select primary papers for review. In this example the authors may consider only studies in which patients were selected at random to receive the treatment drug or a placebo and in which both the investigator and the patient were blind to allocation.

Was the validity of the primary studies assessed?

Authors will come to correct conclusions only if they accurately assess the validity of the primary studies on which the review is based. If all the studies have basic flaws their conclusions may be questionable even if their results are comparable. For example, if the literature on extracranial–intracranial bypass surgery for threatened stroke were reviewed before the results of a recent randomized controlled trial[42] were published, a large number of studies with positive results but of suboptimal design and thus open to bias would have been found. The appropriate conclusion would have been that the procedure's effectiveness was still open to question, despite the volume of studies with positive results; indeed, the subsequent trial showed no benefit of surgical over medical therapy.

Methodologic guidelines for studies of etiology,[10,43] diagnosis,[8] prognosis[9] and therapy[11,44] are available. In a study of therapy one is interested in whether the allocation to treatment was random, whether the subjects and investigators were blind to the allocation, and whether all the relevant outcomes were monitored. Important aspects of the design and conduct of each primary study should be critiqued and the standard used in these critiques made explicit. Critiques should be reported in sufficient detail to allow readers to judge the methodologic quality of the primary studies. Although a study-by-study critique can be tedious, presentation of the methodologic assessment in a table may allow a rapid assessment of validity. Readers should be wary of any review that focuses on the results of studies without thoroughly discussing the methods that were used to arrive at the results.

When information about the methods or results has been omitted from a published report the authors of a review can contact the writers of the report to obtain the missing information. A review is strengthened if the authors have discussed the implications of missing information and have attempted to collect the relevant data.

Was the assessment of the primary studies reproducible and free from bias?

Expert assessment of primary research studies generally results in a level of disagreement that is both extraordinary and distressing. For example, correlations measuring agreement about the decision to publish or not publish primary research studies are almost always less than 0.5 and average about 0.3,[28,45,46] a level not much higher than one would expect to achieve by chance.

Not only do assessments lack reproducibility, but also they are often biased. In one study Peters and Ceci[47] resubmitted previously published articles from respected institutions after they substituted the names of the authors and the institutions with fictitious names. Mahoney[35] submitted an article to different referees, varying the results without altering the methods. These studies found that the articles that came from respected institutions and reported positive results were more readily accepted. Furthermore, in Peters and Ceci's study many of the articles were rejected because of "serious methodological flaws", and in Mahoney's study the article was judged as having weaker methods when it described negative results.

It is even possible for authors to disagree on the results of a study. Numerous conflicting reviews have been reported in which an author who favoured a particular treatment classified the primary study as positive, whereas an author who did not favour the treatment classified the study as negative. For example, Miller[48] found five reviews that compared drug therapy plus psychotherapy with drug therapy alone for psychiatric patients. Of the 11 studies cited in two or more of the reviews the results of 6 were interpreted as positive in at least one review and as negative in at least one other.

Problems with reproducibility and bias can affect two stages of the review process: the decision about which papers to include and judgement of the quality of the papers included. Such problems can be minimized if explicit criteria are used. However, many of the criteria will require considerable judgement of the author of a review. In an example we used earlier one of the criteria for inclusion in a review of treatment with cholinomimetic agents for Alzheimer's disease was a definition of the population as patients with senile dementia in whom causes other than Alzheimer's disease were excluded. Is a statement in the text such as "standard methods for diagnosing Alzheimer's disease were used" adequate or does one require details of how other causes of dementia were ruled out?

Explicit criteria offer little advantage if they cannot be reproduced by other authors. Ideally, all

the potential primary studies should be assessed for inclusion by at least two authors, each blind to the other's decision, and the extent of agreement should be recorded. Reproducibility should be quantified with a statistical measure that quantitates agreement above and beyond that which would have occurred by chance, such as an intraclass correlation coefficient[49] or a κ statistic.[50] A similar process should be used to assess the reproducibility of the criteria used to determine the validity of the primary studies.

Even if the criteria for study inclusion or validity can be reproduced there is no guarantee that bias has not intruded. For example, if the authors believe that a new treatment works they may apply inclusion criteria by which studies with negative results are systematically excluded; the validity of such studies that are included may be judged more harshly. What can be done to prevent this sort of bias?

In randomized controlled trials bias is avoided if both the patients and the clinicians are blind to whether the patients are taking the active drug or a placebo. In an assessment of primary studies the major possible sources of bias are related to the authors, their institution and the results. However, one can assess the content and quality of a study through its methods without knowing this information; the relevant sections of the paper can simply be "whited out" so that the reviewers are blind to the authors' institutions and results. Decisions about study inclusion and validity ideally should be made under these conditions. This added precaution will strengthen the review.

Was variation in the findings of the relevant studies analysed?

Authors of reviews are certain to encounter variability in the results of studies addressing the question of interest. Indeed, if all the results of primary research were the same a review article would probably not be necessary. It is the authors' task to try to explain this variability.

Possible sources of variability are the study design, chance and differences in the three basic study components (the population, the exposure or intervention and the outcome).[51] If randomized controlled trials, before-and-after studies and studies with historical controls are all included in a review, and if the randomized controlled trials consistently show results that differ systematically from those of the other studies, the study design probably explains the differences. For example, Sacks and colleagues[52] found that randomized controlled trials consistently show smaller effects than studies that use historical controls.

A second explanation for differences in study results is chance. Even if two investigations use comparable methods and the true size of the effects is identical the play of chance will lead to apparent differences in the size. If the samples are small, chance alone may lead to apparently large differences in the size of the effects. Some trials of acetylsalicylic acid (ASA) in patients with transient ischemic attacks have shown a trend in favour of a placebo, whereas others have shown reductions in risk of up to 50% with ASA.[53] However, the confidence intervals, which represent the upper and lower limits of the size of the effects consistent with the observed results, overlap. Thus, although the apparently discrepant results might suggest hypotheses for testing in subsequent studies, they are all consistent with a reduction in risk of between 15% and 30% with ASA.

In other instances differences in study results may be so large that they cannot be explained by chance. The authors must therefore look to differences in the population, exposure or intervention and outcome. In our example of cholinomimetic agents in patients with Alzheimer's disease the studies with negative results may have included a larger number of severely affected patients than the studies with positive results. One might then assume that the intervention works only in mildly affected patients. However, the intervention may have differed — that is, higher doses or different agents may have been given in the studies with positive results. Finally, the tests used to determine memory and functional status may have been different; some tests are more responsive to changes in patient status. Horwitz[51] has documented many ways in which differences in the methods of randomized controlled trials can lead to differing results.

Readers of a review should be alert to whether these five explanations for differing study results have been considered and should be sceptical when differences are attributed to one explanation without adequate consideration of the others.

Were the findings of the primary studies combined appropriately?

Meta-analysis (the use of several statistical techniques to combine the results of different studies) is becoming increasingly popular, especially as a method of combining results from randomized controlled trials. However, it remains controversial, and clinical readers cannot be expected to judge the merits of a particular statistical technique used by the authors of a meta-analysis. Nevertheless, there are issues that clinical readers can address.

The crudest form of meta-analysis, in which the number of studies with positive results is compared with the number of those with negative results, is not satisfactory. This "vote count" ignores the size of the treatment effects and the sample sizes of each study. The most satisfactory meta-analysis yields two pieces of information: the magnitude of the overall treatment effect and the likelihood that this effect would have occurred by chance if the true effect were zero. The former may

be expressed as a percentage risk reduction, the latter as a p value.

The primary advantage of meta-analysis is that the results of different studies can be combined accurately and reliably to determine the best estimate of the average magnitude of the effects of the exposure or intervention of interest. Before the results are combined, however, one should consider whether it is appropriate to aggregate across the studies. Study designs, or the three basic study elements, may differ sufficiently that a statistical combination of the results does not make sense. Meta-analysis can be used to analyse the variation in study results to generate or test hypotheses about the source of the differences. However, it is on strongest ground when the methods of the primary studies are similar and the differences in the study results can be explained by chance.

Reviews in which the results are not statistically combined should state explicitly the basis for the conclusions and should attempt to explain the conflicting results. Readers should beware of reviews that conclude that there is no effect without having considered the studies' power to detect a clinically important effect. When several studies do not show a significant difference there is a tendency for reviewers who have not used meta-analysis to conclude that there is no effect even when statistical aggregation demonstrates otherwise. Cooper and Rosenthal[54] demonstrated this experimentally by assigning reviewers at random to either use or not use meta-analysis to combine the results of several studies, including some that did not show significant results. Another investigator made the same observation when he polled researchers who had conducted trials of tamoxifen citrate as adjuvant therapy for breast cancer (Rory Collins: personal communication, 1987). Most of the researchers concluded from the available information that tamoxifen did not produce a longer disease-free interval; however, statistical aggregation of all the available results demonstrated a clinically important, statistically significant effect.

It is important to remember that all the other guidelines we have discussed still apply whether or not the authors of a review have used meta-analysis.

Table III — Guidelines for assessing the strength of a causal inference

Is the temporal relation correct? (A positive answer is necessary, but it does not, in itself, confer strength on the inference.)
Is the evidence strong?
Is the association strong?
Is there consistency between studies?
Is there a dose–response relation?
Is there indirect evidence that supports the inference — that is, evidence relating to intermediate outcomes, evidence from studies of different populations (including animals) and evidence from analogous relations (i.e., related exposures or interventions)?
Have the plausible competing hypotheses been ruled out?

Were the reviewers' conclusions supported by the data cited?

Whether or not authors have used meta-analysis, the results of individual primary studies should be reported in sufficient detail that readers are able to critically assess the basis for the authors' conclusions. The method of presenting individual study summaries will depend on the question addressed. For questions of treatment effectiveness and prevention the size of the effects and its confidence interval give the key information. Reviews of diagnostic tests may provide sensitivities, specificities and likelihood ratios (and their confidence intervals).[8] Survival curves may efficiently depict the main results of studies of prognosis.

With questions of etiology and causation for which randomized controlled trials are not available the authors can evaluate the evidence with criteria for causal inference. Variations of these criteria have been presented by several investigators,[10,44,55,56] but common ingredients include the size and consistency of the association between the causal agent and the outcome and the necessity for demonstrating the appropriate temporal relation. Our version of these criteria is presented in Table III. The authors' comments on each of these criteria should, of course, refer directly back to the data in the primary studies cited.

Conclusion

A literature review is a scientific endeavour, and, as with other scientific endeavours, standards are available for conducting the review in such a way that valid conclusions are reached. Just as readers of the clinical literature who are unable to critically appraise the methods of primary studies may arrive at incorrect conclusions, readers who are unable to assess the scientific quality of a review are apt to be misled. We have offered eight guidelines for readers interested in answering a clinical question relevant to their everyday practice. Application of these guidelines will allow readers to quickly discard review articles that are irrelevant or scientifically unsound, to detect potential sources of bias and to be confident of conclusions made from a systematic evaluation of the available research.

We thank Drs. Geoff Norman, David Streiner, David L. Sackett and Brian Hutchison, and Professor Mike Gent for their help in developing the guidelines.

This work was supported in part by the Ontario Ministry of Health.

References

1. Haynes RB, McKibbon KA, Fitzgerald D et al: How to keep up with the medical literature: I. Why try to keep up and how to get started. *Ann Intern Med* 1986; 105: 149–153

2. Idem: How to keep up with the medical literature: II. Deciding which journals to read regularly. Ibid: 309-312
3. Idem: How to keep up with the medical literature: III. Expanding the number of journals you read regularly. Ibid: 474-478
4. Idem: How to keep up with the medical literature: IV. Using the literature to solve clinical problems. Ibid: 636-640
5. Idem: How to keep up with the medical literature: V. Access by personal computer to the medical literature. Ibid: 810-824
6. Idem: How to keep up with the medical literature: VI. How to store and retrieve articles worth keeping. Ibid: 978-984
7. Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre: How to read clinical journals: I. Why to read them and how to start reading them critically. *Can Med Assoc J* 1981; 124: 555-558
8. Idem: How to read clinical journals: II. To learn about a diagnostic test. Ibid: 703-710
9. Idem: How to read clinical journals: III. To learn the clinical course and prognosis of disease. Ibid: 869-872
10. Idem: How to read clinical journals: IV. To determine etiology or causation. Ibid: 985-990
11. Idem: How to read clinical journals: V. To distinguish useful from useless or even harmful therapy. Ibid: 1156-1162
12. Idem: How to read clinical journals: VI. To learn about the quality of clinical care. *Can Med Assoc J* 1984; 130: 377-382
13. Labarthe DR: Mild hypertension: the question of treatment. *Ann Rev Public Health* 1986; 7: 193-215
14. Haber E, Slater EE: High blood pressure. In Rubenstein M, Federman DD (eds): *Scientific American Medicine*, 9th ed, Sci Am, New York, 1986: sect 1, VII1-VII29
15. Risks of antihypertensive therapy [E]. *Lancet* 1986; 2: 1075-1076
16. Sacks HS, Chalmers TC, Berk AA et al: Should mild hypertension be treated? An attempted meta-analysis of the clinical trials. *Mt Sinai J Med* 1985; 52: 265-270
17. Guyatt GH, Webber CE, Mewa AA et al: Determining causation — a case study: adrenocorticosteroids and osteoporosis. *J Chronic Dis* 1984; 37: 343-352
18. Baylink DJ: Glucocorticoid-induced osteoporosis. *N Engl J Med* 1983; 309: 306-308
19. Swan HJC, Ganz W: Hemodynamic measurements in clinical practice: a decade in review. *J Am Coll Cardiol* 1983; 1: 103-113
20. Robin ED: The cult of the Swan-Ganz catheter: overuse and abuse of pulmonary flow catheters. *Ann Intern Med* 1985; 103: 445-449
21. Harrington JT, Isner JM, Kassirer JP: Our national obsession with potassium. *Am J Med* 1982; 73: 155-159
22. Kaplan NM: Our appropriate concern about hypokalemia. *Am J Med* 1984; 77: 1-4
23. Mulrow CD: The medical review article: state of the science. *Ann Intern Med* 1987; 106: 485-488
24. Sacks HS, Berrier J, Reitman D et al: Meta-analyses of randomized controlled trials. *N Engl J Med* 1987; 316: 450-455
25. Williamson JW, Goldschmidt PG, Colton T: The quality of medical literature. An analysis of validation assessments. In Bailar JC III, Mosteller F (eds): *Medical Uses of Statistics*, NEJM Bks, Waltham, Mass, 1986: 370-391
26. Halvorsen KT: Combining results from independent investigations: meta-analysis in medical research. Ibid: 392-416
27. Oxman AD: *A Methodological Framework for Research Overviews*, MSc thesis, McMaster U, Hamilton, Ont, 1987: 23-31, 98-105
28. Light RJ, Pillemer DB: *Summing Up: the Science of Reviewing Research*, Harvard U Pr, Cambridge, Mass, 1984
29. Jackson GB: Methods for integrative reviews. *Rev Educ Res* 1980; 50: 438-460
30. Cooper HM: *The Integrative Research Review: a Systematic Approach*, Sage, Beverly Hills, Calif, 1984
31. Glass GV, McGaw B, Smith ML: *Meta-Analysis in Social Research*, Sage, Beverly Hills, Calif, 1981
32. Poynard T, Conn HO: The retrieval of randomized clinical trials in liver disease from the medical literature. *Controlled Clin Trials* 1985; 6: 271-279
33. Dickersin K, Hewitt P, Mutch L et al: Perusing the literature: comparison of MEDLINE searching with a perinatal trials database. Ibid: 306-317
34. Simes RJ: Publication bias. The case for an international registry of clinical trials. *J Clin Oncol* 1986; 4: 1529-1541
35. Mahoney MJ: Publication prejudices: an experimental study of confirmatory bias in the peer review system. *Cognit Ther Res* 1977; 1: 161-175
36. Devine EC, Cook TD: Effects of psycho-educational intervention on length of hospital stay: a meta-analytic review of 34 studies. In Light RJ (ed): *Evaluation Studies Review Annual*, 8th ed, Sage, Beverly Hills, Calif, 1983: 417-432
37. Simes RJ: Confronting publication bias: a cohort design for meta-analysis. *Stat Med* 1987; 6: 11-29
38. Cooper HM: Literature searching strategies of integrative research reviewers: a first survey. *Knowledge* 1986; 8: 372-383
39. Huth EJ: Needed: review articles with more scientific rigor. *Ann Intern Med* 1987; 106: 470-471
40. Messer J, Reitman D, Sacks HS et al: Association of adrenocorticosteroid therapy and peptic-ulcer disease. *N Engl J Med* 1983; 309: 21-24
41. Conn HO, Blitzer BL: Nonassociation of adrenocorticosteroid therapy and peptic ulcer. *N Engl J Med* 1976; 294: 473-479
42. EC/IC Bypass Study Group: Failure of extracranial-intracranial arterial bypass to reduce the risk of ischemic stroke. Results of an international randomized trial. *N Engl J Med* 1985; 313: 1191-1200
43. Hill AB: *Principles of Medical Statistics*, 9th ed, *Lancet*, London, 1971: 312-320
44. Chalmers TC, Smith H, Blackburn B et al: A method for assessing the quality of a randomized controlled trial. *Controlled Clin Trials* 1981; 2: 31-49
45. Bailar JC III, Patterson K: Journal peer review: the need for a research agenda. In Bailar JC III, Mosteller F (eds): *Medical Uses of Statistics*, NEJM Bks, Waltham, Mass, 1986: 349-369
46. Marsh HW, Ball S: Interjudgmental reliability of reviews for the *Journal of Educational Psychology*. *J Educ Psychol* 1981; 73: 872-880
47. Peters DP, Ceci SJ: Peer-review practices of psychological journals: the fate of published articles, submitted again. *Behav Brain Sci* 1982; 5: 187-255
48. Miller TI: *The Effects of Drug Therapy on Psychological Disorders*, PhD dissertation, U of Colorado, Boulder, 1977
49. Shrout PE, Fleiss JL: Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; 86: 420-428
50. Cohen J: A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960; 20: 37-46
51. Horwitz RI: Complexity and contradiction in clinical trial research. *Am J Med* 1987; 82: 498-510
52. Sacks H, Chalmers TC, Smith H: Randomized versus historical controls for clinical trials. *Am J Med* 1982; 72: 233-240
53. Sze PC, Pincus M, Sacks HS et al: Antiplatelet agents in secondary stroke prevention. A meta-analysis of the available randomized control trials [abstr]. *Clin Res* 1986; 34: 385A
54. Cooper HM, Rosenthal R: Statistical versus traditional procedures for summarizing research findings. *Psychol Bull* 1980; 87: 442-449
55. Susser M: Reviews and commentary: the logic of Sir Karl Popper and the practice of epidemiology. *Am J Epidemiol* 1986; 124: 711-718
56. Guyatt GH, Newhouse MT: Are active and passive smoking harmful? Determining causation. *Chest* 1985; 88: 445-451

Source: Glaser BG, Strauss AL (1967) The discovery of grounded theory: Strategies for qualitative research. Aldine; New York.

Selecting Comparison Groups

In this section we focus on two questions: which groups are selected, why and how?

Which Groups?

The basic criterion governing the selection of comparison groups for discovering theory is their *theoretical relevance* for furthering the development of emerging categories. The researcher chooses any groups that will help generate, to the fullest extent, as many properties of the categories as possible, and that will help relate categories to each other and to their properties. Thus, as we said in Chapter II, group comparisons are conceptual; they are made by comparing diverse or similar evidence indicating the same conceptual categories and properties, *not* by comparing the evidence for its own sake. Comparative analysis takes full advantage of the "interchangeability" of indicators, and develops, as it proceeds, a broad range of acceptable indicators for categories and properties.[6]

Since groups may be chosen for a single comparison only, there can be no definite, prescribed, preplanned set of groups that are compared for all or even most categories (as there are

5. For example, "The entire design of the study did not permit me to propose hypotheses . . . it simply permitted me to describe what I found," Stanley H. Udy, Jr., "Cross Cultural Analysis: A Case Study," Hammond, *op. cit.*, p. 173, and *passim* for more examples. Merton has developed a research design for interweaving the standard procedures of preplanned data collection and data analysis in order to keep adjusting to discovered relevances. For a synopsis see Hanan C. Selvin, "The Interplay of Social Research and Social Policy in Housing," *Journal of Social Issues*, Vol. VII, (1951), pp. 180-81.

6. Paul F. Lazarsfeld and Wagner Theileus, Jr., *Academic Mind* (New York: Free Press of Glencoe, 1958), pp. 402-08.

in comparative studies made for accurate descriptions and verification). In research carried out for discovering theory, the sociologist cannot cite the number and types of groups from which he collected data *until* the research is completed. In an extreme case, he may then find that the development of each major category may have been based on comparisons of different sets of groups. For example, one could write a substantive theory about scientists' authority in organizations, and compare very different kinds of organizations to develop properties associated with the diverse categories that might emerge: authority over clients, administration, research facilities, or relations with outside organizations and communities; the degree or type of affiliation in the organization; and so forth. Or the sociologist may wish to write a formal theory about professional authority in organizations; then the sets of comparison groups for each category are likely to be much more diverse than those used in developing a substantive theory about scientists, since now the field of possible comparison is far greater.

Our logic of *ongoing inclusion* of groups must be differentiated from the logic used in comparative analyses that are focused mainly on accurate evidence for description and verification. That logic, one of preplanned inclusion and exclusion, warns the analyst away from comparing "non-comparable" groups. To be included in the planned set, a group must have "enough features in common" with the other groups. To be excluded, it must show a "fundamental difference" from the others.[7] These two rules represent an attempt to "hold constant" strategic facts, or to disqualify groups where the facts either cannot actually be held constant or would introduce more unwanted differences. Thus in comparing variables (conceptual and factual), one hopes that, because of this set of "purified groups," spurious factors now will not influence the findings and relationships and render them inaccurate. This effort of purification is made for a result impossible to achieve, since one never really knows what has and has not been held constant.

7. For example see Janowitz, *op. cit.*, Preface and Chapter 1; and Edward A. Shils, "On the Comparative Study of New States" in Clifford Geertz (Ed.), *Old Societies and New States* (New York: Free Press of Glencoe, 1963), pp. 5, 9.

To be sure, these rules of comparability are important when accurate evidence is the goal, but they hinder the generation of theory, in which "non-comparability" of groups is irrelevant. They prevent the use of a much wider range of groups for developing properties of categories. Such a range, necessary for the categories' fullest possible development, is achieved by comparing *any* groups, irrespective of differences or similarities, as long as the data apply to a similar category or property. Furthermore, these two rules divert the analyst's attention away from the important sets of fundamental differences and similarities, which, upon analysis, become important qualifying conditions under which categories and properties vary. These differences should be made a vital part of the analysis, but rules of comparability tend to make the analyst inattentive to conditions that vary findings by allowing him to assume constants and to disqualify basic differences, thus nullifying their effort before the analysis.

It is theoretically important to note to what degree the properties of categories are varied by diverse conditions. For example, properties of the effect of awareness contexts on the interaction between the nurse and the dying patient within a hospital can usefully be developed by making comparisons with the same situation in the home, in nursing homes, in ambulances, and on the street after accidents. The similarities and differences in these conditions can be used to explain the similar and diverse properties of interaction between nurse and patient.

The principal point to keep clear is the purpose of the research, so that rules of evidence will not hinder discovery of theory. However, these goals are usually not kept clear (a condition we are trying to correct) and so typically a sociologist starts by applying these rules for selecting a purified set of groups to achieve accurate evidence. He then becomes caught up in the delights of generating theory, and so compares everything comparable; but next he finds his theory development severely limited by lack of enough theoretically relevant data, because he has used a preplanned set of groups for collecting his information (see Chapter VI). In allowing freedom for comparing any groups, the criterion of theoretical relevance

used for each comparison in systematically generating theory controls data collection without hindering it. Control by this criterion assures that ample data will be collected and that the data collection makes sense (otherwise collection is a waste of time). However, applying theoretical control over choice of comparison groups is more difficult than simply collecting data from a preplanned set of groups, since choice requires continuous thought, analysis and search.

The sociologist must also be clear on the basic *types* of groups he wishes to compare in order to control their effect on generality of both *scope* of population and *conceptual level* of his theory. The simplest comparisons are, of course, made among different groups of exactly the same substantive type; for instance, federal bookkeeping departments. These comparisons lead to a substantive theory that is applicable to this one type of group. Somewhat more general substantive theory is achieved by comparing different types of groups; for example, different kinds of federal departments in one federal agency. The scope of the theory is further increased by comparing different types of groups within different larger groups (different departments in different agencies). Generality is further increased by making these latter comparisons for different regions of a nation or, to go further, different nations. *The scope of a substantive theory can be carefully increased and controlled by such conscious choices of groups.* The sociologist may also find it convenient to think of subgroups within larger groups, and of internal and external groups, as he broadens his range of comparisons and attempts to keep tractable his substantive theory's various levels of generality of scope.

The sociologist developing substantive or formal theory can also usefully *create* groups, provided he keeps in mind that they are an artifact of his research design, and so does not start assuming in his analysis that they have properties possessed by a natural group. Survey researchers are adept at creating groups and statistically grounding their relevance (as by factor analysis, scaling, or criteria variables) to make sure they are, in fact, groups that make meaningful differences even though they have been created: for example, teachers high, medium, and low on "apprehension"; or upper, middle, and lower class;

or local-cosmopolitan.[8] However, only a handful of survey researchers have used their skill to create multiple comparison subgroups for discovering theory. This would be a very worthwhile endeavor (see Chapter VIII on quantitative data).

The tactic of creating groups is equally applicable for sociologists who work with qualitative data. When using only interviews, for instance, a researcher surely can study comparison groups composed of respondents chosen in accordance with his emergent analytic framework. And historical documents, or other library materials, lend themselves wonderfully to the comparative method. Their use is perhaps even more efficient, since the researcher is saved much time and trouble in his search for comparison groups which are, after all, already concentrated in the library (see Chapter VII). As in field work, the researcher who uses library material can always select additional comparison groups after his analytic framework is well developed, in order to give himself additional confidence in its credibility. He will also—like the field worker who sometimes stumbles upon comparison groups and then makes proper use of them—occasionally profit from happy accidents that may occur when he is browsing along library shelves. And, again like the researcher who carefully chooses natural groups, the sociologist who creates groups should do so carefully according to the scales of generality that he desires to achieve.

As the sociologist shifts the degree of conceptual generality for which he aims, from discovering substantive to discovering formal theory, he must keep in mind the *class* of the groups he selects. For substantive theory, he can select, as the same substantive class, groups regardless of where he finds them. He may, thus, compare the "emergency ward" to all kinds of medical wards in all kinds of hospitals, both in the United States and abroad. But he may also conceive of the emergency ward as a subclass of a larger class of organizations, all designed to render immediate assistance in the event of accidents or break-

8. In fact, in backstage discussions about which comparative groups to create and choose in survey analysis, the answer frequently is: "Where the breaks in the distribution are convenient and save cases, and among these choose the ones that give the 'best findings.'" Selvin, however, has developed a systematic method of subgroup comparison in survey research that prevents the opportunistic use of "the best finding" criteria. See *The Effects of Leadership* (Glencoe, Ill.: Free Press, 1960).

downs. For example, fire, crime, the automobile, and even plumbing problems have all given rise to emergency organizations that are on 24-hour alert. In taking this approach to choosing dissimilar, substantive comparative groups, the analyst must be clear about his purpose. He may use groups of the more general class to illuminate his substantive theory of, say, emergency wards. He may wish to begin generating a formal theory of emergency organizations. He may desire a mixture of both: for instance, bringing out his substantive theory about emergency wards within a context of some formal categories about emergency organizations.[9]

On the other hand, when the sociologist's purpose is to discover formal theory, he will definitely select dissimilar, substantive groups from the larger class, while increasing his theory's scope. And he will also find himself comparing groups that seem to be non-comparable on the substantive level, but that on the formal level are conceptually comparable. Non-comparable on the substantive level here implies a stronger degree of apparent difference than does *dissimilar*. For example, while fire departments and emergency wards are substantially dissimilar, their conceptual comparability is still readily apparent. Since the basis of comparison between substantively non-comparable groups is not readily apparent, it must be explained on a higher conceptual level.

Thus, one could start developing a formal theory of social isolation by comparing four apparently unconnected monographs: *Blue Collar Marriage*, *The Taxi-Dance Hall*, *The Ghetto* and *The Hobo* (Komarovsky, Cressey, Wirth, Anderson).[10] All deal with facets of "social isolation," according to their authors. For another example, Goffman has compared apparently non-comparable groups when generating his formal theory of stigma. Thus, anyone who wishes to discover formal theory should be aware of the usefulness of comparisons made on high level conceptual categories among the seemingly non-comparable; he should actively seek this kind of comparison; do it with flexibility; and be able to interchange the apparently

9. Cf. Shils, *op. cit.*, p. 17.

10. Respectively, Mirra Komarovsky (New York: Random House, 1962); Paul Cressey (Chicago: University of Chicago Press, 1932); Louis Wirth (Chicago: University of Chicago Press, 1962 edition); and Nels Anderson (Chicago: University of Chicago Press, 1961 edition).

non-comparable comparison with the apparently comparable ones. The non-comparable type of group comparison can greatly aid him in transcending substantive descriptions of time and place as he tries to achieve a general, formal theory.[11]

Source: Glaser B.G., Strauss A.L. (1967) The discovery of grounded theory Strategies for qualitative research. Aldine; New York.

V

The Constant Comparative Method of Qualitative Analysis*

Currently, the general approaches to the analysis of qualitative data are these:

1. If the analyst wishes to convert qualitative data into crudely quantifiable form so that he can provisionally test a hypothesis, he codes the data first and then analyzes it. He makes an effort to code "all relevant data [that] can be brought to bear on a point," and then systematically assembles, assesses and analyzes these data in a fashion that will "constitute proof for a given proposition."[1]

2. If the analyst wishes only to generate theoretical ideas—new categories and their properties, hypotheses and interrelated hypotheses—he cannot be confined to the practice of coding first and then analyzing the data since, in generating theory, he is constantly redesigning and reintegrating his theoretical notions as he reviews his material.[2] Analysis after the coding operation

* We wish to thank the editors of *Social Problems* for permission to publish this paper as Chapter V. See Barney G. Glaser, *Social Problems*, 12 (1965), pp. 436-45.

1. Howard S. Becker and Blanche Geer, "The Analysis of Qualitative Field Data" in Richard N. Adams and Jack J. Preiss (Eds.), *Human Organization Research* (Homewood, Ill.: Dorsey Press, Inc., 1960), pp. 279-89. See also Howard S. Becker, "Problems of Inference and Proof in Participant Observation," *American Sociological Review*, (December, 1958), pp. 652-60; and Bernard Berelson, *Content Analysis* (Glencoe, Ill.: Free Press, 1952), Chapter III, and p. 16.

2. Constantly redesigning the analysis is a well-known normal tendency in qualitative research (no matter what the approach to analysis), which occurs throughout the whole research experience from initial data collec-

would not only unnecessarily delay and interfere with his purpose, but the explicit coding itself often seems an unnecessary, burdensome task. As a result, the analyst merely inspects his data for new properties of his theoretical categories, and writes memos on these properties.

We wish to suggest a third approach to the analysis of qualitative data—one that combines, by an analytic procedure of constant comparison, the explicit coding procedure of the first approach and the style of theory development of the second. The purpose of the constant comparative method of joint coding and analysis is to generate theory more systematically than allowed by the second approach, *by using explicit coding and analytic procedures*. While more systematic than the second approach, this method does not adhere completely to the first, which hinders the development of theory because it is designed for provisional testing, not discovering, of hypotheses.[3] This method of comparative analysis is to be used jointly with theoretical sampling, whether for collective new data or on previously collected or compiled qualitative data.

Systematizing the second approach (inspecting data and

tion through coding to final analysis and writing. The tendency has been noted in Becker and Geer, *op. cit.*, p. 270, Berelson, *op. cit.*, p. 125; and for an excellent example of how it goes on, see Robert K. Merton, *Social Theory and Social Structure* (New York: Free Press of Glencoe, 1957), pp. 390-92. However, this tendency may have to be suppressed in favor of the purpose of the first approach; but in the second approach and the approach presented here, the tendency is used purposefully as an analytic strategy.

3. Our other purpose in presenting the constant comparative method may be indicated by a direct quotation from Robert K. Merton—a statement he made in connection with his own qualitative analysis of locals and cosmopolitans as community influentials: "This part of our report, then, is a bid to the sociological fraternity for the practice of incorporating in publications a detailed account of the ways in which qualitative analyses *actually* developed. Only when a considerable body of such reports are available will it be possible to *codify* methods of qualitative analysis with something of the clarity with which quantitative methods have been articulated." *Op. cit.*, p. 390. This is, of course, also the basic position of Paul F. Lazarsfeld. See Allen H. Barton and Paul F. Lazarsfeld, "Some Functions of Qualitative Analysis in Social Research," in Seymour M. Lipset and Neil J. Smelser (Eds.), *Sociology: the Progress of a Decade* (Englewood Cliffs, N.J.: Prentice-Hall, 1961). It is the position that has stimulated the work of Becker and Geer, and of Berelson, cited in Footnote 1.

redesigning a developing theory) by this method does not supplant the skills and sensitivities required in generating theory. Rather, the constant comparative method is designed to aid the analyst who possesses these abilities in generating a theory that is integrated, consistent, plausible, close to the data—and at the same time is in a form clear enough to be readily, if only partially, operationalized for testing in quantitative research. Still dependent on the skills and sensitivities of the analyst, the constant comparative method is not designed (as methods of quantitative analysis are) to guarantee that two analysts working independently with the same data will achieve the same results; it is designed to allow, with discipline, for some of the vagueness and flexibility that aid the creative generation of theory.

If a researcher using the first approach (coding all data first) wishes to discover some or all of the hypotheses to be tested, typically he makes his discoveries by using the second approach of inspection and memo-writing along with explicit coding. By contrast, the constant comparative method cannot be used for both provisional testing and discovering theory: in theoretical sampling, the data collected are not extensive enough and, because of theoretical saturation, are not coded extensively enough to yield provisional tests, as they are in the first approach. They are coded only enough to generate, hence to suggest, theory. Partial testing of theory, when necessary, is left to more rigorous approaches (sometimes qualitative but usually quantitative). These come later in the scientific enterprise (see Chapter X).

The first approach also differs in another way from the constant comparative method. It is usually concerned with a few hypotheses couched at the same level of generality, while our method is concerned with many hypotheses synthesized at different levels of generality. The reason for this difference between methods is that the first approach must keep the theory tractable so that it can be provisionally tested in the same presentation. Of course, the analyst using this approach might, after proving or disproving his hypotheses, attempt to explain his findings with more general ideas suggested by his data, thus achieving some synthesis at different levels of generality.

A fourth general approach to qualitative analysis is "analytic

induction," which combines the first and second approaches in a manner different from the constant comparative method.[4] Analytic induction has been concerned with generating and proving an integrated, limited, precise, universally applicable theory of causes accounting for a specific behavior (*e.g.*, drug addiction, embezzlement). In line with the first approach, it tests a limited number of hypotheses with *all* available data, consisting of numbers of clearly defined and carefully selected cases of the phenomena. Following the second approach, the theory is generated by the reformulation of hypotheses and redefinition of the phenomena forced by constantly confronting the theory with negative cases, cases which do not confirm the current formulation.

In contrast to analytic induction, the constant comparative method is concerned with generating and plausibly suggesting (but not provisionally testing) many categories, properties, and hypotheses about general problems (*e.g.*, the distribution of services according to the social value of clients). Some of these properties may be causes, as in analytic induction, but unlike analytic induction others are conditions, consequences, dimensions, types, processes, etc. In both approaches, these properties should result in an integrated theory. Further, no attempt is made by the constant comparative method to ascertain either the universality or the proof of suggested causes or other properties. Since no proof is involved, the constant comparative method in contrast to analytic induction requires only saturation of data—not consideration of *all* available data, nor are the data restricted to one kind of clearly defined case. The constant comparative method, unlike analytic induction, is more likely to be applied in the same study to any kind of qualitative information, including observations, interviews, documents, articles, books, and so forth. As a consequence, the constant comparisons required by both methods differ in breadth of purpose, extent of comparing, and what data and ideas are compared.

Clearly the purposes of both these methods for generating theory supplement each other, as well as the first and second

4. See Alfred R. Lindesmith, *Opiate Addiction* (Bloomington: Principia, 1947), pp. 12-14; Donald R. Cressey, *Other People's Money* (New York: Free Press of Glencoe, 1953), p. 16 and *passim;* and Florian Znaniecki, *The Method of Sociology* (New York: Farrar and Rinehart, 1934), pp. 249-331.

approaches. All four methods provide different alternatives to qualitative analysis. Table I locates the use of these approaches to qualitative analysis and provides a scheme for locating additional approaches according to their purposes. The general idea of the constant comparative method can also be used for generating theory in quantitative research. Then one compares findings within subgroups and with external groups (see Chapter VIII).

TABLE I. USE OF APPROACHES TO QUALITATIVE ANALYSIS

Generating Theory	*Provisional Testing of Theory*	
	Yes	No
Yes	Combining inspection for hypotheses (2) along with coding for test, then analyzing data (1)	Inspection for hypotheses (2)
	Analytic induction (4)	Constant comparative method (3)
No	Coding for test, then analyzing data (1)	Ethnographic description

The Constant Comparative Method

We shall describe in four stages the constant comparative method: (1) comparing incidents applicable to each category, (2) integrating categories and their properties, (3) delimiting the theory, and (4) writing the theory. Although this method of generating theory is a continuously growing process —each stage after a time is transformed into the next—earlier stages do remain in operation simultaneously throughout the analysis and each provides continuous development to its successive stage until the analysis is terminated.

1. *Comparing incidents applicable to each category.* The analyst starts by coding each incident in his data into as many categories of analysis as possible, as categories emerge or as data emerge that fit an existing category. For example, the category of "social loss" of dying patients emerged quickly from comparisons of nurses' responses to the potential deaths of their patients. Each relevant response involved the nurse's appraisal

of the degree of loss that her patient would be to his family, his occupation, or society: "He was so young," "He was to be a doctor," "She had a full life," or "What will the children and her husband do without her?" [5]

Coding need consist only of noting categories on margins, but can be done more elaborately (*e.g.*, on cards). It should keep track of the comparison group in which the incident occurs. To this procedure we add the basic, defining rule for the constant comparative method: *while coding an incident for a category, compare it with the previous incidents in the same and different groups coded in the same category.* For example, as the analyst codes an incident in which a nurse responds to the potential "social loss" of a dying patient, he also compares this incident, before further coding, with others previously coded in the same category. Since coding qualitative data requires study of each incident, this comparison can often be based on memory. Usually there is no need to refer to the actual note on every previous incident for each comparison.

This constant comparison of the incidents very soon starts to generate theoretical properties of the category. The analyst starts thinking in terms of the full range of types or continua of the category, its dimensions, the conditions under which it is pronounced or minimized, its major consequences, its relation to other categories, and its other properties. For example, while constantly comparing incidents on how nurses respond to the social loss of dying patients, we realized that some patients are perceived as a high social loss and others as a low social loss, and that patient care tends to vary positively with degree of social loss. It was also apparent that some social attributes that nurses combine to establish a degree of social loss are seen immediately (age, ethnic group, social class), while some are learned after time is spent with the patient (occupational worth, marital, status, education). This observation led us to the realization that perceived social loss can change as new attributes of the patients are learned. It also became apparent, from studying the comparison groups, under what conditions (types of wards and hospitals) we would find clusters of patients with different degrees of social loss.

5. Illustrations will refer to Barney G. Glaser and Anselm L. Strauss, "The Social Loss of Dying Patients," *American Journal of Nursing*, 64 (June, 1964), pp. 119-121.

As categories and their properties emerge, the analyst will discover two kinds: those that he has constructed himself (such as "social loss" or 'calculation" of social loss); and those that have been abstracted from the language of the research situation. (For example, "composure" was derived from nurses' statements like "I was afraid of losing my composure when the family started crying over their child.") As his theory develops, the analyst will notice that the concepts abstracted from the substantive situation will tend to be current labels in use for the actual processes and behaviors that are to be explained, while the concepts constructed by the analyst will tend to be the explanations.[6] For example, a nurse's perception of the social loss of a dying patient will affect (an explanation) how she maintains her composure (a behavior) in his presence.

After coding for a category perhaps three or four times, the analyst will find conflicts in the emphases of his thinking. He will be musing over theoretical notions and, at the same time, trying to concentrate on his study of the next incident, to determine the alternate ways by which it should be coded and compared. At this point, the second rule of the constant comparative method is: *stop coding and record a memo on your ideas.* This rule is designed to tap the initial freshness of the analyst's theoretical notions and to relieve the conflict in his thoughts. In doing so, the analyst should take as much time as necessary to reflect and carry his thinking to its most logical (grounded in the data, not speculative) conclusions. It is important to emphasize that for joint coding and analysis there can be no scheduled routine covering the amount to be coded per day, as there is in predesigned research. The analyst may spend hours on one page or he may code twenty pages in a half hour, depending on the relevance of the material, saturation of categories, emergence of new categories, stage of formulation of theory, and of course the mood of the analyst, since this method takes his personal sensitivity into consideration. These factors are in a continual process of change.

If one is working on a research team, it is also a good idea to discuss theoretical notions with one or more teammates. Teammates can help bring out points missed, add points they

6. Thus we have studies of delinquency, justice, "becoming," stigma, consultation, consolation, contraception, etc.; these usually become the variables or processes to be described and explained.

have run across in their own coding and data collection, and crosscheck his points. They, too, begin to compare the analyst's notions with their own ideas and knowledge of the data; this comparison generates additional theoretical ideas. With clearer ideas on the emerging theory systematically recorded, the analyst then returns to the data for more coding and constant comparison.

From the point of view of generating theory it is often useful to write memos on, as well as code, the copy of one's field notes. Memo writing on the field note provides an immediate illustration for an idea. Also, since an incident can be coded for several categories, this tactic forces the analyst to use an incident as an illustration only once, for the most important among the many properties of diverse categories that it indicates. He must look elsewhere in his notes for illustrations for his other properties and categories. This corrects the tendency to use the same illustration over and over for different properties.

The generation of theory requires that the analyst take apart the story within his data. Therefore when he rearranges his memos and field notes for writing up his theory, he sufficiently "fractures" his story at the same time that he saves apt illustrations for each idea (see Step 4). At just this point in his writing, breaking down and out of the story is necessary for clear integration of the theory.

2. *Integrating categories and their properties.* This process starts out in a small way; memos and possible conferences are short. But as the coding continues, the constant comparative units change from comparison of incident with incident to comparison of incident with properties of the category that resulted from initial comparisons of incidents. For example, in comparing incident with incident we discovered the property that nurses constantly recalculate a patient's social loss as they learn more about him. From then on, each incident bearing on "calculation" was compared with "accumulated knowledge on calculating"—not with all other incidents involving calculation. Thus, once we found that age was the most important characteristic in calculating social loss, we could discern how a patient's age affected the nurses' recalculation of social loss as they found out more about his education. We found that education was most influential in calculations of the social loss of a middle-aged

adult, since for a person of this age, education was considered to be of most social worth. This example also shows that constant comparison causes the accumulated knowledge pertaining to a property of the category to readily start to become integrated; that is, related in many different ways, resulting in a unified whole.

In addition, the diverse properties themselves start to become integrated. Thus, we soon found that the calculating and recalculating of social loss by nurses was related to their development of a social loss "story" about the patient. When asked about a dying patient, nurses would tell what amounted to a story about him. The ingredients of this story consisted of a continual balancing out of social loss factors as the nurses learned more about the patient. Both the calculus of social loss and the social loss story were related to the nurse's strategies for coping with the upsetting impact on her professional composure of, say, a dying patient with a high social loss (*e.g.*, a mother with two children). This example further shows that the category becomes integrated with other categories of analysis: the social loss of the dying patient is related to how nurses maintain professonal composure while attending his dying.[7] Thus the theory develops, as different categories and their properties tend to become integrated through constant comparisons that force the analyst to make some related theoretical sense of each comparison.

If the data are collected by theoretical sampling at the same time that they are analyzed (as we suggest should be done), then integration of the theory is more likely to emerge by itself. By joint collection and analysis, the sociologist is tapping to the fullest extent the in vivo patterns of integration in the data itself; questions guide the collection of data to fill in gaps and to extend the theory—and this also is an integrative strategy. Emergence of integration schemes also occurs in analyses that are separate from data collection, but more contrivance may be necessary when the data run thin and no more can be collected. (Other aspects of integration have been discussed in Chapter II.)

3. *Delimiting the theory.* As the theory develops, various

7. See Glaser and Strauss, "Awareness and the Nurse's Composure," in Chapter 13 in *Awareness of Dying* (Chicago: Aldine Publishing Co., 1965).

delimiting features of the constant comparative method begin to curb what could otherwise become an overwhelming task. Delimiting occurs at two levels: the theory and the categories. First, the theory solidifies, in the sense that major modifications become fewer and fewer as the analyst compares the next incidents of a category to its properties. Later modifications are mainly on the order of clarifying the logic, taking out nonrelevant properties, integrating elaborating details of properties into the major outline of interrelated categories and—most important—reduction.

By reduction we mean that the analyst may discover underlying uniformities in the original set of categories or their properties, and can then formulate the theory with a smaller set of higher level concepts. This delimits its terminology and text. Here is an illustration which shows the integration of more details into the theory and some consequent reduction: We decided to elaborate our theory by adding detailed strategies used by the nurses to maintain professional composure while taking care of patients with varying degrees of social loss. We discovered that the rationales which nurses used, when talking among themselves, could all be considered "loss rationales." The underlying uniformity was that all these rationales indicated why the patient, given his degree of social loss, would, if he lived, now be socially worthless; in spite of the social loss, he would be better off dead. For example, he would have brain damage, or be in constant, unendurable pain, or have no chance for a normal life.

Through further reduction of terminology we were also discovering that our theory could be generalized so that it pertained to the care of all patients (not just dying ones) by all staff (not just nurses). On the level of formal theory, it could even be generalized as a theory of how the social values of professionals affect the distribution of their services to clients; for example, how they decide who among many waiting clients should next receive a service, and what calibre of service he should be given.

Thus, with reduction of terminology and consequent generalizing, forced by constant comparisons (some comparisons can at this point be based on the literature of other professional areas), the analyst starts to achieve two major requirements of

theory: (1) *parsimony* of variables and formulation, and (2) *scope* in the applicability of the theory to a wide range of situations,[8] while keeping a close correspondence of theory and data.

The second level for delimiting the theory is a reduction in the original list of categories for coding. As the theory grows, becomes reduced, and increasingly works better for ordering a mass of qualitative data, the analyst becomes committed to it. His commitment now allows him to cut down the original list of categories for collecting and coding data, according to the present boundaries of his theory. In turn, his consideration, coding, and analyzing of incidents can become more select and focused. He can devote more time to the constant comparison of incidents clearly applicable to this smaller set of categories.

Another factor, which still further delimits the list of categories, is that they become *theoretically saturated*. After an analyst has coded incidents for the same category a number of times, he learns to see quickly whether or not the next applicable incident points to a new aspect. If yes, then the incident is coded and compared. If no, the incident is not coded, since it only adds bulk to the coded data and nothing to the theory.[9] For example, after we had established age as the base line for calculating social loss, no longer did we need to code incidents referring to age for calculating social loss. However, if we came across a case where age did not appear to be the base line (a negative case), the case was coded and then compared. In the case of an 85-year-old dying woman who was considered a great social loss, we discovered that her "wonderful personality" outweighed her age as the most important factor for calculating her social loss. In addition, the amount of data the analyst needs to code is considerably reduced when the data are obtained by theoretical sampling; thus he saves time in studying his data for coding.

8. Merton, *op. cit.*, p. 260.

9. If the analyst's purpose, besides developing theory, is also to count incidents for a category to establish provisional proofs, then he must code the incident. Furthermore, Merton has made the additional point, in correspondence, that to count for establishing provisional proofs may also feed back to developing the theory, since frequency and cross-tabulation of frequencies can also generate new theoretical ideas. See Berelson on the conditions under which one can justify time-consuming, careful counting; *op. cit.*, pp. 128-34. See Becker and Geer for a new method of counting the frequency of incidents; *op. cit.*, pp. 283-87.

Theoretical saturation of categories also can be employed as a strategy in coping with another problem: new categories will emerge after hundreds of pages of coding, and the question is whether or not to go back and re-code all previously coded pages. The answer for large studies is "no." The analyst should start to code for the new category where it emerges, and continue for a few hundred pages of coding, or until the remaining (or additionally collected) data have been coded, to see whether the new category has become theoretically saturated. If it has, then it is unnecessary to go back, either to the field or the notes, because theoretical saturation suggests that what has been missed will probably have little modifying effect on the theory. If the category does not saturate, then the analyst needs to go back and try to saturate it, provided it is central to the theory.

Theoretical saturation can help solve still another problem concerning categories. If the analyst has collected his own data, then from time to time he will remember other incidents that he observed or heard but did not record. What does he do now? If the unrecorded incident applies to an established category, after comparison it can either be ignored because the category is saturated; or, if it indicates a new property of the category, it can be added to the next memo and thus integrated into the theory. If the remembered incident generates a new category, both incident and category can be included in a memo directed toward their place in the theory. This incident alone may be enough data if the category is minor. However, if it becomes central to the theory, the memo becomes a directive for further coding of the field notes, and for returning to the field or library to collect more data.

The universe of data that the constant comparative method uses is based on the reduction of the theory and the delimitation and saturation of categories. Thus, the collected universe of data is first delimitated and then, if necessary, carefully extended by a return to data collection according to the requirements of theoretical sampling. Research resources are economized by this theoretical delimiting of the possible universe of data, since working within limits forces the analyst to spend his time and effort only on data relevant to his categories. In large field studies, with long lists of possibly useful categories

and thousands of pages of notes embodying thousands of incidents, each of which could be coded a multitude of ways, theoretical criteria are very necessary for paring down an otherwise monstrous task to fit the available resources of personnel, time, and money. Without theoretical criteria, delimiting a universe of collected data, if done at all, can become very arbitrary and less likely to yield an integrated product; the analyst is also more likely to waste time on what may later prove to be irrelevant incidents and categories.

4. *Writing theory.* At this stage in the process of qualitative analysis, the analyst possesses coded data, a series of memos, and a theory. The discussions in his memos provide the content behind the categories, which become the major themes of the theory later presented in papers or books. For example, the major themes (section titles) for our paper on social loss were "calculating social loss," "the patient's social loss story," and "the impact of social loss on the nurse's professional composure."

When the researcher is convinced that his analytic framework forms a systematic substantive theory, that it is a reasonably accurate statement of the matters studied, and that it is couched in a form that others going into the same field could use—then he can publish his results with confidence. To start writing one's theory, it is first necessary to collate the memos on each category, which is easily accomplished since the memos have been written about categories. Thus, we brought together all memos on calculating social loss for summarizing and, perhaps, further analyzing before writing about it. One can return to the coded data when necessary to validate a suggested point, pinpoint data behind a hypothesis or gaps in the theory, and provide illustrations.[10]

Properties of the Theory

Using the constant comparative method makes probable the achievement of a complex theory that corresponds closely to

10. On "pinpointing" see Anselm Strauss, Leonard Schatzman, Rue Bucher, Danuta Ehrlich and Melvin Shabshin, *Psychiatric Ideologies and Institutions* (New York: Free Press of Glencoe, 1964), Chapter 2, "Logic, Techniques and Strategies of Team Fieldwork."

the data, since the constant comparisons force the analyst to consider much diversity in the data. By *diversity* we mean that each incident is compared with other incidents, or with properties of a category, in terms of as many similarities and differences as possible. This mode of comparing is in contrast to coding for crude proofs; such coding only establishes whether an incident indicates the few properties of the category that are being counted.

The constant comparison of incidents in this manner tends to result in the creation of a "developmental" theory.[11] Although this method can also be used to generate static theories, it especially facilitates the generation of theories of process, sequence, and change pertaining to organizations, positions, and social interaction. But whether the theory itself is static or developmental, its generation, by this method and by theoretical sampling, is continually in process. In comparing incidents, the analyst learns to see his categories in terms of both their internal development and their changing relations to other categories. For example, as the nurse learns more about the patient, her calculations of social loss change; and these recalculations change her social loss stories, her loss rationales and her care of the patient.

This is an inductive method of theory development. To make theoretical sense of so much diversity in his data, the analyst is forced to develop ideas on a level of generality higher in conceptual abstraction than the qualitative material being analyzed. He is forced to bring out underlying uniformities and diversities, and to use more abstract concepts to account for differences in the data. To master his data, he is forced to engage in reduction of terminology. If the analyst starts with raw data, he will end up initially with a substantive theory: a theory for the substantive area on which he has done research (for example, patient care or gang behavior). If he starts with the findings drawn from many studies pertaining to an abstract sociological category, he will end up with a formal theory pertaining to a conceptual area (such as stigma, deviance, lower class, status congruency, organizational careers, or reference groups).[12] To be sure, as we described in Chapter IV, the level of generality of a substantive theory can be raised to a formal theory. (Our theory of dying patients' social loss could be raised to the level of how professional people give service to clients according to their respective social value.) This move to formal theory requires additional analysis of one's substantive theory, and the analyst should, as stated in the previous chapter, include material from other studies with the same formal theoretical import, however diverse their substantive content.[13] The point is that the analyst should be aware of the level of generality from which he starts in relation to the level at which he wishes to end.

The constant comparative method can yield either discussional or propositional theory. The analyst may wish to cover many properties of a category in his discussion or to write formal propositions about a category. The former type of presentation is often sufficiently useful at the exploratory stage of theory development, and can easily be translated into propositions by the reader if he requires a formal hypothesis. For example, two related categories of dying are the patient's social loss and the amount of attention he receives from nurses. This can easily be restated as a proposition: patients considered a high social loss, as compared with those considered a low social loss, will tend to receive more attention from nurses.

11. Recent calls for more developmental, as opposed to static, theories have been made by Wilbert Moore, "Predicting Discontinuities in Social Change," *American Sociological Review* 29 (1964), p. 322; Howard S. Becker, *Outsiders* (New York: Free Press of Glencoe, 1962), pp. 22-25; and Barney G. Glaser and Anselm Strauss, "Awareness Contexts and Social Interaction," *op. cit.*

12. For an example, see Barney G. Glaser, *Organizational Careers* (Chicago: Aldine Publishing Co., 1967).

13. ". . . the development of any one of these coherent analytic perspectives is not likely to come from those who restrict their interest exclusively to one substantive area." From Erving Goffman, *Stigma: Notes on the Management of Spoiled Identity* (Englewood Cliffs, N.J.: Prentice-Hall, 1963), p. 147. See also Reinhard Bendix, "Concepts and Generalizations in Comparative Sociological Studies," *American Sociological Review*, 28 (1963), pp. 532-39.