BASIC PRINCIPLES OF

SURVEY RESEARCH

by

BRAM HAMOVITCH, M.A.

A Project

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Arts (Teaching)

McMaster University

MASTER OF ARTS (1989)   McMASTER UNIVERSITY
(TEACHING)   Hamilton, Ontario

TITLE: Basic Principles of Survey Research

AUTHOR: Bram Hamovitch,   B.A.   (McGill University)

                              M.A.   (McMaster University)

                              B.Ed. (University of Toronto)

SUPERVISORS: Professor D. Ralph Matthews

                Professor Jane Synge

NUMBER OF PAGES: v, 131

ABSTRACT:

This book represents an attempt to assist the senior high school student in his or her attempt to complete original survey research. The book elaborates on a number of the basic principles within social science methodology. There are discussions on how to choose an appropriate topic, the formation of hypotheses, proper sampling techniques, questionnaire development, data organization and analysis. The book also helps the student understand *why* the research must be done in a certain manner, in order for it to be scientific in nature. Detailed instructions are given on how the student is to go about writing up the results of his or her research efforts.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

An Introductory Note to the Instructor

The goal of this book is to provide the reader with a simplified understanding of the complex and sometimes controversial world of social science research methods. It is aimed primarily at high school students, although we hope that it will also be of value to students taking community college and university courses, who wish to have a straightforward introduction to the issues involved in social scientific research.

The author of this book is a high school teacher who teaches social science in the Ontario school system. This book is his response to recent changes in the curriculum for Ontario high school students. Newly introduced guidelines from the Ontario Ministry of Education now require that all students taking the Family Studies O. A. C. course The Canadian Family in Perspective must complete an original independent research project (worth 20 percent of the student's final grade) in order to

> . . . gain experience in research methods, such as interviewing and administering questionnaires, using instruments they have developed with the teacher's guidance. They can then summarize the data thus produced and tentatively identify trends. . .
> (Ontario Ministry of Education, 1987, p. 102).

Thus, Ontario teachers are being asked to give their students a first-hand experience in using the methods of

social science survey research.

Survey research involves the application of the scientific method to the study of *social* phenomena. It uses the questionnaire or interview schedule as the primary method of gathering data. We chose to limit our writing to this type of research for a number of reasons. First, it is a method which is commonly used in the social sciences. Also, this kind of study encompasses a large number of the main ideas of how social science research is practised. Thus, we are focusing on survey research because it serves as a focal point through which the student can learn a great deal about the scientific method.

We consider the idea of having students complete individual research studies to be a valid one for several reasons. The most important reason is the belief that any well educated person living in this society should have a basic knowledge of the advantages, limitations and frustrations of doing this type of research. This is primarily because of the pervasive impact of survey research on our society. The community is constantly being reminded of the results of public opinion surveys and social science research, on a wide variety of topics. Students, then, should be helped to understand the complexities of survey research, in order to more critically assess research results which are presented to them as "truth".

Another important educational consideration lies in the fact that this type of assignment requires the student to organize and incorporate a large number of ideas in completing a single project. The student is not only required to deal with the topic which he has chosen to study, but he must also pay careful attention to a rather large and detailed volume of information related to the manipulation of that chosen topic. In short, the carrying out of a survey research project represents a challenge to *all* students, forcing them to organize their time, as well as their academic efforts.

This book has three main sections. They reflect different stages in completing a research project. The first stage involves reading about the chosen topic, and formalizing one's ideas into *Hypotheses*, supported by *Theories*. The next stage directs the students to develop their questionnaires, and choose their samples with a number of considerations in mind. The last main stage of research involves the process of putting the answers to the questionnaires into proper *Tables*, and analyzing the results. *Chapter 1* introduces the readers to this whole process, thus allowing them to have some reasonable sense of direction while they are completing each stage. *Chapters 2-4* describe the above-mentioned process in the kind of detail necessary for the students to actually carry

out their own research. *Chapter 5* shows students how to
complete their reports on their projects with proper
Introductions and *Conclusions*.

There are a number of ethical issues which the
instructor must settle for him or herself, before guiding
students through an exercise such as this. It is important
that the instructor maintain the right to have the last
"say" before students go out into the community with their
questionnaires. In order to do this, the instructor must
be willing to take in and check the students' work, prior
to its completion. Our practice has been to check the
students' progress after the development of a *Theory*
section (*chapter 2*), as well as after the development of
the questionnaire and sample plan (*chapter 3*). This allows
the teacher to maintain control over how the students
impact upon the community. It also has the advantage of
allowing the instructor to ensure that all the students are
more or less "on track" in completing their assignments.

In the text of this book, we have stressed the need
for the students to ensure the privacy of their
respondents' answers. That is basic to this type of
research, and must be firmly and repeatedly stressed to the
students. Some of the student researchers may ask their
respondents to reveal rather sensitive areas of their
lives. They may report on their sexual behaviour, illegal
activities, family finances, and so on. Particularly when

sensitive information is being elicited from respondents,
we would suggest that you ensure that the students hand
their questionnaires out in a manner that ensures the
anonymity of each respondent. Even if this is done, there
may be situations where the student researchers are in a
position to recognize the identity of one or more of their
respondents. This would certainly be the case when the
students choose to interview their respondents. When this
occurs, the students must be made to understand that they
are engaged in a process which is bigger than themselves,
and which requires their commitment to "keep a secret". As
well, the instructor is going to have to come to his or her
own personal decision as to what types of questionning are
too "personal" to be acceptable at this level of study. We
suggest that this does not have to be decided rigidly in
advance, since very often, questions can be reworded to get
around this issue, and yet stay within the student's
original area of interest. In short, we feel that we would
be remiss if we did not warn teachers that there are likely
to be some ethical concerns in completing a project such as
this with a class. A frank and honest "coming to terms"
with these issues is advisable if you are to assign an
original research project. In this regard, you have a
responsibility to your students and yourself. But you must
also consider the implications of the students' work upon

the school, the community, and "social science" as a whole.

We have emphasized the challenging nature of
this type of student assignment. It should also be
mentioned that this type of work reaps its own rewards, as
well. Not only will most students gain the self-confidence
of having learned a new skill, but most will also feel that
they are learning something important about a small corner
of the society in which they live. In one word, doing this
type of research tends to be something which is "relevant"
to students.

Chapter 1   <u>Survey Research: An Overview</u>

The project which you are about to embark upon will
be unique and original to yourself. At the same time, it
will be very similar to those completed by your fellow
classmates. It will be unique in the sense that <u>you</u> <u>alone</u>
will pick the topic, and <u>you</u> <u>alone</u> will decide what
specific research to complete within that topic area.
Also, you will have to make a large number of personal
judgements or decisions at practically every stage of the
project. You will find that there will be ample
opportunity for you to use your imagination and your
thinking abilities throughout this project. In doing so,
you will depend upon many of the skills that have taken you
this far within the education system (eg. reading, writing
and research skills). On the other hand, you will also be
called upon to use skills which are often overlooked within
the formal education system, but which most of us have
developed within our daily lives. These are skills like
"common sense", "judgement", and just plain "taking a
chance" on something, based on the idea that it somehow
"looks right".

What do we mean when we say that your research
project will in many ways resemble those of your

classmates? In simple terms, this means that your research
will take place within a complex set of guidelines or rules
which have been assembled under the heading "the scientific
method". Most students have come across this idea before
while studying physical sciences, such as physics and
chemistry. Many high school students have completed
science "labs" which attempt to acquaint them with the "how
to" of using the scientific method. The fundamental aim of
this book is to instruct the student in "how to" use the
scientific method when studying "social" phenomena. You will
no doubt recognize many of the terms as being similar to
those used in your science classes. Yet social scientists
cannot use telescopes or test tubes in the study of human
behaviour. Instead, human behaviour can be best studied by
utilizing other techniques. One can use material from
interviews, questionnaires, case studies, etc. The purpose
of this book is to show you how social scientists actually
use these tools while applying common "scientific"
principles to the study of human beings. It will take many
hours of reading, thinking, discussing and writing to come
to a basic understanding of the "how to" of completing
survey research. This chapter will give you an
overview of the entire process.

Before you commence your own project, you should
have some understanding of where it fits within the context
of social science research as a whole. There are many

different methods which researchers employ in gathering information about the social world around them, and this book will only be focusing on one of them: the method that is known as *survey research*. We have chosen to focus on survey research for a number of reasons already mentioned in the *Introduction*. In general terms, survey research is the method of gathering information about people around us by using questionnaires or interviews. Another popular method of gathering information is *participant observation*. As the name implies, the researchers gather information by *participating* in the group that they are studying, and by *observing* the behaviours of the members of that group. *Content analysis* is another method used in social science. In this method, the researchers categorize what they observe (usually the written word) into categories which they have defined, and they then look for patterns within those materials. The last method which we would like to discuss is *historical analysis*. This method very often attempts to look for long term trends in social behaviour. This is usually done by analysing the content of existing historical documents, and comparing these findings with present-day observations. A wide variety of historical documents are used in these studies: for example, census data, newspapers, diaries, and letters. The point which we would like you to understand at the outset is that the

survey research which you are about to embark upon is not the only method of social science research, even though it is widely used and very important. Other methods of collecting information are valid, necessary and important in the sceintific attempt to make sense of and understand our human behaviour.

Here is a brief outline of the stages of a research project. The very first thing that students must consider at the outset is choosing a research topic. After the topic has been introduced, the students must formalize their thoughts by making a number of *hypotheses*. Next, they must develop a *research design* which systematically goes about the process of measuring and testing the above hypotheses. At this point, the *data* will be *collected*, usually by means of an instrument such as a questionnaire. They will then be *analysed* by the researcher, that is, explained and put into words. Finally, the research will be completed when an appropriate *conclusion* can be made.

Let us now consider the above five stages in turn. We will then go into more detail on their precise meaning in later chapters.

A. Research Topics

You are the person who chooses the research topic. After all, it is your research project. For some

students, choosing a topic will be a relatively simple, straightforward task. Others will find it very difficult to make a decision, and will choose and reject several before committing themselves to one. Here are a few "tips" on choosing a topic for your research project.

When studying in a new discipline, it is useful to consider the areas that have traditionally been of interest to those who work within it. Here is an example. One of the areas of intense concern to social scientists has been the question of social class, and how it affects a person's personality, values, and behaviour. A typical study in this area might start off by observing that other researchers have noticed various differences between classes, and seek to validate or refute some of those findings. Social scientists write about topics that are of historical interest to their discipline, as well as what is going on in their society around them at the moment. Thus, any list of topics must be understood as partial and changing with the times. The following list is in no particular order of importance or interest.

    Sociology of religion (or religious beliefs of
        different sub-groups, e.g. by gender, social
        class, age)
    Sex roles (and various types of behaviour, identity,
        and personality which are associated with gender)
    Race, ethnic or religious groups (practices, values,
        beliefs and commitments of different sub-groups)
    Sociology of education (e.g. behaviours, norms,
        beliefs and patterns within the system, its
        functioning, its perceived weaknesses)
    Sexual behaviour (or beliefs of a sub-group within the

population, e.g. by sex, age, social class, stage
of life, marital status, sexual orientation)
Sociology of the family (e.g. the effects of divorce,
birth position, stage of life, social class,
child-rearing practices, upon individuals)
Social class structure and social class behaviour (the
beliefs, values, and/or personality of different
social classes, and their relationship to other
aspects of society such as ethnicity, religion,
gender)
The effects of the media on our population (e.g. tele-
vision viewing of various types of people, their
reading of newspapers, listening to music)
Sociology of sports (e.g. their effects upon
participants and observers)
Formal organizations (e.g. places of employment,
churches, schools, hospitals, etc., and the
effects which they have on individuals)
Self-concept, personality, values or attitudes (and
their effects on individual's behaviour)
Deviance (e.g. law-breaking, drug abuse,
prostitution and membership in "alternative sub-
groups". Generally, researchers explore
"causes", "effects" and/or the "world-view" of a
deviant group.)
Stages of life (e.g. the behaviours and ideas of
adolescents, young married couples, couples with
children at home, the "empty nest" stage, old
age, infancy)
Politics (e.g. why people vote as they do, their
attitude toward power)
Rural vs. suburban vs. urban behaviours (attitudes)
Death and dying (how people think about and prepare
for their own or others' death)

Social scientists generally choose to enter areas

of investigation which are familiar to them on some level.

They start with the well known, and transform it from the

mundane by making us "see it in a new light" (Berger, 1963,

p. 21). This is done by looking for patterns which give us

new insights into the behaviour of the people around us,

and ultimately ourselves. These patterns indicate certain

relationships to us, showing us how our values, our

personalities, and our behaviours are related to our
membership in groups.   The truly exciting part of doing
social science research is that "things are not always what
they seem." To the voter, one's ballot is cast *as one sees
fit!* The social scientist sees the voter's behaviour as
being related to membership in various groups. For
example, voting patterns vary with social class,
neighbourhood, one's personality, attitudes, etc. Whose
view is more correct? Let us just say at this point that
the social scientist would argue that <u>both</u> views are
useful, in that they help us to gain perspective through
which we can better understand ourselves and others.

     In choosing your topic, try to start with an area
of research that is relevant to you, either because of your
past experiences in life, your present life situation, or
where you see yourself headed in life. Start out by
reading some articles or books written by social
scientists, or summaries of these, written by journalists.
Journalists write articles for newspapers and magazines,
whereas social scientists publish their research in books
and social scientific journals. Journalists usually obtain
their information about social phenomena by questioning
people who are chosen in a somewhat unsystematic and often
haphazard fashion. Social scientists, on the other hand,
take a great deal of time and effort in designing just how
and from whom they will obtain information, in a controlled

manner. Your job, then, as a beginning social scientist, will be to let the existing literature (or your ability to find it) help you define your topic area. If you have a firm idea of what you want to write about, write these ideas down in the form of questions, and then search your library for social scientific answers to your questions.

B. Hypotheses

Now that you have determined what your topic is going to be, your next task will involve "narrowing it down" to a manageable size, and putting it into a systematic form. In social science literature, you will notice that *variables* are shown to be related to each other. A variable may be thought of as a characteristic such as religion, income, or personality type, which <u>varies</u> from individual to individual, or from group to group. A person may differ from others in his religious practices and beliefs. An identifiable group may earn less income than our national average. When social scientists say that variables are *related*, they mean that they affect each other in some way. This means that changes in one characteristic are associated with changes in the other characteristic. Age and criminal behaviour are often found to be *related*. So too with income and voting behaviour.

You must do the reading on your topic area with the

following questions in mind:

1. Can I find any relationships or patterns between two or more "variables" or characteristics which are suggested by the literature, and which I consider to be interesting or significant enough to be the focus of my paper?

2. Can I find an "acceptable" level of academic "support" for each of the above relationships?

Let us consider these two questions, in turn. From your reading, you must develop a number (we suggest 2-3) of formal statements which systematically propose the existence of relationships between variables within your topic area. These are referred to as *hypotheses*. While these may simply be based on a hunch, they are more typically based on other researchers' studies, or on a more general theory (set of ideas) which logically suggests a relationship. Here is an example: "I think that a student's probability of succeeding in high school is directly related to the number of books present in his or her parent's home." This is only a hunch, because it came directly out of my head, without doing any specific reading to support it. Nevertheless, it could be formally stated as a hypothesis, upon which I could base my research study. Here is this idea once again, stated as a more formal hypothesis:

Hypothesis 1: Students whose parents' homes have a large number of books are more likely to attain higher grades in high school than those whose parents' homes have few books.

In sum, then, a hunch is precisely what one would think it is. It represents an intelligent (or thoughtful or commonsense) guess about the real world, which could later be supported with logic, facts and/or ideas.

As mentioned above, the vast majority of hypotheses are derived directly from the written works of previous researchers. In our experience, many students perceive this as a "rip off" of another person's work. This perception is correct in some ways. However, a more accurate thinking about this is to realize that you are "replicating" another person's work, to see if it indeed holds true today, in your own community, and among the actual people you include in your study. Allow us to give you an example of this process to show what we mean. Suppose that you are reading through a book, and the following words catch your attention: "The aptitude [potential] of Jews for suicide is always less than that of Protestants; in a very general way it is also, though to a lesser degree, lower than that of Catholics" (Durkheim, 1951, p. 155, originally published in 1897). Without much thinking or originality on your part, this brief statement can be quickly turned into the following hypothesis:

"Protestants are more likely than Catholics, who are in turn more likely than Jews to commit suicide."

This type of hypothesis is little more than a formal

restatement of the "hypothesis" and findings of the
social scientist quoted above.

The last type of hypothesis that we wish to share
with you is the type that is still based on the ideas of
another researcher, but does not apply directly to the
situation or topic area that you wish to describe. Most
commonly, the student may come across an appealing idea
which has been stated rather broadly or generally by the
researcher. The idea has a bearing upon the topic about
which the student is interested in developing hypotheses,
but only indirectly. In cases such as this, the student
must be prepared to argue logically that the researcher's
general statement likely applies in this context, for a
number of stated reasons. The following example may serve
to clarify this process. Suppose that you are interested
in the topic of social class, and come across the following
statement in your reading: "Difficulty will arise because
the value system held by members of a class will contain
specific values about their abilities, aptitudes and
potential that will constrain the aspirations, motivations,
expectations, etc. of that class" (Squibb, 1975, p. 145).
Certainly, a number of formal hypotheses can logically flow
from this stated idea: I will limit myself to the
following.

Hypothesis 2: A lower class student is more likely than a

middle or upper class student to believe that his/her ability to succeed in school is weaker than the average student.

Please note two things about this hypothesis. First, it is a clear, unequivocal statement which will generally be clearly understandable to the reader. Second, it makes a number of assumptions about what was meant by the original statement from which it was derived. What we mean by this is that the original author did not necessarily intend this exact hypothesis, but it is a logical extension of what he is saying more generally.

The last step in this stage of developing and stating hypotheses involves looking around for proper logical "support" for each of these hypotheses. This is done by referring to the research and conclusions of other social scientists in lending support for the "validity" of each of your hypotheses. If your hypothesis is based on a hunch, or is of the third type cited above, you must provide a convincing logical argument in support of the probable validity of the hypothesis. Basically, what you are doing here is answering a simple question about each hypothesis. You are asking, "What is there out there in the literature on this topic which leads me to believe that my hypothesis has credibility?" *Chapter 2* will go into considerably more detail on how to develop hypotheses from theoretical and research literature.

C. Research Design

This section of your paper will explain just how you plan to go about determining the validity of your hypotheses in the "real" world. Up until this point, you have mainly used the kinds of research skills which you have been developing since elementary school. You have been deciding upon a topic, reading about it, refining or limiting which aspects of the topic upon which you care to focus, and summarizing the ideas of others in some coherent fashion. At this stage, you are very likely to be entering an entirely new arena. Allow us to give you this warning in advance. As with most new things, you will probably experience some feelings of uneasiness along the way. We believe that this is mostly due to the fact that you cannot possibly see the finished product while you are still putting it together for the first time. Just remember, then, that having feelings of apprehension is normal under the circumstances. If you persist, in the end, you will come to understand a rather complex and important set of principles about how to do social science research properly.

Now, try to imagine the scientific process. To this point, you have come to understand, through reading articles and books, what the current thinking is in a given

subject area. You have formalized what it is that you want to test, by way of stating hypotheses. Now you must somehow determine if your "best educated guesses" (i.e. hypotheses) are correct or not. If they in fact turn out to be more or less correct, you can pat yourself on the back for having chosen valid ideas from perceptive researchers. However, if they turn out to be incorrect, it is entirely possible that you have stumbled upon something of real significance. Perhaps you have helped, in a small way, to push out the boundaries of our understanding in your area of research. Only by rigorously testing the ideas of others, can we confirm them, refine them or reject them altogether. This is the scientific process and method. In doing social science, you must respect this method or approach.

Part of showing respect for the scientific method means that you must develop a method of testing your hypotheses which is fair and open to the view of others who may be interested. This means that your research design section must explain to the reader exactly how you went about doing this testing. A number of decisions must be made at this point in your project about how you are going to proceed to determine the validity of your hypotheses. This involves a number of complex ideas which are better left to *Chapter 3*. However, in brief, it centers on the development of a questionnaire or an interview schedule,

whose questions are carefully designed to measure each variable in your hypotheses. It is only fair to point out in advance, that developing questions and asking them is a difficult and time-consuming process. *Chapter 3* will explore the whole question of how to go about developing questions which can be used to test your hypotheses. Several examples of questions will be discussed, which you can use as models for the development of your own questionnaire.

The *research design* section of your paper also involves explaining your choice of those persons to whom you will administer your questionnaire. This is referred to as selecting a *sample*, and it involves a number of serious considerations. These will be discussed on pages 74-8. You must keep in mind that you are going to select a small number of people (say 35-75 or so) from all those around you in your town or city, and you will ask them for some of their time to answer your questions. Suffice it to say, then, that the research design section of your paper will involve a full disclosure on your part of all the methods which you used in contacting your respondents, and in gathering and organizing the answers they give to your questions.

D. Organizing and Writing Up the Data

As the above title implies, there are two major tasks to be accomplished next in your journey toward completing your first original social science research project. The first involves successfully lifting the answers to your questions off your 35-75 completed questionnaires, and arranging the answers in *Tables* which are organized such that they can clearly reveal the meaning of those answers and their relationship to one another. This is not a difficult task, but it is very time-consuming if it is to be done properly. Professional social scientists have the advantage of computers and research assistants to help them in this process. It remains a time-consuming task, even with such help. *Chapter 4* will explain in precise detail just what is included in this process. Stated briefly, it involves coding or scoring each respondent's answer to each question, and then systematically adding up and cross-tabulating groups of people that answered alike.

The second task which needs to be accomplished in order to successfully complete this section of your paper involves answering this question: What do I do with the numbers on the *Tables* which took hours to calculate? Each *Table* must be accompanied by words which explain it to the reader. These words will discuss the extent to which the data do or do not support each of your hypotheses. This

section has only one aim. The purpose is to explain to the reader just what the data say to you about the nature of that one small corner of the world that you have been studying. I hear some of you asking, "What if my data don't support my hypotheses? What if I've done all this work _for nothing_?!" This is not at all unlikely, given the small number or respondents in your sample. Many real differences and trends are relatively minor within the population as a whole, and are only likely to show up consistently in larger samples. However, fortunately there is no need to worry about this problem. It does not matter whether or not your initial hypotheses are confirmed or rejected by the data from your sample. What counts here is that you have had the experience of "getting down to the truth". You will be given much more detail about how to "write up" your data in _Chapter 4_. For now, let us briefly consider the last important section of your research paper, that is, your _Conclusion_.

E. Conclusion

If you look up the word 'conclusion' in the dictionary, you will find the following kind of definition:

> Termination; final result; inference; decision;
> (Logic) proposition deduced from previous ones
> (Fowler & Fowler, 1964).

We draw your attention to these ideas in the hope that you

will avoid the mistake that many novice writers make in thinking that one's concluding remarks are mostly "summary", with very little "conclusion". While it is true that a proper *Conclusion* should seek to summarize important ideas or findings within the preceding paper, it must go well beyond mere summary. It should always refer back to and reinforce what was previously stated as the goals of the paper in the *Introduction*. Doing this ensures that the paper has a certain unity and cohesion to it. A well written *Conclusion* should attempt to say something which integrates and elevates the analysis which precedes it. The writer should use judgement in speculating about what can be surmised or inferred from the research. Thoughtful social scientists will often ask about the significance of their work for the society as a whole. Who could benefit from having the knowledge generated by this research effort? How could they benefit? What future research is suggested as being potentially fruitful? Why?

In short, your *Conclusion* must attempt to 'infer' what are the implications or is the importance of your efforts. Frequently, this is the worst done section of even the best students' papers. We can understand why: this involves standing back from all your work (almost as if it belonged to someone else), and thinking about it seriously. In your *Conclusion*, you have nobody to rely upon but yourself. It is up to you to give your work some

meaning. *Chapter 5* will provide further information with respect to writing up the concluding and introductory sections of your research effort.

Chapter Summary

In this chapter, we have outlined the five main steps involved in doing survey research. This was described as an application of the scientific method to the study of social phenomena, using questionnaires or interviews. It may also be thought of as research involving the use of surveys, i.e. questionnaires, as the principal means of gathering information from a sample of persons. In attempting to complete survey research following the scientific method, a number of steps must be carefully followed. First, a *research topic* must be chosen. This usually involves the study of social factors such as behaviour, values, ideas, etc. as they pertain to different groups of individuals. Once the student has done some serious reading on the chosen topic, a number of testable *hypotheses* are developed. These form the basis for investigation. At this point, the researcher then develops a *research design* to aid in testing the stated hypotheses. This involves, among other things, the development of questions to ask respondents, and a plan of action as to how those questions can help us better

understand the chosen topic. This research design also involves the careful consideration of who will be chosen to make up the *sample* to be questioned. After all the respondents have answered the questions, the data must then be systematically presented to the reader in *Tables*, and then *analysed* for the reader. Finally, a good researcher will give careful consideration to precisely what can be thoughtfully *concluded* from having gone through the above process with the chosen topic.

Chapter 2   Developing Hypotheses: Your Theory Section

In this chapter, we will try to give you a
practical sense of how to go about creating formal,
testable hypotheses when you are in the early stages of
your research project. In order to do this properly, you
must develop an appreciation for precisely what you are
doing when you "do" social science research. At the end of
the chapter, you will be provided with step-by-step
instructions on how to write up the *Theory* section of your
own research project. Let us start by considering the
most basic question. We need to know just what a
hypothesis *is*?

Hypotheses

It is not unusual for students who are taking
courses in the social sciences to feel that the material,
ideas, etc. are nothing more than "common sense". Students
often remark to others or think to themselves that they
could have told you roughly ninety percent of the ideas
contained in the text. This kind of reaction is quite
understandable. We all live out our daily lives by
carrying around and using a large number of generalizations

(general "truths") which we have stored in our heads. For example, when adult Canadians meet somebody for the first time, they can immediately _see_ the age and sex of the other person. Often, the very first piece of information that they seek from that other person is their "occupation". By gaining access to that single piece of information, they are able to make all kinds of hidden "generalizations" about that new acquaintance. They might predict the approximate amount of education that person possesses, the likely amount he or she earns, the probable size of their family home, the type of vacations which that person might be expected to take, the value which that person places on education, personal standards of cleanliness, honesty, etc. Based on these generalizations, a person might intuitively decide what kind of relationship he or she would like to have with this new acquaintance. For example, if you were introduced to a person who holds the occupation "labour union organizer", this might very well have an impact on how you perceive and behave with that person. All these predictions are based on generalizations, or general understandings as to what the world is usually like, which we all carry around in our heads. These are extremely useful to us, in that they allow us to successfully negotiate our way through our daily lives.

The development of generalizations is also the main focus of social science. Like the ordinary person, social

scientists also attempt to generalize in order to understand how our social world is ordered. The generalizations of social scientists are often given the different title: "hypothesis", or "hypotheses" in its plural form. Common sense and hypotheses are similar in that they both make a definite statement (or proposition) about the exact nature of part of our social world. However, the two ideas are not completely similar. Hypotheses tend to be broader than common sense generalizations, in that they attempt to go beyond the experience of any one individual. This means that social scientists look for general "truths" which hold true for many different people, often over significant time periods. Social scientists are usually not interested in generalizations which relate to unique occurrences, or only hold true for a very small number of people.

Why, then, do social scientists generate these hypotheses? They make these proposals in order to focus their studies on a particular aspect of a topic, rather than skipping all over the place from subtopic to subtopic. Notice also, that researchers need not necessarily believe in the hypotheses that they put forth in their research papers. Hypotheses have usually been suggested by the researched facts, ideas or theories of other social scientists, or from the researcher's own observation of the

surrounding world. Researchers can focus on hypotheses with the hope of finding them to be not supported, as well as supported by the data which they gather. The more usual situation, however, is that researchers slowly develop a belief in and commitment to the formally stated predictions which they have developed as hypotheses.

We will call the specific type of hypotheses that we will be working with in this book *associative hypotheses*. They are "associative" in the sense that they propose that one particular characteristic or occurrence is a factor which is often associated with another characteristic or occurrence. For example, one might hypothesize that "children from homes with fewer children are more likely to be more successful at school than those from homes with larger numbers of children". In this example, we are proposing that the characteristic "school success" is associated with or related to (among other things) "number of children in the home". In short hand notation, this hypothesis may be diagrammed as follows:

No. of children ↑ =====> School success ↓

[As the number of children in families increases, this is associated with a lesser level of school success by the children in those families.]

This type of notation is useful in that it allows the researcher to see clearly and concisely just what is being proposed. Notice that the hypothesis itself does not make

any attempt to answer the question as to why things are this way. It simply focuses our attention on the very narrow aspect of school performance, and tries to suggest one (among several) possible factors associated with success or lack of success.

Where do these hypotheses come from? This question is generally asked by students who are new to the field of social science. Often, they are contained on every page or two of social science readings that students undertake. Yet students often have a difficult time discovering or identifying them. This is because authors often do not formally state their generalizations (or those of others) as hypotheses. Instead, they are often contained within the texture of the writing itself. After a thorough search of the literature, if you cannot find a particular hypothesis or generalization (which you have in your head), it may be cautiously proposed as a working hypothesis, worthy of testing. However, in such a case, the student should be very careful to show how the proposed relationship (hypothesis) can be seen as related to or derived from established facts or ideas.

When we say that a hypothesis is a "generalization", this suggests that it is a general statement which attempts to summarize a number of occurrences that regularly repeat themselves in the social

world. For example, the statement "All little green people come from Mars" is a statement which suggests the origin of a whole group of individuals with at least one thing in common: their colour. As such it is a generalization. It is a statement that could be of interest to social scientists in that it deals with social issues like "place of origin" and "race".

With all hypotheses, however, the student must be careful to insist on the appropriateness of its *level of generalizability*. Every hypothesis must be at an appropriate level of abstraction. This means that it must be neither too broad nor too narrow. The researcher must always keep in mind that the hypothesis must not be so general that it includes too many individual cases in the world. Nor should it be so narrow that it only relates to a small number of individuals. An example of a generalization that is too broad for social scientific purposes is "God loves all His children." It is an assertion of a proposed general truth, but it is so broad as to include everybody (and, as such, gives us little useful information about any individual or group of individuals). The generalization "Sally gets up every day and puts on her shoes" is an example of the other extreme. While it may be true that Sally "generally" does what this statement suggests, it lacks enough scope to be of interest to social scientists, unless it could be shown that other

"Sally's" are likely to act in a similar fashion. What we are looking for lies somewhere in between these two extreme examples.

An acceptable level of abstraction in social science includes statements about identifiable categories of people within society. For example, one might hypothesize that "girls" are more likely than "boys" to think of themselves as "religious". There are at least four categories of people suggested by this hypothesis: "girls", "boys", "those who are religious", and "those who are not religious". An acceptable hypothesis, then, must recognize that people vary (or differ) from each other along certain dimensions: that is, they are *more or less* "religious"; they are "male" or "female"; they are *more or less* "loving"; etc. A hypothesis must contain conceptions which vary from type or category of individual to a different type or category of individual.

Here are a number of hypotheses that our students have successfully researched. Notice that they all include a reference to at least four categories of people, and they have an appropriate level of generalizability: neither too broad nor too narrow.

<u>Hypothesis 1</u>: People who watch a great deal of violence on television are more likely to act aggressively than those who watch little or no violence on television.

<u>Hypothesis 2</u>: People who belong to a lower social class

are more likely to: (1) act aggressively; and (2) commit violent crimes than those who are members of a higher social class.

Hypothesis 3: Students with a high "need to achieve" are more likely to have firm plans to go on to university or college than those with a low "need to achieve".

Hypothesis 4: Children who are brought up with strict discipline are more likely to have "self-discipline" in doing their school work than those brought up with lax discipline.

Hypothesis 5: Children who are brought up with strict discipline are more likely to be "conservative" about their sexuality than those brought up with lax discipline.

Hypothesis 6: Elder siblings are more likely to have greater suicidal tendencies than younger siblings.

Hypothesis 7: Children from homes with fewer siblings are more likely to be "successful" at school than those from homes with larger numbers of siblings.

These have been included here to serve as models to assist

you in developing your own generalizations or hypotheses.

Here are some more possibly testable hypotheses (in an

incomplete form), which could give you ideas about other

possible topics for research:

1. Working women are more likely than housewives to . . .

2. Higher achieving students are more likely than lower achieving students to . . .

3. Working class children are more likely than middle or upper class children to . . .

4. People who vote for the N.D.P. are more likely than Conservative voters to . . .

5 (a) Men are more likely than women to . . .
   (b) Boys are more likely than girls to . . .

6. People from rural areas are more likely than city-

dwellers to . . .

7. Protestants are more likely than Catholics to. . .

8. Principals are more likely than teachers to . . .

9. History teachers are more likely than science or mathematics teachers to . . .

10. People who grow up in large cities are more likely than those growing up in small towns or rural areas to . . .

11. People earning large salaries are more likely than those earning small salaries to . . .

12. Those who watch a lot of T.V. are more likely than those who watch little or no T.V. to . . .

13. Those who regularly break criminal laws are more likely than "honest" people to . . .

14. Those who are "leaders" within their peer (or other) group are more likely than "followers" to . . .

15. Mothers are more likely than fathers to . . .

16. Athletic people are more likely than sedentary people to . . .

17. Blacks are more likely than Caucasians to . . . [Name your favourite ethnic group] are more likely than "WASPS" to . . .

18. Children who are brought up with strict discipline are more likely than those brought up with lax discipline to . . .

There are countless others mentioned or suggested in the social science literature: all you have to do is start reading the literature.

<u>Variables</u>

In <u>social</u> science, it is often very difficult to

demonstrate that any given characteristic (X) is associated with or influences another characteristic (Y). Yet this is what hypotheses attempt to do. These characteristics (like 'X' and 'Y') are called *variables*, because they are capable of varying in amount, type, or some other manner. *Income* is an example of a variable, since individuals can be thought of as varying considerably from each other in the amount of income they receive. *Religion* is also a variable, in the sense that individuals have a variety of religious beliefs, and also vary in their level of commitment to their religion (the latter is referred to as "religiosity"). In short, all variables must possess a range of "variance" which would "vary" from individual to individual.

Hypotheses of the type which we will be dealing with always contain two different variables: one *independent* and one *dependent*. The independent variable may be thought of as existing "independently of" the other (dependent) variable. The dependent variable, on the other hand, "depends upon" the independent variable for at least part of its variance. When we say that variable 'B' is dependent upon variable "A", we are in fact suggesting that "A" contributes to "B's" variance. Very often, social scientists diagram these relationships between variables, to clarify them to the reader. These diagrams often look like this:

$$A =====> B$$

If we imagine that a fictitious individual was, say "high"
in "A", we could expect him to be, say, "low" in "B". The
independent variable can be easily spotted, since it is
generally mentioned first in a hypothesis (or diagram).
Logically, then, the dependent variable is usually
mentioned second. But knowing this does you no good, when
you are attempting to construct hypotheses yourself! If
you have two variables, how can you determine which one is
dependent or independent? Think of it this way. The
dependent variable "depends upon" the independent variable
for its variance. For example, let us consider the
following two variables, "ethnicity" and "social class".
Ask yourself which variable is responsible for variance in
the other? Does one's social class influence one to be a
more likely member of one ethnic group as opposed to
another? Or does being born into a given ethnic group
predispose one to be more likely to be a member of one
social class or another? Which of the following diagrams
is more correct?

$$social\ class =====> ethnicity$$

OR

$$ethnicity =====> social\ class$$

Since one's ethnicity is usually determined at birth, ethnicity is the independent variable, i.e. it is independent of variance in the other variable (social class). Since one's social class can be thought of as being at least in part influenced by one's ethnic group affiliation, social class is the dependent variable. Hence, the second diagram best describes this relationship.

To help you determine dependent versus independent, these two ideas may be helpful:

(1) the independent variable is the one which usually occurs first in people's lives;

(2) the independent variable can be logically explained as being responsible for variance in the dependent variable.

Let us briefly reconsider two of the above-mentioned hypotheses (see pages 33-4). *Hypothesis 1* proposes a relationship between watching violent T. V. programs and aggressive behaviour. Which comes first: watching violent programs or the aggression? We suggest the use of common sense in answering this question. Does one's aggressive behaviour predispose one to watch more (or perhaps less) violence on T.V.? Or does viewing many hours of violent programs on T. V. predispose one to act aggressively after viewing? What do you think? Our answer to this question is contained in the order in which the two variables appear in our formal statement of the hypothesis

(see pg. 33). More precisely, we consider the aggressive behaviour to be the dependent variable, since our understanding is that the amount of violent television which people watch has some influence upon their aggressiveness.

Let us also consider *Hypothesis 3* (above). The two variables thought to be associated are one's "need to achieve" (an attitude) and one's future educational plans. One could argue this hypothesis both ways: one's attitude logically influences one's educational goals, OR one's educational goals logically influence one's attitude. We would guess that the researcher thought of the attitude as being the independent variable in this example, mainly because it is generally established early in life, and remains relatively inflexibly throughout a lifetime. The researcher would know this, because prior to developing this hypothesis, (s)he no doubt came across the research of other social scientists which suggests the relative inflexibility of one's attitudes (at least in this aspect of one's personality).

Sometimes it is difficult to know for sure which variable influences the other (even with the above suggestions). Consider the following example:

Hypothesis 8(a): Those who perceive the punishment for theft to be "light" are more likely to commit theft than those who perceive it to be "severe".

In this example, "perception" is presented as the

independent variable (mentioned first), and "theft

behaviour" is the dependent variable. The originator of

this hypothesis no doubt thought that theft behaviour was

dependent, at least in part, upon one's attitude toward

getting caught. The thinking here was that the attitude

comes first, followed by the behaviour (theft or non-

theft). But consider the following:

Hypothesis 8(b): Those who commit theft are more likely
to consider the punishment for theft to be "light" than
those who do not commit theft.

In this reversed hypothesis, the behaviour is the

independent variable, influencing the attitude (dependent

variable). Is it not possible that those who have

experience with an activity (eg. theft) slowly develop a

change in attitude that comes about as a result of one's

participation in that activity? The question to be

answered in cases such as the above is this. Which usually

comes first? Is it the attitude or the behaviour that

comes first? In instances such as this, there is no exact

"formula" for determining the "correct" approach. We

suggest that you recognize the problem, and try your best

to deal with it logically. You may end up by concluding

that the issue cannot be resolved, and that you are looking

for an *association* between the two variables, without

determining a <u>direction</u> of influence. In most cases

however, the issue can be resolved by using your judgement.

Read the literature to find out how others have sorted out

the problem. From your understanding, make the best

decision you can about the direction of <u>most</u> of the

variance. Does "A" generally predispose "B" to occur, or

vice-versa? Or should I simply state that the relationship

runs in both directions?

****

<u>Student Exercise</u>:

How would <u>you</u> handle the above example, involving theft and
perception of punishment? Why would you handle it this
way?

****

Sometimes researchers are convinced that they have

found a relationship between two variables, when in reality

there is no relationship at all. Let us consider the

researcher who is pondering over the relationship between

the rooster crowing and the sun rising. The researcher has

observed that every morning, just before the sun appears

over the horizon, one is able to hear the rooster crow.

After considering this for a while, the researcher might

propose the following hypothesis:

<u>Hypothesis 9</u>: The rooster's crowing causes the sun to
rise every morning.

We know, of course, that the rooster does not possess the powers attributed to it by our fictitious researcher! But, to a certain extent, we can "relate to" his hypothesis, which was based on his observations. The point of this story is to warn you about proposing similar hypotheses of your own. It is not enough to argue that "factor A" is usually associated with "factor B". When you propose a hypothesis, you must assure yourself (as well as your reader) that "factor A" is indeed responsible for at least some of the variation in "factor B". This is done by explaining precisely how "A" acts upon "B", so that "B" can be thought of as a logical consequence of "A's" acting upon "B".

Choosing Testable Hypotheses

In deciding upon the hypotheses for your research project, you must keep one other important idea in mind. You must be practical. You must remember that your time, resources and experience are all severely limited, compared to that of professional researchers. As a result, you must be extremely selective in choosing which hypotheses you wish to pursue. A couple of examples will assist you in understanding this point. Try to imagine yourself testing the following hypotheses.

Hypothesis 10: Today's high school students are less racist than high school students were 20 years ago.

Hypothesis 11: Today's seventeen-year-old high school students are more likely than those of the last generation to have a part (or full) time job during the school year.

Hypothesis 12: Teenagers from single-parent) families are more likely than teenagers from two-parent families to break the law and be convicted.

Before you set about seriously researching and testing any hypotheses, you must consider one very important issue: Is there a group of people in the community, who can help me test my hypothesis, and are accessible enough to answer my questionnaires?! *Hypothesis #10* (above) looks promising right away, because you could probably gain access to high school students relatively easily (they are nearby, usually willing, and often gathered in one place!) But how could you handle the problem of comparing students today with those of twenty years ago? We suggest that you could not test this hypothesis, precisely because you have no readily available sample of high school students from twenty years ago. The only exception to this advice would be when you are able to find a prior study in the library which would give you information on the racist attitudes of students twenty years ago. This information could then be used as a comparison with today's students. However, it would not be acceptable to find some older people today who were high school students twenty years ago, and ask them about their

racist attitudes at that time. They simply would not be able to remember accurately enough (for social science purposes). They could guess or give you their impressions, but those are not the same as determining their actual "racism". In sum, watch out for hypotheses involving extended periods of "time" as one of your variables: they are often difficult (though not always impossible) for you to test at this level.

You might have better luck with *Hypothesis 11* (above). This one also involves "time" as one of its variables. It postulates a relationship between "time" (one generation -- twenty years) and "work habits of high school students". This is more testable than *Hypothesis 10*, in that middle aged respondents could reasonably be expected to accurately remember whether or not they were employed when they were seventeen. Alternatively, you could possibly obtain statistics on the employment rate among high school students twenty years ago in the social science research literature.

****

Student Exercise:

How would you find people to ask about their employment at age seventeen (both today and twenty years ago)?

****

Hypothesis 12 should pose little difficulty for the

researcher. The hypothesis demands that you look for a group of people to answer some questions about these two variables: "type of family" and "trouble with the law". Yes, it is true that people often find both of the above types of information to be highly personal. Who is going to tell you about their family's status, and whether or not they have ever been convicted of a crime? The simple answer is: "Almost everybody." People will answer the most intimate kinds of questions, providing:

(1) they are clearly phrased so that they are readily understood;

(2) they are politely approached;

(3) they are assured of the confidentiality (i.e. anonymity) of their answers;

and (4) they are told _why_ you are asking.

We have had students successfully request information about all kinds of sensitive areas in respondents' lives: eg. sexuality, death of a loved one, depression and suicide, income of family, etc. Remember that if people are properly approached (considering the above 4 points), they are usually pleased to be part of the scientific process: it makes them feel helpful and perhaps even important! However, be advised that you must never breach the promise of confidentiality which you have assured your respondents. To do so is to put the whole process of doing social science

research into jeopardy. To help assure anonymity, social scientists generally omit names from their questionnaires and interview schedules.

In sum, then, be cautious when choosing your hypotheses. Make sure they are testable by yourself, with a reasonable expenditure in time and effort. Avoid hypotheses that demand that you hand out your questionnaire in all kinds of impossible situations: eg. twenty years ago, in a prison, mental ward, another country, etc. -- all of which are inaccessible to you! Try to choose a topic which allows you to develop hypotheses about areas of social life that are of significance to you and in which you have a keen interest. This will ensure that you stay interested in the project. This will also give you the additional "edge" of having a personal stake or interest in the outcome of your testing.

Supporting Your Hypotheses: The Theory Section Of Your Paper

After you have selected your hypotheses, you must then carefully explain why you put forward the hypotheses you did. The *Theory* section of your final paper will be divided into as many sections as you have hypotheses (we suggest 2-3 for most beginners). Each section will start with a formal statement of your hypothesis. Look once again at the examples of hypotheses throughout this

chapter, and ensure that your hypotheses are in that form.
Then go on to "support" each of your hypotheses with ideas
and facts collected from the research of other social
scientists (or summaries of the same by journalists). Make
sure that the discussion which follows each hypothesis
clearly explains any ambiguous terms which may be contained
within that hypothesis. For example, in a hypothesis
relating gender to personality differences, you need not
bother explaining precisely what is meant by the term
"gender" (since it is a widely understood term). However,
you must clarify precisely what you mean by these
"personality differences", so that the exact meaning of
your prediction is clear to the reader.

The bulk of your discussion in support of your
hypothesis, however, must involve an examination of how or
why your two variables are related to each other. You must
carefully answer the question, "How does 'factor A'
influence 'factor B'?" When you explain that one factor
influences another, try to show how it actually impacts
upon people in a significant manner. In attempting to do
this, you should feel free to use the comments and findings
of other researchers, but do not let them dominate or
completely take over control of your support. Remember two
things: any outside research which you cite must be
relevant to the specific hypothesis with which you are
dealing, and it must be you who maintains control of the

discussion (not your references). Carefully link sentences together to form a paragraph, and a number of paragraphs together to convince the reader that your hypothesis is a reasonable and logical judgement about the nature of some aspect of social reality. You should also make sure that you properly reference the sources of your supporting quotations and paraphrasings (see appendix 3 on citing references).

When social scientific support cannot be found, you must be very careful to develop a logical argument in support of your hypothesis. This must be based on researched ideas and facts. You should not be encouraged to pick an idea out of the sky and simply "run with it". Conjecture and intuition are acceptable in social science, but hypotheses based on these must be well argued and supported with good solid evidence of why they should be treated seriously.

Chapter Summary

We started this chapter by pointing out the role that *Hypotheses* play in helping researchers focus their studies on a specific topic and area of interest within that topic. The type of hypotheses which we will be using in this book are called *associative hypotheses*, because they include two variables, one of which has a determining

influence on the other. These are referred to as
independent and *dependent variables*. It was noted that
hypotheses must be at an appropriate *level of
generalizability*: neither too broad nor too specific.
Hypotheses in social science tend to separate people into
categories (by their various characteristics). Social
scientists then propose that these characteristics
influence (or are influenced by) our behaviour, attitudes,
etc. The beginner researcher must be very cautious in the
choice of which hypotheses to select for use in this
project: make sure your hypotheses are testable by you.
The chapter ended with a discussion of how to organize the
Theory section of your final paper, by focusing on the need
to find acceptable academic support for each hypothesis.

Chapter 3   <u>Measuring Variables and Choosing a Sample</u>:

                <u>Your Research Design Section</u>

Now that you have chosen your specific research topic, and have narrowed it down to a statement of two or three hypotheses, you are ready to see if your ideas hold true for a sample of individuals that you will carefully select. In order to do this, you must pay close attention to some of the rules and methods which social scientists have developed over the years, in their attempt to ensure "fair" or "accurate" testing of hypotheses. This chapter explains some of these basic rules and methods, so that you can understand them and apply them to your specific research project. This chapter will discuss two main areas of concern which need to be considered by you in order to accomplish this goal. First, we will discuss the measurement of your variables, before turning to the proper choice of a sample. While we can only consider these two aspects separately here, please bear in mind that they are interrelated aspects of your research project, and must be considered as such by you. How you measure a variable has an impact on who you should ask to answer your questionnaire. And who you decide to include in your sample influences the type of questions that you should

ask.

It will be assumed in this book that you will be developing a questionnaire or interview schedule with the aim of distributing it (or administering it) to approximately 35-75 people in your city or town. The closer that you come to the higher figure, and the more carefully you select your sample, the more likely you will be to obtain data which *accurately* reflect actual trends within the population that you are studying. For example, imagine the folowing situation. You are studying a community to determine its reaction to a proposed nearby highway extension. Would it be better to question 35 residents on your street, or 75 residents randomly chosen from across the entire community? Clearly, the latter choice is superior. This is true first of all because the larger number of respondents allows the researcher to get closer to the true range of thinking by community residents on this issue. Secondly, and equally importantly, having residents from all over the community is clearly superior to only one location. If you chose only one street, you must consider that its location in relation to the new highway might have a large impact on the residents' thinking on this issue. So when you come to choose your own sample, try to make it close to 75 respondents, and as truly representative as you can.

You will also have to decide upon *how* you will

gather information from your sample. Will you use the questionnaire or the interview method? The advantage of the questionnaire over the interview is that it allows you to quickly distribute it to your total sample. This is especially true when you can gather your sample or part of your sample together, and administer the questionnaire to them as a group. The interview has the disadvantage of being time-consuming, but researchers often feel that it allows for a more thorough communication between the respondent and the researcher on the topic being explored. The choice is up to you. But in either case, you must go through the same process of carefully developing appropriate questions to ask your respondents. Although these are by no means the only methods of data collection used by social scientists, they remain very useful and popular methods. Every year, thousands of survey research projects are completed around the world by social scientists. These studies attempt to test social scientists' ideas against how things actually occur in the "real world". It is this tradition into which you are entering.

If you are careful and thoughtful in completing this stage of your project, the data which you collect will give you very useful information about the nature of the world. If this stage is done poorly, you will experience

frustration in determining the accuracy of your original hypotheses. For now, let us consider the issue of developing the questions for your questionnaire.

Developing your Questionnaire

Just what is a questionnaire? Most people have answered them at least once during their lifetime, so you have probably had some experience with questionnaires. A questionnaire is a series of questions which the researcher will request a number of people to answer. To a social scientist, it represents more than just a series of questions on a sheet of paper. It is a carefully developed document which attempts to measure the position of each respondent along a number of dimensions. It is well planned in advance, in terms of how the questions will be phrased, as well as the proper ordering of them. This chapter will attempt to acquaint you with some of the problems which must be dealt with from the point of view of the "questioner". We will assume that this is new territory for you.

Before you sit down to compose your questions, you must be aware that what you are essentially doing in your questionnaire is measuring variables. It will be recalled that *Chapter 2* explained the process of hypothesis formation as being at least partly a matter of linking

related independent and dependent variables. When you think about one of your variables in terms of how to measure it, only think about the exact meaning of that one variable. Do not think about any of the <u>other</u> variables which you will also be measuring in each of your respondents. If you have developed three hypotheses (as per the instructions in *Chapter 2*), you will have to measure up to six separate variables (two per hypothesis). Since your research project should probably <u>focus</u> on one variable, you will have only three or four variables to consider for measurement (since one is repeated in two or more hypotheses). For example, your three hypotheses may take this form: 'A' =====> 'B'
                    'A' =====> 'C'
                    'A' =====> 'D',
where 'A' is gender, 'B', 'C' and 'D' are different attitudes, values or behaviours.

Before you attempt to measure your variable, you must <u>define it precisely</u>. You must know exactly what it is that you want to measure before you go ahead and do so. This process of moving from an exact definition of a variable toward actually measuring it with a question or series of questions will be called *operationalizing* the variable. Let us consider the following hypothesis:

<u>Hypothesis</u>:   Boys are more likely than girls to be
          independent and achieveing.

Before proceeding to develop questions which determine each
respondent's position along the range from "independent and
achieving" to "dependent and not achieving", you will have
to come to grips with an answer to these questions: Just
what do I mean when I say "independent"? What does
"achieving" really mean, as I have used the term? This
process of closely defining each variable is essential to
the process of developing good questions.

Validity and Reliability

Your operationalization must be valid in the sense
that it should accurately reflect "reality". If you are
attempting to measure respondents' attitudes toward what
they watch on television, your questions must be carefully
designed to measure that exact attitude, and not, for
example, what the respondents think is the "proper"
attitude to hold toward what they watch on television.
There are several methods that social scientists have
developed in an attempt to insure the validity of their
questions. On page 57, we discuss the methods of using
known experts or *known groups*. Most social scientists
attempt to assure themselves that their questions are valid
by the use of logic (Goode and Hatt, 1952). They carefully
define the variable which they are attempting to measure,
and then develop a question or questions which directly

relate to that definition. So, if the researcher is attempting to measure the "conservatism" of each respondent, he or she would first define it carefully. In doing this, the researcher would quickly come to realize that it is not just "conservatism" that is at issue here. Questions must be developed that reflect the reality that people have a variety of attitudes which range from "conservative" to "liberal". The rersearcher's job in developing questions is to "properly" place each respondent along the range of that dimension. In doing so, he or she might ask questions of the respondents, asking them to report their attitudes on various issues: e.g. divorce, premarital sex, how the country should be run, etc. The researcher would then classify the responses by referring to the researched definitions of "conservative" or "liberal". Think of it in the following way. You must design questions which accurately spread your respondents out along the range of variation of your variable, as you have defined it. Your questions must be designed so that respondents' answers spread them out over the range of the variable, as it exists in the real world.

How can you know when your questions are "valid" in the sense that they are measuring precisely what you intend them to measure? In the strictest sense, you cannot. But social scientists have developed a number of techniques for checking on the validity of questions. One is to check the

apparent validity of the questions with a number of *known experts* in the field which the questions are attempting to measure. For example, one might ask a veteran politician for his or her opinion on the validity of a series of questions which attempt to measure "conservatism". A second approach is to pretest your questions in *known groups*. So if you ask some representatives of a "known" conservative group to answer your questions, as well as some members of a "known" liberal group, you could check their responses against what you expected.

Sometimes researchers will think that they have developed valid questions measuring their variables, when the reality is otherwise. For example, if you are attempting to measure *religiosity* ("religiousness"), you could start out by considering a question such as the following:

Do you believe in God? Yes___   No___

What, in fact, does this question really measure? As far as we can determine, it measures precisely what it looks like it measures: belief (or nonbelief) in a deity. Yet, if the researcher accepted answers to this question as an indication of the *religiosity* of the respondent, he or she would be sadly mistaken. For one thing, this question would too rigidly divide the respondents into only two

types or categories of people: "believers" and "nonbelievers". Most of us are aware that people are not simply reducible to two "types" in this regard. Also, we know that there are some "believers" who are not very religious at all, at the same time that there are some "nonbelievers" who are quite religious (in the broad sense of that term). So beginner researchers must be extremely cautious to ensure that their questions actually measure what they claim to measure.

What this all gets back to is the central idea that the researcher must know precisely what it is that is being measured before the job can be done successfully. Do you wish to determine *religiosity* or *belief in a deity?* Do not confuse the measurement of one as being an indication of the other! Do not think to yourself: "These two variables are usually closely associated in people anyway, so by measuring one, I am in fact getting a good measure of the other." This would be an unacceptable thing to do, because when you do this, you are not actually measuring the exact variable which you had in your hypothesis. If you want to measure a variable, it is often advisable to be as <u>direct</u> as possible. To refer back to our example once again, one question used by social science researchers to measure religiosity is the following, which gives the respondent five points to choose from in answering:

How religious a person do you consider yourself to be?
(Please circle one):
Extremely   Somewhat   Undecided   Somewhat   Not very
Religious   Religious   Religious   Religious
                                                   At All

This question is referred to by social scientists as being
subjective. It is subjective in the sense that it asks the
"subject" to look into his "self", in order to report back
to you (the researcher) on his feelings about himself. It
has the advantage of limiting the options which the
respondent has to choose from, so that his position along
the range of the variable is not measured in an infinite
number of steps. [Notice that a question that allows for an
open-ended answer would be much more difficult to assess in
terms of actually placing each respondent along the range
of the variable under consideration.] Also, please note
that this question allows for a variety of answers (thus
reflecting the complexity of the real population, which
contains a variety of positions on this issue). The number
of possible answers is limited to a convenient number (in
this case, five) which has the advantage of simplifying the
sample nicely into five categories, without oversimplifying
the variable.

Aside from subjective questions, there is yet
another type of question referred to as *objective*. These
questions ask the respondents to report on what is
obviously real to them, without any need to interpret or

give their feelings about a subject. Information about a
topic can often be elicited in an objective OR a subjective
manner. For example, suppose that one of your variables is
"social class". You might try either of the two following
approaches, both of which will "work" to elicit information
on the social class position of one's respondents:

Subjective question:

Which of the following social classes do you consider
yourself to be a member of? (Please check one)

        Working class ___
        Lower-middle class ___
        Middle class ___
        Upper-middle class ___
        Upper class ___

Objective questions:

1. What is your occupation? _____

2. What is the highest level of education that you
   have reached? _____

3. What is your yearly income? _____

****

<u>Student Exercise</u>:

>Look once again at the subjective question and the
>objective questions asked above. Consider this
>additional fact: Almost all Canadians consider
>themselves to be "middle class".

>1. If your research project required you to
>   develop a measure of social class, which of the
>   above two approaches would <u>you</u> adopt? Support
>   your answer with good solid reasoning.

****

In developing your questions, you must attempt to
minimize *bias* and maximize the *reliability* of the data
collected (Selltiz et al., 1966). Bias occurs when the
researcher asks questions in such a manner that the
respondent is encouraged to answer in one way as opposed to
others. Consider the following question:

Drug abuse is a major problem in our society. Do you use
marijuana? ___ Yes   ___No

Clearly, the researcher has biased the respondent in a
direction away from complete honesty in this question. The
respondent who smokes marijuana has been systematically
discouraged from answering "Yes". The wording of this
question is <u>not</u> neutral. Bias has been introduced. You
must be sensitive to the little ways in which bias can

creep into your own questions. Ask yourself: am I asking this in such a manner so that everyone feels encouraged to answer it honestly?

The researcher can consider the data to be *reliable* when he or she has confidence that similar results would be achieved if the data was collected for a second time with a new sample. For example, many radios are reliable, in that if they are tuned to a given frequency on the dial, they will repeatedly reproduce the same station. But survey questions may lack reliability. It has been shown, for example, that the same questions will get different responses, depending on their location in the survey, or the time at which they are administered. All we can suggest at this point is that you should be aware that reliability is always a concern for researchers. In fact, many social scientists deliberately use other researchers' questions (or variations of these), so that they know that they are comparing like measures of the same variable.

How many categories (of answers) is appropriate for the questions which you will be developing? The simple answer to this question is that there is no simple answer! If you want to categorize respondents by their *gender*, a question such as the following with two answers will generally be sufficient:

What is your sex? Male____ Female____ (check one)

However, if you want to categorize respondents by their *religion*, you have a choice to make. You can ask a multiple-answer question, such as the following:

1. What is your religion?   (Please check one):

        Agnostic ____
        Atheist ____
        Anglican ____
        Baptist ____
        Islam ____
        Jewish ____
        Greek Orthodox ____
        Lutheran ____
        Presbyterian ____
        Roman Catholic ____
        United Church ____
        Other (Please specify):_____

Clearly, even this long list is only partial, and could easily be expanded significantly. Notice also that an "other" category is included, to allow respondents who don't fit any of the above categories to express themselves to you. The other possibility that you have open to you in measuring this variable is to do some prior *grouping* of your respondent's answers, in a way which is related to your hypothesis (or the theory behind your hypothesis). Do you remember the hypothesis mentioned in *Chapter 2*, which relates suicidal tendencies to one's religion (either Protestantism, Roman Catholicism, or Judaism)? If that was your hypothesis, you could phrase the religion question as follows:

2. What is your religion? (Please check one):

    Jewish ___
    Protestant ___
    Roman Catholic ___
    Other (please specify):_____

Given this hypothesis, you don't care about the intricacy concerning which subdivision of the religion the respondents belong to (because the hypothesis is silent in its prediction about those sub-groups). As such, in this case, question #2 should be asked, rather than #1 (above).

The careful reader will have noticed that all our examples of questions have not only asked the question, but have also provided the respondent with a series of answers from which to choose. This type of question is referred to as being *closed*. Closed questions offer a great advantage, especially for the beginner researcher, in that these questions are much easier to quantify later on. Students are often tempted to ask *open* questions, such as:

What is your family income? _____

What are your goals or aspirations for the future? _____

There is nothing *unscientific* about asking open questions such as the above. However, closed questions offer several advantages. In the case of the family income question, respondents may feel inhibited about writing down an exact

figure, but they may be much more willing to identify the income level into which their family falls. The goals/aspirations question, stated as is, would require considerable sophistication on the part of the researcher, who would have to "group" the answers when they all come in. Our advice, then, is for you to devise questions which are closed instead of open, for the reasons discussed above.

****

Student Exercise:

For each of the above variables ("family income" and "goals":

1. Invent your own hypothesis which includes the variable which the question is attempting to measure, as well as another variable made up by yourself.

2. Compose your own *closed* alternative to the *open* question provided.

****

In summary, then, researchers must strive for *validity* and *reliability* in developing their questions, as well as an absence of *bias*. This means that the questions must measure the variables in such a way that they mirror as accurately as possible each respondent's real "position"

along the dimension of each measured variable. At the same time, the researcher must strive to ask questions which yield results that coincide with those of other researchers. You might look in the social science literature to see how <u>other</u> researchers have measured your variables. Whenever possible, we would recommend that you avoid the use of *open* questions, and limit yourself to *closed* ones. As we have already suggested, you would do well to consider seeking the assistance of *known experts*, (<u>knowledgeable</u> friends or acquaintances) by asking them whether <u>they</u> think your questions measure the exact variable that you hope they measure. It is standard procedure among professionals to *pretest* their questions by administering them and discussing them with others, <u>before</u> giving them to their samples. Also, it is sometimes useful for you to administer your questions to a number of *known groups*, and then compare your predictions with the results. One problem with making up questions for a questionnaire is that, after a short while, you lose your objectivity about what you are doing and can no longer put yourself in the role of the respondent. Making up questions is not an easy task: in fact, it requires considerable thought and sensitivity.

Scoring and Grouping Variables

Some variables cannot be measured by simply asking the respondent one question to answer. In situations such as these, you must be prepared to develop a *score*, which reflects the position of each respondent along the range of that variable. Since each respondent will have two or more answers, each of those answers must be assigned a score by you, according to a predetermined, logical scheme developed by yourself. Let us consider two examples of scores in detail: first "current grades", and then "social class".

It soon became obvious to these researchers that we could not measure a student's current grades accurately by asking only one question, which requires a single response. So we developed the following question, which requires a number of answers from each respondent.

In the space below, please indicate the course codes, names and most recent grades of each of the courses you are presently taking, as reported on the most recent report.

Subject Code	Subject Name	Advanced?	Mark achieved on the most recent report
eg. NF00A1	Canadian Family	Advanced	79%
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____

This question will be scored according to the following instructions:

> We are attempting to provide each student-respondent with a fair assessment of his academic skill level, in relation to his peers. In attempting to do this, we run into the problem of what to do with the difference between "general" and "advanced" students' marks. To meet this problem, we subtracted 20% from the actual mark achieved in each general course. In this way, it is hoped that we were able to "even out" the effect of taking different course levels upon a student's average, and achieve a reasonable comparison. This process will hopefully determine the relative success of each respondent in school (rather than his or her actual grade average). The last step is to classify the respondents' grade scores into three categories: a "high", a "medium", and a "low" category. This will be done by grouping the 1/3 of the respondents with the lowest grade scores into the "low" category; the middle 1/3 will be grouped and called the "medium" category; and the highest 1/3 will be called the "high" category.

Let us now consider how we would score social class, which cannot be measured without asking a series of questions of each respondent. When one investigates how this concept is traditionally defined within social science, one typically arrives at a definition such as the following:

> The most commonly used objective criteria of class are income, occupation, property ownership, and education, all of which are ways of expressing objective economic differences among members of society (Porter, 1971, p. 10).

Since we do not have easy access to the current market value of all our respondents' properties, we have interpreted the above definition of social class to include three different aspects: *occupation, education* and *income.* As such, the following series of questions can help to determine a measure of social class for each respondent.

1. What is your occupation? _____

   Please include a brief description of your actual job:

   _____

   _____

2. What is highest level or grade of schooling that you have _successfully_ completed?

___ Some elementary school, but no high school
___ Some high school
___ Completed high school
___ Some post-secondary schooling
___ Completed an apprenticeship or community college program
___ Completed a 3 or 4 year university degree
___ Five or more years of successful post-secondary study

3. Into which of the following groups would you say your income will fall this year? (Please check one)

___ Under $5,000
___ $5,000 - $9,999
___ $10,000 - $14,999
___ $15,000 - $19,999
___ $20,000 - $24,999
___ $25,000 - $29,999
___ $30,000 - $34,999
___ $35,000 - $39,999
___ $40,000 - $44,999
___ $45,000 - $49,999
___ Over $50,000

Let us now consider how to score the concept of social class. Our first question deals with the _occupational_ component. We will review each respondent's answer to this question, and compare it to the extensive list of occupations and their comparative prestige levels, developed by Pineo and Porter (Pineo & Porter, 1967, p. 64-68). Each occupation is apparently rated out of 100. We will assign an occupational prestige score to each respondent by noting what they have written down as their occupation, and comparing it to the Pineo-Porter list. If the exact occupation of the respondent cannot be located on the list, we will extrapolate from other known _similar_ occupations.

The second component of social class is _education_. Since the occupational score was based on the range 0-100, we will do the same with our educational score (to ensure more or less equal weighting of the two elements of social class). We will assign "12.5" to those respondents who answered

"some elementary school, but no high school"; "25" to those who answered "some high school"; . . . "87.5" to those who answered "five or more years of successful post-secondary study". This attempt to weight education equally with occupation flows directly from the definition of social class which we are working with.

The final factor, *income*, will be scored as follows. Since the occupational and educational scores were based on the range 0-100, we will do the same with our income score (to ensure more or less equal weighting of the three elements of social class). We will arbitrarily assign "5" to the 10% of our respondents with the lowest income; "15" to the next 10%; "25" to the next 10%;. . . and "95" to the last 10%.

Lastly, we will add up the three scores (occupational prestige, education and income) for each respondent, and we will call this total the respondent's *social class score*. At first glance, it may appear that we are adding "apples" and "oranges". Consider our operating definition of social class once again. Using the definition as a guide, we are almost directed to weight all three elements of social class equally.

Our last task is to group the respondents into two or three groups: "High" and "Low", or "High", "Medium", and "Low", as per the method previously discussed in the "current grades" score above.

The Measurement of your Variables: A Brief Summary

The scoring and grouping of variables is often a long and tedious process, involving a number of steps or operations. First, researchers must ensure that they are entirely clear on the exact meaning which they will assign to their variables to be measured. At the same time, they should have a good idea of the range which those variables will likely assume in the population under study. Only then can they proceed to the composing of questions which

measure those variables in each of their respondents. In designing questions, they should take special care to consider that their questions are measuring what they intend them to measure (*validity*), that they do not systematically encourage or discourage certain types of responses (*bias*), and that they would reveal similar results if repeated again (reliability). *Pretesting* with the assistance of an *expert* or a *known group* is invaluable in this process. Finally, the questions should be constructed with a *grouping* and/or *scoring* plan in mind. Researchers must have a plan for each question or group of questions, which direct them in dividing the respondents into groups or categories, such as "high", "medium", and "low"; "rich" and "poor"; "self-confident" and "self-deprecating"; "Catholic", "Protestant" or "Jewish"; etc.

Here is a summary which lists a few tips on questionnaire construction and administration, to be used as quick reference:

1. Only ask questions which are directly related to the variables in your your hypotheses. Do not ask irrelevant questions.

2. Variables must be measured independently of each other. When you are developing questions for a variable, try not to think about the other variable which this variable is linked to in the hypothesis.

3. Think about what your variable means before attempting to develop questions measuring it. Define your concepts carefully and clearly.

4. Think through your questions carefully: Does each
   question actually measure what you think it will
   measure?

5. Whenever you ask a question, be sure that you have a
   good idea as to what you will do with the answers to
   it! Avoid questions whose answers will be hard to
   quantify (put into categories) later. Usually, it is
   best to provide your own answers, and give each
   respondent a series of choices.

6. Try to avoid "yes - no" answers, where a range of
   response would more accurately reflect the respondents'
   opinions on the matter.

7. Put options (answers) in logical order for your
   respondents. [eg. smallest to largest]

8. The order of your questions is important: as a
   general rule, order them from the least difficult or
   offensive to the most difficult or offensive. Ensure
   that your questionnaire has a certain logic or flow to
   it -- check with those whom you pretest it with.

9. Keep the distinction between subjective and objective
   questions in mind: sometimes it is wise to combine
   both types for a single variable.

10. Use everyday language that your respondents can
    readily understand. You will know if your language is
    unclear when you pretest your questionnaire before
    administering it to your sample.

11. Use correct punctuation, grammar, etc.

12. Be neat and organized.

13. When you explain your scoring (in your *Research Design*
    section) explain why you should score it that way!
    Scoring must be systematic -- using actual numbers or
    scores.

14. Be aware that the cut-off points (eg. between "low"
    and "high") usually cannot be decided in advance of
    collecting the data. They will often be determined
    later by trying to create even groups within your
    sample.

15. Be sure to start your questionnaire with a brief
    statement assuring anonymity, and end your

questionnaire with a thank you (see appendix 1 for a sample questionnaire). Never breach that promise of anonymity!

16. When you administer your questionnaire, be conscious of the fact that you are not just representing yourself. You are also representing your school and the discipline of social science.

17. When administering your questionnaire, never insist that a respondent answer any given question. Remember that participation is strictly voluntary.

18. When social scientists administer questionnaires, they often tell their respondents that they will "brief" them later on the results of the survey. This acts as an incentive to the respondents, and encourages future participation in similar ventures. Seriously consider doing this for your respondents.

19. You must always be the one handing out your questionnaire to the respondents. This allows you to set a reasonably serious mood, and you are there to properly answer any questions which come up.

Choosing Your Sample

You will be handing your questionnaires to a number of individuals, requesting their cooperation in filling them out. The ones who agree to do so are referred to as your *sample*. The main question which we will concern ourselves with in this section is this: How shall you choose who to approach with your questionnaire in hand? As you no doubt have come to expect by now, there is no simple, direct answer to this question. In order to help you answer this question for yourself, you must consider what you are attempting to do, when you sample.

When you make a prediction (hypothesis) about the

real world, you are secretly identifying a *population* about
which you are referring. For example, when you predict
that "rural people are more likely than city-dwellers
to. . .", you are including "rural people" and "city-
dwellers" in your population. But think carefully, once
again: do you mean all of the above two categories of
people living in Canada? The world? Your county or
province? Do you mean all age groups, or should your study
limit itself to a specific age range? Will those who have
recently moved from the city to the country be included or
excluded? Just what do you mean by the city, anyway?
Where does the city end and the rural area begin? As you
can see, you have a great deal to consider before you can
decide upon who you want to include in the population from
which you will draw your sample. One thing remains a
certainty, however: you cannot do a good job in selecting
your sample until you have seriously considered who it is
that you want to include in the population of your study.

Once you have settled the issue of population, you
are ready to select a sample from that population. The
main goal to keep in mind when selecting this sample is
that it should be "representative" of the population from
which it is drawn. This means that it should share many of
the basic characteristics of that population. For example,
it should have roughly the same proportion of males and

females, Protestants and Catholics, old and young, rich and poor, etc. as the population itself. In order to accomplish this, the researcher should attempt, as much as possible, to draw individuals *randomly* from the total population.

In random sampling, individuals within the population each have an equal chance to be part of the sample. Often this is accomplished by methods like drawing names from a hat, or numbering each member of the population and selecting numbers blindly. Let us say, for example. that you wanted to randomly draw a sample from the grade nine population of your school. In order to do this, you would have to elicit the help of the office, to supply you with a list of names including all the students registered in grade nine. You would then develop a strategy of randomly drawing from that list. You could put all their names in a hat, and draw them out until you reach your desired sample size. Or you could use a random numbers Table, to help you select your sample. Now you have your sample chosen, but you still have the problem of how to make contact with all these students, at a time and in a place where the questionnaire or interviews could be administered. Many students would respond favourably to a well-worded request to meet at a certain time and place, in order to help you out with your research. If some fail to show up, or respond to you that they want no part of your

survey, such is life for the social scientist. Draw another set of names from those left on the grade nine list, and repeat the procedure! If you were able to select a purely random sample from your population, this would allow you to generalize your findings about your sample to the population as a whole. So if, for example, your sample told you that where you live has an impact on your relationship with your neighbours, you would then be able to generalize this finding to the population (from which the sample was drawn).

Since truly random sampling is sometimes not practicable, social scientists have developed what they consider to be the next best thing: *cluster sampling*. This involves randomly selecting some smaller units from the population as a whole. Then, from this smaller unit, the researcher randomly selects a still smaller unit to interview. So, if "high school students in Town X" is your population, you may list the 10 high schools within your community, and randomly select two of them, by using the "name-in-hat" technique. Then, you may list the 8000 classes within the two schools, and randomly select 5 of them for the survey. This method of sampling closely approximates the above purely random method. It also allows the researcher to generalize his results to the population as a whole, in this case, all high school

students in Town X.

The last method of sample selection which we will consider is called *accidental sampling*. It is called "accidental" because the choice of respondents often depends upon the accident of the researcher and the respondent being in the same place at the same time. In accidental sampling, the researcher gives up on the idea of being able to generalize the results to the stated population. The researcher's goal is much more limited. The goal is confined to reporting on what was found in the sample, and the researcher cannot make any claims to have any systematic knowledge about the population. This kind of sampling technique is sometimes used in social science, but it clearly is not the preferred method. The following are examples of accidental sampling: standing outside a shopping centre, and approaching every fifth person who heads for the entrance door; standing in a cafeteria, and approaching every person in the back left-hand quarter of the room between the hours of 12:00 p.m. and 12:30 p.m. on three successive days; choosing one teacher randomly, and handing out your questionnaire to every one of his students present on a given day, with a request for cooperation in filling it out honestly. The advantages of accidental sampling are that it is cheap, quick and usually very effective in obtaining a sample.

In summary, then, you must develop a definite plan

as to how you are going to choose the people who will
potentially answer those questions that you will be so busy
developing. In order to do this properly, you must first
decide upon the *population* about which you wish to
generalize. This information can come from you thinking
about your hypotheses, and considering who you would
reasonably sample (given your time and monetary
restrictions). Be prepared to compromise on this point.
For example, none of your classmates are going to be able
to draw a sample from the total population of your town or
city (unless you are living in an extremely small
community). Next, you must decide on a method of selecting
potential respondents from that population, in as close to
a purely *random* fashion as possible.

The Research Design Section of Your Paper

The *research design* section of your paper will
explain to the reader just how you went about choosing a
sample and developing the questionnaire. Start with a
brief introduction, describing what you are
attempting to accomplish in this section of the paper.
Then explain how you *operationalized* some of your key
variables, and how you went about dividing them into
categories like "high", "medium" and "low". [Use some of
the examples on pages 68-71 as your guide.] If you

developed any *scores* for your variables, you must explain
(at least in brief) <u>how</u> you went about doing the scoring.
For each scored variable, you must give a concrete example
of how a question or series of questions were scored. You
must also include a discussion of how you chose your *sample*
to represent a larger *population*. In this discussion, you
must include a brief statement of why you decided to sample
as you did. You must also answer basic questions about
your sample: e.g. What is your sample size? What is its
age and sex composition?

Chapter 4   Organizing and Writing Up the Data

Now that you have asked your respondents to complete your questionnaire, what will you do with that rather sizable pile of paper sitting in front of you?! The most helpful answer is that you must carefully organize the responses into proper *Tables*, which then allow you to understand what your sample can communicate to you about the nature of your population. If you have completed all the steps mentioned in this book with care, your *Tables* will give you useful results, which will allow you to make judgements about the validity of your hypotheses. "Writing up" what your *Tables* tell you about your respondents is called *data analysis*. This important stage will be discussed in detail in the second part of this chapter. We will begin with a detailed explanation of how to move the answers to your questionnaires or interview schedules from the sheets into *Tables*.

A. Organizing Your Data into *Tables*

The reader may recall that we emphasized in *Chapter 3* that each variable must be measured independently of the others. The reason that we stressed this idea was that it

is only now that we can start to bring these variables together in pairs to try to determine if one has any apparent relationship to the other. The last steps which we left you with in *Chapter 3* involved the scoring and grouping of each respondent's answers to your questions. While some variables require scoring (i.e. those having questions with two or more answers which need combining), other variables can by-pass that stage, and can be grouped without the need to develop a score first. Let us first consider the situation where the variable must be scored, before it can be grouped.

Scoring and Grouping Your Variable:

Before you proceed, we recommend that you follow these instructions by starting with the independent variable of one of your hypotheses. The very first thing that you should do with a variable that must be scored (because its questions contain more than one answer) is to actually *carry out* the scoring instructions which you gave yourself to follow when you developed your questions for your variable. If your variable is not going to be scored, proceed to the instructions on page 87.

For a variable requiring scoring, you must go through each respondent's questionnaire and add up the total point score for that variable. You should write that

total score on each respondent's questionnaire in the same
convenient location, using the same colour pen or pencil.
Your next step may only be attempted when all of your
questionnaires have been answered. Put your actual
questionnaires in order, from the lowest score in the
variable score which you are concerned with, to the highest
score (or vice versa). Then, you must decide whether you
wish to divide the variable into two or three sub-groups.
The choice of two or three is up to you, and should be
based on a number of considerations. Look back at your
Theory section to see if cut-off points or groups are
suggested there. Look at how your own sample varied in
terms of score (if the variable was scored), and try to
look for "natural" cutting points between "low" and
"medium", "high" and "low", etc. Consider the *size* of your
sample: a large sample of approximately 65-75 may be able
to "handle" division into three as opposed to two groups.
Let us consider some examples of data, and discuss how we
would make our decision with respect to grouping.

Example #1

Variable: Sex

Sample

Males: 35

Females: 31
          ____
            66

Example #2

Variable: Family type

Sample

Two-parent: 24

Single parent: 12

Blended family: 14

Living apart from family: 2
                          ____
                            52

<u>Example #3</u>

<u>Variable</u>: Individualism score (The higher the score,
          the more "individualistic" the respondent)

<u>Sample</u>

  1:     4

  2:     6

  3:     5

  4:     7

  5:     9

  6:     6

 7+:   4
        ———
       41

The grouping of the first example is rather clear-cut: the researcher has no choice but to divide the data into two (as opposed to three) groups. Likely, the theory which suggested the hypothesis to the researcher is working with the assumption that the world is divided into two (not three or more) sexes.

The second example represents more of a challenge to the researcher. Again one must go back to the hypothesis and ask oneself what one is looking for in the data. If the hypothesis is comparing respondents raised in a "two-parent" family with "all others": divide the data into two groups. If the hypothesis has something to say about three or four different types of family types: divide the data into three or four groups. Notice in

advance that you do not have enough respondents (only 2!)
in the "living apart from family" group to hope to learn
something meaningful about that part of your sample. As
such, you must consider the options of either including
this group with another of your groups (this does not seem
to be a viable option in this case), or eliminating this
group of respondents from your calculations for this
particular hypothesis. As a general rule, if you have
fewer then 10 respondents in any group, you are near the
limit of an acceptable number of respondents. In such
cases, you should think about *regrouping* the data.

The third example involves considerable judgement
as well. The most obvious first choice is to divide the
data into two groups. If the researcher divides the 41
respondents into a "low" individualism group (including the
22 respondents with scores of 1-4) and the 19 "high"
individualism score respondents having scores of 5-7+),
this will successfully accomplish the above goal. What
about the option of dividing this sample into 3 groups? We
would suggest that a sample of 41 is getting toward the low
end of the acceptable range for considering such a
division. But if the researcher wanted to go ahead with
such a division, one could decide that the "low"
individualism group included the 15 respondents with scores
between 1 and 3, the "moderate" individualism group

included the 16 respondents with scores ranging from 4-5,
and the "high" individualism group is comprised of the 10
respondents with scores ranging from 6-7+.

In general, we would recommend that you limit
yourself to dividing your respondents into two groups or
categories, unless you feel strongly otherwise. It has the
advantage of simplicity, and is not unsuited to samples of
the size recommended here. Sometimes researchers have been
known to develop *Tables* using 2 and 3 groupings, and to
only present the *Table* which "looked" better, in terms of
supporting his hypothesis. This behaviour is acceptable,
as long as the cut-off points are not chosen to
deliberately distort the data.

Let us consider an example of a variable which does
not involve scoring, but must be properly grouped: "self-
image". In our example, we have decided to operationalize
"self-image" with the following question, requiring a
single response by each respondent:

Most days I feel pretty satisfied with myself. (circle one)
Strongly   Agree   Undecided   Disagree   Strongly
 Agree   Disagree

The big question is: how does one "group" answers to a
question such as the above? In a situation such as this,
you would be best to physically divide your completed
questionnaires into five groups, representing the five
different possible answers to the question. Suppose, for

example, that the piles of questionnaires ended up like
this:

Pile	# in Pile
Strongly Agree	8
Agree	13
Undecided	7
Disagree	6
Strongly Disagree	4
Total =	38

At this point, you have a couple of decisions to make:
Should I divide my respondents into 2 or 3 groups (along
the range of this dimension)? Where do I draw a line
between "low" self-esteem and "high" self-esteem (or
between "low", "medium" and "high" self-esteem)? The
simple answer to these questions is that you use your own
judgement. If your question was well formulated to measure
what you set it out to measure, you must realize that the
people who answered your question with the first response
are somehow different from those who answered the second
response, etc. But in order to be left with a manageable
number of groups (we strongly advise no more than 3), you
must combine neighbouring groups. This is to be done so
that you create two groups of respondents, with as close to
half of your total sample within each group as possible.
Another possibility is that you may choose to divide your
respondents into three groups. Do this so that you are

left with *as close to* a third of the respondents in each group as possible.

****

Student Exercise:

1. How would you handle the above example? (see pg. 87-8)

2. Why?

3. How would you handle the following example?

How much television do you watch during the average week? (circle one)

a) none

b) 1-4 hours (about 1/2 hour a day)

c) 5-9 hours (about 1 hour a day)

d) 10-17 hours (about 2 hours a day)

e) 18-24 hours (about 3 hours a day)

f) 25-31 hours (about 4 hours a day)

g) 32-38 hours (about 5 hours a day)

h) 39 or more hours (over 5 hours a day)

Pile	# in Pile
a) none	3
b) 1/2 hr.	2
c) 1 hr.	5
d) 2 hr.	10
e) 3 hr.	7
f) 4 hr.	4
g) 5 hr.	2
h) over 5 hr.	0

****

It should now be extremely easy to actually divide your respondents into groups. Take the top half of your questionnaires, and call them the "low" group (or the "high" group, or the "smart", "tall", "ambitious", "athletic", etc. group -- depending upon the exact nature of your variable). Then the remaining half of your questionnaires will be identified by you in some appropriate manner as being the opposite of the first group. Do not be surprised or disturbed if you cannot divide your sample into exactly equal groups. Just do the best you can to make the groups approximately equal. If you choose to divide your sample into three rather than two categories, select the first third, and call them "low" or "high", etc. Then take the next third, and call them something like the "medium" group, with the remaining third being called some other appropriate label. Now go back inside each questionnaire, and make sure to *identify each respondent's categorization by name*, in some convenient location. If it is a scored variable, write this categorization beside the variable score which you marked down previously.

Your next step will be to set up a blank *Table* on a separate sheet of paper, so that it looks something like *Table 1A* below. Please note that the independent variable in this example is "social class", *and that it must appear down the left hand column of the Table*!

TABLE 1A: SOCIAL CLASS AND PHYSICAL ACTIVITY

SOCIAL CLASS	PHYSICAL ACTIVITY	
	LOW	HIGH
LOWER		
UPPER		

Before you begin to actually record the data, you
must set up your blank *Table* as above, with the independent
and dependent variables in their correct locations.   You
must also make sure that you have completed the above steps
on the (scoring and) grouping of your independent and
dependent variables.   When these steps are complete, you
can begin the process of filling the empty cells of your
Table in with data.   Pick up the first questionnaire, and
determine whether the respondent has been grouped "low" or
"high" along the dimension of the independent variable of
the hypothesis with which you are concerning yourself.   At
the same time, determine how the respondent has been
grouped along the dimension of the dependent variable.   In
our example, the dependent variable is "amount of physical
activity", which must appear to the right of the
independent variable in your *Table*, as shown.   For example,
if your first respondent scored "low" in the independent
variable, and "low" in the dependent variable, mark a "1"
in the cell which is beside the "low" class heading, and

under the "low" physical activity heading. If the next respondent scored "low" in the independent variable, and "high" in the dependent variable, mark a "1" in the cell beside the "low" class heading, and under the "high" physical activity heading. Please note that if respondents fail to (properly) answer the question(s) which measure one of the variables, you will simply fail to record a ʹ1ʹ in any of the cells for those individuals. Also, there may be times when you will have to exclude certain answers, such as when a respondent answers "other" to one of your questions. This is not unusual, and should not be of concern to you unless this happens repeatedly. Just proceed through your entire pile, until your *Table* looks something like this:

TABLE 1B: SOCIAL CLASS AND PHYSICAL ACTIVITY

SOCIAL CLASS   PHYSICAL ACTIVITY

                    LOW   HIGH

LOWER   11111 11111 = 10   11111 11111 111 = 13

UPPER   11111 11111 1111 = 14   11111 111 = 8

There are two more steps before we reach our final destination: a *Table* in final form, ready to be analysed. The first step involves the elimination of the rough "1ʹs", as well as the addition of a third column, indicating the

total number of respondents (denoted by "N", which is the
first letter in "number of respondents") within each
category of the independent variable. The "N" is
determined by adding horizontally across the *Table*. The
next version of your *Table* should look like this:

TABLE 1C: SOCIAL CLASS AND PHYSICAL ACTIVITY

SOCIAL CLASS	PHYSICAL ACTIVITY		
	LOW	HIGH	N
LOWER	10	13	23
UPPER	14	8	22
			———
			45

The final version of your *Table* will be given in
percentages, in addition to raw numbers. This helps the
researcher in looking for trends within the data. This
will be discussed later in this chapter. The percentages
are calculated as follows. Each "cell" (number) is
calculated as a percent of the "N" to the right of that
number. In the above example: "10" is calculated as a
percent of "23"; "13" as a percent of "23"; "14" as a
percent of "22"; and "8" as a percent of "22". The final
Table looks like this:

TABLE 1D: SEX AND PHYSICAL ACTIVITY

SOCIAL CLASS	PHYSICAL ACTIVITY		
	LOW	HIGH	N
LOWER	43% (10)	57% (13)	23
UPPER	64% (14)	36% (8)	22
			45

```
****************************************************************
*                                                              *
* Summary of the Steps Involved in Moving the Data from the    *
*                                                              *
* Questionnaires to Tables:                                    *
*                                                              *
*                                                              *
* 1.   Some variables must be scored before they can be        *
*      grouped:  others require grouping only.                 *
*                                                              *
* 2.   Choose an independent variable to start with.           *
*                                                              *
* 3.   For variables requiring scoring:  carry out your        *
*      scoring instructions, and write the total score for     *
*      that variable on each respondent's questionnaire.       *
*                                                              *
* 4.   Put the questionnaires in order, from lowest to         *
*      highest score (or vice versa).                          *
*                                                              *
* 5.   Divide the questionnaires into 2 or 3 more or less      *
*      equal groups:  you decide which is most appropriate     *
*      for your variable and your sample.                      *
*                                                              *
* 6.   Identify each respondent as being "low" or "high",      *
*      etc. on that person's questionnaire, beside that        *
*      person's score for that variable.                       *
*                                                              *
* 7.   Follow steps 3-6 for the dependent variable related     *
*      to your chosen independent variable.                    *
*                                                              *
* 8.   Set up a blank Table, as per the above instructions,    *
*      making sure to place the independent variable and the   *
*      dependent variable in the correct places.               *
*                                                              *
* 9.   Carefully follow the instructions on how to transfer    *
*      the data from the questionnaires to your Table.  This   *
*      is extremely tedious work, well suited to computers:    *
*      for your next survey research project, consider that    *
*      option!                                                 *
*                                                              *
* 10.  If one of your variables does not need to be scored     *
*      before it is grouped, follow the instructions           *
*      immediately above.                                      *
*                                                              *
* 11.  Repeat the above steps, until you have one well         *
*      prepared Table for each of your hypotheses.             *
*                                                              *
****************************************************************
```

****

<u>Student Exercise</u>:

Using the above information, gather the following data into a properly constructed *Table*. When you have completed this task, examine the data carefully, and express what the *Table* tells you *in words*.

1. What is your sex? _____

2. What level of education do you want to achieve?

_____

ANSWERS:

<u>RESPONDENT</u>	<u>QUESTION #1</u>	<u>QUESTION #2</u>
#1	MALE	HIGH
#2	FEMALE	LOW
#3	FEMALE	LOW
#4	MALE	LOW
#5	FEMALE	HIGH
#6	FEMALE	LOW
#7	MALE	LOW
#8	MALE	LOW
#9	MALE	HIGH
#10	FEMALE	HIGH
#11	FEMALE	LOW
#12	MALE	HIGH
#13	MALE	HIGH
#14	FEMALE	LOW
#15	FEMALE	HIGH
#16	MALE	HIGH
#17	FEMALE	LOW
#18	FEMALE	HIGH
#19	MALE	HIGH
#20	MALE	LOW

****

B. <u>Analysing Your Data</u>

Data analysis is precisely what you would logically think it to be: the careful examination of your *Tables*, with the intent to explain the information gained from them

to the reader. This is the long-awaited presentation of the results of all your hard work. Each *Table* must be accompanied by words which clarify its meaning. The aim of this section of your paper is to explain to the reader just what the data say to you about the nature of the world out there. Do the data tend to support or reject my hypotheses? Do they say nothing at all conclusive one way or the other about the hypotheses? Why did the data turn out the way they did? You should realize that it is not important whether or not your initial hypotheses are confirmed, rejected, or neither confirmed nor rejected by your sample. What counts here is how you compare your results with the research and theory presented at the beginning of your paper. Below, we will explain in general terms what to include in your data analysis section, followed by a concrete example of one.

The *data analysis* section of your paper should be divided into as many sections as you have hypotheses. Each subsection should start with a clear, formal restatement of your hypothesis. You should very briefly remind the reader of the exact meaning of your hypothesis, as well as its origin in research (i.e. quickly summarize your *Theory* section). Then introduce and present the *Table* which summarizes your findings in relation to the hypothesis under consideration. Translate the numbers into words, by

pointing out to the reader general trends in the data, *as well as any items of surprise or interest hidden within the numbers*. Do not assume that the reader can look at the *Table* for him or herself, and learn anything from it. *Your* job is to explain what the data mean to you, that is, to *interpret*. It is not unusual for social scientists to quote from their respondents' answers, to capture the essence of the data. Do not insult the reader, though, by repeating each number and percent in your *Table*. The reader is quite capable of reading these without assistance. At this point, the reader has the right to expect you to indicate how your findings compare with your prediction (in terms of confirmation or rejection of your hypothesis). If the data failed to confirm your hypothesis, it is customary to review for the reader just what you think went wrong. For example, is it possible that the theory does *not* apply to your specific population? Did you make errors (eg. in wording your questions, or choosing your sample) which influenced the outcome? Readers are often impressed when the researcher is honest and generous in criticizing him or herself. If the data supported your hypotheses (even if it did so only "moderately"), you should mention this fact _modestly_ to the reader. Indicate that you take this to be a confirmation of (never "proof" of) the theory underlying your hypothesis.

It would no doubt be beneficial for you to have a concrete example of part of a data analysis section, in order to help you understand the style in which these are typically written. This would give you the advantage of being able to copy that style, when you come to do your own writing up. The following example is the work of a former student. The notations in the left margin are there to give you an indication of what the student has done correctly in his section of the paper.

Brief theory review.

It will be recalled that in the *Theory* section of this paper, it was suggested that as one moves from "lower" to "higher" status jobs in our society, there is a change in peoples' approach to their job. More specifically, we predicted that higher status job-holders are more "instrumental" in their approach to their work, when compared to holders of lower status jobs. Stated formally, *Hypothesis 1* summarizes our position.

Formal re-statement of hypo-thesis.

Hypothesis #1: The basis upon which our society establishes its system of job ranking is the valuing of instrumental patterns of action. Thus, we expect those positions generally given high status to be more likely to be filled by people who orient themselves toward their work in an instrumental manner than lower status jobs.

.

Instrumental action involves the following:

(1) The feeling that the achievement of the goal would be satisfying in and for itself;

(2) An awareness of the fact that there is a relationship between one's goal for working and the actual behaviour which one performs on the job;

(3) The feeling that one is absolutely instrumental to the process of completing one's goal -- i.e. it will not get done without one's own interference;

(4) The feeling that one must actively pursue any knowledge which may possibly help toward the attainment of one's goal; and

(5) The feeling that one's work should outweigh immediate gratification interests.

The final instrumental score was based upon the addition of each respondent's points, as allotted by the answers to five questions.

The results of our study very strongly agree with the above prediction. *Table 2* summarizes these results.

	Table 2: Prestige of Occupation by respondent's Instrumental Score.		
Table in proper form.			

PRESTIGE OF OCCUPATION	INSTRUMENTAL SCORE		N
	Low	High	
Low	65% (13)	35% (7)	20
Medium	52% (11)	48% (10)	21
High	21% (6)	79% (22)	28
			69

General trend in the data.

      The above *Table* reveals a strong relationship between the prestige of an occupation and the instrumentalness of its holder. Because these data so strongly support *Hypothesis 1*, they lend support to the theory which underlies it, as discussed above in the *Theory* section. The main point,

Reference back to *Theory* section.

as expressed by Parsons, was that societies need to encourage their most competent and hard-working members to complete the most

Always be cautious in your claims for your data.

difficult tasks. We cannot say that these data definitely confirm Parsons' ideas to be true, but it leads us to believe that he is possibly correct in his understanding of this issue.

In short, it appears that our society has

Summing up.

indeed structured things so that the most "instrumental", hard-working, concerned

individuals are indeed in possession of the
most high status jobs.

Introduction   Let us now consider our second
of the next
sub-section.   hypothesis. . .

You may have noticed that the above researcher felt
confident in asserting that "the above *Table* reveals a
strong relationship between. . ." He or she was able to
look at *Table 2*, and *see* the relationship. This is not an
obvious thing to see, and needs further elaboration here.

How to Read a *Table*

Statistically, we are allowed to infer the general
credibility which *Table 2* (see above, pg. 101) lends its
Theory, because tha data visually *appear* to run in the
anticipated direction. Looking at that *Table*: as we move
from our "low" prestige respondents to our "high" prestige
respondents, the percentage of "low" instrumental
respondents decreases from 65% to 21%. At the same time,
the percentage of "high" instrumental respondents increases
from 35% to 79%. Social scientists generally rely on
complex and sophisticated statistical tests to tell them
about the credibility of their data. We can, however,
trust our eyes and our common sense to steer us on this
issue. Compare *Table 2* with the following variations on

that *Table*:

Table 2A: Prestige of Occupation by respondent's
Instrumental Score.

| PRESTIGE OF OCCUPATION | INSTRUMENTAL SCORE | | N |
	Low	High	
Low	50% (10)	50% (10)	20
Medium	52% (11)	48% (10)	21
High	50% (14)	50% (14)	28
			69

Table 2B: Prestige of Occupation by respondent's
Instrumental Score.

| PRESTIGE OF OCCUPATION | INSTRUMENTAL SCORE | | N |
	Low	High	
Low	35% (7)	65% (13)	20
Medium	52% (11)	48% (10)	21
High	79% (22)	21% (6)	28
			69

We can easily imagine each of the three *Tables* being the
data which *we found* from our sample. In *Table 2*, the data
fell in the anticipated or predicted manner. Our
hypothesis would lead us to *expect* that as we move from
"lower" prestige occupation respondents, toward "higher"

prestige respondents, that the percentage having "high" instrumental scores would increase. And this is precisely what happened in the first sample. This allows us to find that our data *lends credibility to* our theory. *Table 2A* revealed totally *neutral* results: there is *no* relationship between the two variables. As we move from "low" prestige to "high" prestige respondents, there is absolutely *no* shift in the proportion of respondents having "high" (or "low") instrumental scores. This data may be said to lead the researcher to conclude that the theory is probably false, since it does not run in the anticipated direction. *Table 2B* reveals a relationship between the two variables, but in a manner which is *opposite* to the one that was predicted by the *Theory*. The data indicate that as we move from our "low" prestige toward our "high" prestige respondents, the percentage having "high" instrumental scores actually *decreases* significantly. If your data were to come out in this manner, you would likely conclude that the hypothesis (and the theory underlying it) is false (unless you could determine some error which you made to cause this result).

Concluding From your *Tables*

Philospohically, when we theorize something to be true, we are implicitly stating that it is true for a

variety of populations, in a variety of situations, at various periods of time. We can imagine three different test situations and outcomes, as per the following diagram, which we adapted from Stinchcombe (1968):

Theory: Variable ´A´ =====> Variable ´B´

Test Situation #1	Test Situation #2	Test Situation #3
Data found no relationship or the opposite relationship to the predicted one.	Data found the predicted relationship to be present.	Data found the predicted relationship to be present in a variety of settings & in a variety of forms.
The *Theory* is false.	The *Theory* is credible.	The *Theory* is more credible.

The best that you can do in *your* study, is to lend credibility or support to the *Theory* which underlies your data. This is due primarily to the fact that you do not have the time or resources to extenseively test and retest your hypotheses. On the other hand, your data do have the power to refute or deny the credibility of your hypotheses, since even one refutation is a denial of the general truth of your *Theory*. You should feel free to reject your hypothesis, and the *Theory* underlying it, when you are very sure of the validity of your questions, and the randomness of your sample.

As you can see above, the writing style used in the data analysis section should be "formal" and descriptive.

The researcher explains what was discovered, presenting the material as if the reader was totally unfamiliar with the data. Remember that by this stage, you will have been totally immersed in the data for many long hours. Try to give the reader some of the insights that you have gained from your experience. The description must be in neutral terms. Resist any temptation you might have to moralize about the behaviour or the beliefs of your sample. Do not comment either negatively or in a positive fashion about the behaviours or attitudes of your respondents. Lastly, you should never conclude that you have "proven" anything (least of all, your hypothesis). Be modest about the wider implications of your findings.

Chapter 5   <u>Conclusion and Introduction</u>

At this point, the only two remaining sections of your paper to be completed are your *Introduction* and *Conclusion*. Fortunately, these two sections are related to each other, as well as to all the work you have done so far. As a result, the writing should not be a very arduous task. Please keep in mind that the length of each of these sections should be in proportion to the total length of the main body of your paper. We would cautiously advise that you consider a 5-10% guideline for each of these sections. For example, if your paper is twenty pages in length, you should have an *Introduction* and a *Conclusion* of <u>about</u> one-two pages each. We would also strongly recommend that you write your *Conclusion* section before writing your *Introduction*. This helps to ensure that you introduce your paper properly. You will know by then <u>exactly</u> the nature of your results, and whether or not your data do or do not support your hypotheses.

<u>The Conclusion Section</u>

The *Concluding* section is comprised of two parts.

They normally flow smoothly from one to the other. First there is a summary of the whole topic, the process of investigation, and the major findings. Second, there is a conclusion, based on the above. Keep the above 5-10% guideline in mind when it comes to summing up your paper. Resist the strong temptation to include every single detail in your summary! Try not to supply the reader with a shopping list of points covered within your research paper. Instead, give the reader a brief recounting of how the topic was explored, including an overall restatement of the main findings of your research. You may find it useful to pretend that you are telling the story of your research project to someone who has not heard about or read it, but is interested enough to want to know something about it. This is very much like the answer you provide your friends when they ask you to tell them about the film you saw last night: answer with judgement, clarity, and an "appropriate" amount of detail.

When you have finished your brief summary of the findings of your research, the fun really begins in earnest! You are now expected to conclude something of significance from all those words and numbers. Bring your mind back to what you were thinking when you chose your topic originally. Why was it of significance to you? Did all your work since that time amount to anything of any importance to anybody? Why does it have merit or interest?

Did you learn anything about your topic which could
potentially help a specific group (or groups) of
readers?   Did you learn anything about doing research
which others could benefit from hearing?   Did you make any
mistakes which affected your results?   Your *Conclusion*
should be an attempt to put all that precedes it into some
new context or framework, which helps the reader to view it
in a new light.

We have one last recommendation.   You might end
your report with suggestions to future researchers in this
field.   This is a common practice in social science
research.   Tell the readers what areas appear to be most
promising for future research, and why that seems to be the
case.   This type of conclusion is sensible, in that it
allows you to admit an obvious fact.   Research in the
social sciences tends to raise as many questions as it
answers.   Be sensitive to this fact, and be quick to point
out any *new* questions raised by your research.

Introduction

Finally, you have come to the last section which
you will have to write for your paper!   This will be a
relatively short section.   You will not need to
spend much time writing it, since you already *know* what it
is that you have to introduce.   The *introductory* section of

your research project should <u>introduce</u> the reader to what
you are attempting to do here, with all this effort, and
all these words.

Try not to think of your reader as your "grader".
Rather, think of this person as someone whom you want to
encourage to read your report. We may conceive of three
different types of *Introductions*. First, there is the
"advertising" approach, in which the researcher tells lies
or exaggerations about what glorious facts and ideas the
reader will find inside. Second, there is the "I'm so sick
and tired of this assignment, I just wish it would go away
and die" approach, in which the researcher promises the
reader a long, boring, insignificant string of words pasted
together, with little or no purpose to them. Finally,
there is the "straightforward and matter-of-fact" approach,
which simply tells the reader what the paper will be about,
as a matter of courtesy. We recommend the last approach,
since it gives readers information that they could use to
help them decide whether or not they wish to pursue the
reading of this work. [Of course, this takes some
imagination on your part, since your reader has no choice
but to read your work!]

In the real world of social science, potential
readers will likely only make up their minds about reading
given articles after skimming the *Introduction*. They will

probably only be interested in reading about certain topics (and defined sub-topics within that topic area). So the *Introduction* has a special purpose. It is intended to give the the reader general information about what follows. There is no need for you to work endlessly in search of the "perfect" opening sentence, as writers of fiction are often alleged to do. Your writing style should be clear and concise. It should not be flowery or verbose.

Remember that your *Introduction* and *Conclusion* each have a special role to play in your project. They help integrate it into a unified whole. Your paper needs to be more than the sum of its individual parts. We have spent a good deal of time here, explaining the mechanics of how to successfully fit in to the tradition of doing "proper" social science research. Yet, when it comes right down to it, the paper must be <u>yours</u>. It is a vehicle for your own self-expression. You have some ideas which you think are important enough to research. Tell the reader why are they of importance. Develop your paper around your concern or curiosity or about some particular aspect of our social existence. Communicate your interest and enthusiasm to your reader. If you do this, it will lift the project from being another mundane attempt to please a teacher into a personal expression of concern about the world. Markers will often overlook many a technical mistake, when they sense that they are reading sincere attempts to learn

something of personal significance.

We wish you the best of luck. We know how much work you have put into this. But there is nothing that beats "learning by doing". We hope that when this is all over, you can honestly say to yourself: "Survey research, I know what that's all about!"

Appendix 1   A Sample Questionnaire

WHAT FOLLOWS IS A SERIES OF QUESTIONS, MAINLY ABOUT YOUR
EXPERIENCES AT SCHOOL. PLEASE ANSWER THEM AS BEST YOU CAN.
IF A QUESTION IS UNCLEAR TO YOU, PLEASE ASK FOR HELP. I
WOULD LIKE TO ASSURE YOU THAT ALL INFORMATION WILL BE HELD
IN STRICT CONFIDENCE: NO ONE WILL EVER HAVE ACCESS TO THE
ANSWERS, AND THEY WILL BE DESTROYED AFTER THEY ARE ANALYSED
STATISTICALLY.

1. WHAT IS YOUR SEX?   ___MALE   ___FEMALE

2. WHEN YOU ENTERED GRADE NINE, WAS YOUR ENGLISH COURSE:

    ___ENRICHED?   ___ADVANCED?   ___GENERAL?   ___BASIC?

3. WHEN YOU ENTERED GRADE NINE, WAS YOUR MATHEMATICS COURSE:

    ___ENRICHED?   ___ADVANCED?   -___GENERAL?   ___BASIC?

4. WHEN YOU ENTERED GRADE NINE, WAS YOUR SCIENCE COURSE:

    ___ENRICHED?   ___ADVANCED?   ___GENERAL?   ___BASIC?

5. WHEN DID YOU ENTER HIGH SCHOOL, THAT IS IN SEPTEMBER OF
    WHAT YEAR?
              19_____

6. TO DATE (NOT INCLUDING YOUR PRESENT CLASSES) HOW MANY
    CREDITS HAVE YOU EARNED?
              _____

7. HOW MANY DIFFERENT ELEMENTARY SCHOOLS HAVE YOU ATTENDED,
    FROM KINDERGARTEN TO GRADE 8?

              _____

8. PLEASE TRY TO HONESTLY ESTIMATE THE AVERAGE NUMBER OF
   PERIODS YOU WERE ABSENT FROM SCHOOL LAST SEMESTER:
   CHECK ONE

   ___APPROXIMATELY 1-5 DAYS (I.E. ABOUT ONCE A MONTH)

   ___APPROXIMATELY 6-10 DAYS (I.E. ABOUT ONCE EVERY 2ND
      WEEK)

   ___APPROXIMATELY 11-15 DAYS (I.E. ABOUT TWICE IN THREE
      WEEKS)

   ___APPROXIMATELY 16-20 DAYS (I.E. ABOUT ONCE A WEEK)

   ___APPROXIMATELY 21-25 DAYS (I.E. ABOUT ONCE OR TWICE A
      WEEK)

   ___APPROXIMATELY 26-30 DAYS (I.E. ALMOST TWICE A WEEK)

   ___31 OR MORE DAYS (I.E. TWICE A WEEK OR MORE)

9. WHICH OF THE FOLLOWING MOST CLOSELY DESCRIBES YOUR
   PRESENT FAMILY SITUATION?   CHECK ONE

   ___SINGLE PARENT FAMILY: I AM LIVING WITH MY MOTHER

   ___SINGLE PARENT FAMILY: I AM LIVING WITH MY FATHER

   ___TWO PARENT FAMILY: I AM LIVING WITH BOTH MY PARENTS

   ___TWO PARENT FAMILY: I AM LIVING WITH ONE PARENT AND
      HIS/HER SPOUSE OR "LIVE-IN"

   ___I AM LIVING WITH NEITHER PARENT

   ___NONE OF THE ABOVE (PLEASE EXPLAIN) _____

   _____

10. WHAT DOES YOUR MOTHER DO FOR A LIVING? (PLEASE EXPLAIN
    HER JOB, AS BEST YOU UNDERSTAND IT.)

    JOB TITLE: _____

    JOB DESCRIPTION: _____

    _____

11. WHAT DOES YOUR FATHER DO FOR A LIVING?   (PLEASE EXPLAIN
    HIS JOB, AS BEST YOU UNDERSTAND IT.)

    JOB TITLE: _____

    JOB DESCRIPTION: _____

    _____

12. DO YOU PRESENTLY WORK PART TIME?   ___YES   ___NO

13. IF "YES":

    ON THE AVERAGE, APPROXIMATELY HOW MANY HOURS PER WEEK
    DO YOU USUALLY WORK?
                 _____ HOURS

14. LAST YEAR, DID YOU WORK PART TIME?   ___YES   ___NO

15. IF YES:

    ON THE AVERAGE, APPROXIMATELY HOW MANY HOURS PER WEEK
    DID YOU USUALLY WORK?
                 _____ HOURS

16. TWO YEARS AGO, DID YOU WORK PART TIME?   ___YES   ___NO

17. IF YES:

    ON THE AVERAGE, APPROXIMATELY HOW MANY HOURS PER WEEK
    DID YOU USUALLY WORK?
                 _____ HOURS

18. THREE YEARS AGO, DID YOU WORK PART TIME?   ___YES   ___NO

19. IF YES:

    ON THE AVERAGE, APPROXIMATELY HOW MANY HOURS PER WEEK
    DID YOU USUALLY WORK?
                 _____ HOURS

20. HAVE YOU EVER BEEN INVOLVED IN OUT-OF-CLASS SCHOOL
    ACTIVITIES (EG. SPORTS, CLUBS, ETC.) WHILE YOU WERE IN
    HIGH SCHOOL?

    ___YES   ___NO

    IF "NO", PLEASE GO TO QUESTION #22

    IF "YES", PLEASE CONTINUE WITH #21

21. PLEASE COMPLETE THE FOLLOWING INFORMATION AS WELL AS
    YOU CAN:

<u>WHAT OUT-OF-CLASS (BUT IN SCHOOL)</u>   ON THE AVERAGE,
HOW MANY HOURS
<u>ACTIVITIES WERE YOU INVOLVED IN?</u>   DID EACH ACTIVITY
<u>TAKE PER WEEK?</u>

GRADE 9: _____

_____

GRADE 10: _____

_____

GRADE 11: _____

_____

GRADE 12: _____

_____

THANK YOU VERY MUCH FOR YOUR TIME AND COOPERATION. LET ME
ONCE AGAIN ASSURE YOU THAT NOBODY OTHER THAN MYSELF WILL
EVER SEE OR HEAR ABOUT YOUR PERSONAL ANSWERS. INSTEAD, THE
ANSWERS OF ALL MY RESPONDENTS WILL BE PRESENTED AS A GROUP,
WITH NO ONE PERSON IDENTIFIED.

Appendix 2   <u>A Sample Title Page and Table of Contents</u>

SOCIAL CLASS AND BEHAVIOUR

by

Jody Lemon

Pauline Johnson Collegiate Vocational School

Mr. Hamovitch

NFP OA1

January 3, 19__

TABLE OF CONTENTS

Appendix 3   <u>Citing References</u>

     It is recommended that you use the method of citing

references developed by the <u>American Psychological</u>

<u>Association</u> (APA, 1987). You will want to cite the source

of facts and ideas which you present in your paper that

originate from sources other than yourself. You should not

feel obliged to cite facts or ideas which are universally

accepted or "understood" (eg. "an apple is a fruit").

However, any idea which might conceivably be questioned or

referred to by a reader should be cited (eg. "language

usage varies directly with the social class of one's family

of origin"). As well, all direct quotations from another

source should be properly referenced.

     The APA method of citing references is designed to

achieve a minimum of interference with the reader, as he

reads the text of your paper. It involves giving your

reader the author and date of publication and page of the

source of your information, generally within parentheses.

The interested reader can then check at the back of your

paper for more details as to title, publisher, place of

publication, volume and number, etc. Please check below,

for several examples which should help you cite different

types of references, in different contexts within the

written sentence. We have borrowed heavily from the APA

manual.

In the following citations, the author is

referring to the general idea of the whole book or article

by Smith. As such, no page references are noted:

Smith (1983) compared reaction times. . .

In a recent study of reaction times, (Smith, 1983) it
was observed. . .

In 1983, Smith observed that. . .

In a recent study of reaction times, Smith (1983)
described . . .   . . .Smith also found. . .

Note that in the last entry (above) the second reference to

the same 1983 Smith publication need not be cited: the

reader will automatically assume that it was from the same

publication previously cited.

These citations refer to specific information

learned on given pages within the noted text. All direct

quotations from a source must be footnoted with this

information:

Smith (1983, p. 10) found that. . .

Smith reports on page 10 that. . .

He stated that "the placebo effect disappeared when
behaviours were studied in this manner" (Smith, 1982,
p. 276).

Smith (1982) found that "the placebo effect
disappeared when behaviours were studied in this
manner" (p. 276).

He stated, "The placebo effect disappeared when behaviours were studied in this manner" (Smith, 1982, p. 276), but he did not. . .

Smith (1982) found the following:

The placebo effect, which had been verified in previous studies, disappeared when behaviours were studied in this manner. Furthermore, the behaviours were never exhibited again, even when real drugs were administered. Earlier studies were clearly premature in attributing results to a placebo effect. (p. 276)

Please note that the last quotation is rather lengthy: as such, it is indented and single-spaced.

The following references are to multiple authors or non-individual authors (eg. associations, governments or corporations):

Multiple authors, first citation within your text:

Williams, Jones, Smith, Bradner and Torrington

(1983, p. 25) found. . .

Multiple authors, subsequent citations within your text:

Williams et al. (1983, p. 35) found. . .

Corporate author, first citation within your text:

(National Institute of Mental Health [NIMH], 1981,

p. 27)

Corporate author, subsequent citations within your text:

(NIMH, 1981, p. 29)

Reference List

At the end of your paper, you must provide the reader with the additional information which is necessary for him to identify and retrieve the publications cited within your text. The rule here is clear and unequivocal: your reference list should include only those sources that were actually cited within your paper. At the same time, each reference cited within your paper must be in your reference list (in precisely the same format). Each entry usually contains the following: author, year of publication, title, publisher name and location. We have included here a sample reference list, to act as a guide for you to follow in your paper. Please note that we have included a variety of types of references: eg. single-author book, multiple-author book, journal (magazine) article, one author, two, authors, no known author, etc. Also, please note the form of the reference page: it is double-spaced, in alphabetical order, etc.

REFERENCES

Bernstein, T. M. (1965). The careful writer: A modern guide to English usage. New York: Atheneum

Gardner, H. (1981, December). Do babies sing a universal song? Psychology Today, pp. 70-76.

Letheridge, S., & Cannon, C. R. (Eds.). (1980). Bilingual education: Teaching English as a second language. New York: Praeger.

Paivio, A. (1975). Perceptual comparisons through the mind's eye. Memory & Cognition, 3, 635-647.

Spetch, M. L., & Wilkie, D. M. (1983). Subjective shortening: A model of pigeons' memory for event duration. Journal of Experimental Psychology: Animal Behavior Process, 9, 14-30.

Strubnk, W., & White, E. B. (1979). The elements of style (3rd ed.). New York: Macmillan.

Study finds free care used more. (1982, April). Montreal Star, p. 14.

Appendix 4   <u>Glossary of Key Terms</u>

       This glossary contains definitions of important
terms used throughout the text.   We have drawn from
Chadwick (1984), Reading (1977), Stein (1969), Teevan
(1987) and Zadrozny (1959), in assembling these
definitions.   Our intent is to provide the student with
quick reference material, to aid him in making his way
through the miriad of new terms to which he is introduced
in this book.

Accidental sample: a sample chosen by the accidental
          meeting of the respondent and the researcher.

Anonymity: refers to the idea that the researcher will not
          declare the name of any participant in his study.

Associative hypothesis: see ´hypothesis, associative´.

Attitude: refers to the idea that people have learned
          predispositionss to think or act in a certain way
          toward objects,persons, ideas or situations.

Association between variables: any degree of influence or
          causation among events.

Behaviour: any kind of an action on the part of a person,
          whether observable or not.

Bias in questions: questions which are worded so that the
          respondents systematically distort their answers,
          and do not report their true position to the
          researcher.

Closed question: see ´Question, closed´.

Cluster sample: a sample which is chosen by randomly

selecting smaller groups of individuals from the population as a whole, and then randomly choosing a sample from that smaller group.

Conclusion: the final section of a research paper, which combines summary with an attempt to discuss the overall significance of the study.

Credibility of a hypothesis: refers to the idea that a hypothesis must be shown to be a logical prediction, based on certain facts and assumptions.

Cut-off point: a numerical value in a scored variable where the data divides the respondents into different sub-groups (eg. the "high" as opposed to the "low" group).

Data: facts, information or statistics which are derived by calculation and/or experimentation.

Data analysis (also "writing up" data): expressing the nature of a social entity by identifying component parts and how they relate to each other, as a result of studying the data.

Data collection: the process of systematically gathering information according to scientific principles.

Defining terms: to state or set forth the meaning of a word or phrase; to explain the nature or essential qualities of.

Dependent variable: see 'Variable, dependent'.

Generalization: a statement which claims to express a truth which is valid in a wide number of cases and situations.

Generalizability, level of (also level of abstraction): refers to the size of the class of individuals for which a generalization is claimed to be valid.

Grouping (also categorizing) data: identifying sub-groups of respondents who are similar along the dimension of a given variable.

Hypothesis (plural = hypotheses): a statement of possible relationship between variables in a form suitable for for empirical testing.

Hypothesis, associative: a hypothesis having two
       variables, one of which has a determining influence
       on the other.

Independent variable: see 'Variable, independent'.

Interview schedule: the exact questions the interviewer
       wants to ask, and the order in which they are to be
       presented.

Introduction: the first section of a research paper, in
       which the reader is informed of the general nature
       of the study.

Known expert: Someone whom the researcher considers to
       have expert knowledge of the variable he is
       attempting to measure. This assists the researcher
       who is concerned about the validity of his measure
       of a variable.

Known group: A group of individuals who are known by the
       researcher to have a certain position along the
       dimension of a given variable. This assists the
       researcher who is concerned about the validity of
       his measure of that variable.

Level of abstraction: see 'generalizability, level of'.

Level of generalizability: see 'generalizability, level
       of'.

Measurement of a variable: the process of moving a concept
       from the realm of ideas into a state which has
       quantity or dimension.

Objective question: see 'Question, objective'.

Open question: see 'Question, open'.

Operationalize: see 'measurement of a variable'.

Personality: a construct which includes a person's goals,
       attitudes, opinions, habits, moral ideas, and his
       conceptions and evaluations of himself.

Population: the total group the researcher wishes to
       study. The entire group from which the sample is
       drawn.

Pretest: the process of administering the questionnaire or
       interview schedule to individuals who will not form

part of the sample, in order to refine or modify the questionnaire.

Question, closed: a question included in a questionnaire or interview schedule that forces the respondent to choose one or more responses that the researcher provides.

Question, objective: any question whose answer does not depend on any of the opinions or prejudices of the respondents.

Question, open: a question included in a questionnaire or interview schedule that allows the respondent to answer in his or her own words.

Question, subjective: any question whose answer depends upon the respondent giving his opinion on a certain matter.

Questionnaire: a set of written questions which the respondent answers by him or herself.

Random sample: the selecting of individuals from the population so that each member of that population has the same probability of being included in the sample.

Reliability: the degree to which repeated measurements of the same variable, using the same or equivalent methods of measurement, would give the same result.

Respondent: a person who agreed to answer a questionnaire, and hence become part of the sample.

Research: diligent and systematic inquiry or investigation into a subject in order to revise facts, theories, etc.

Research design: the systematic description of the researcher's plan for gathering data.

Sample: the individuals selected for study; the results are generalized to the population they represent.

Scientific method: the method of assessing the validity of ideas about reality by a systematic study and observation, combined with the recording of observations and how they were obtained so that the resulting "facts" may be checked and modifiied by others.

Score: a value assigned by the researcher to a response to a question; the sum of individual scores by an individual.

Social class: a group who share approximately the same economic position, prestige, occupational rank, power, value orientations, etc.

Social science: any of the related sciences that systematically study and analyse the significant aspects of human social behaviour. These disciplines include sociology, political science, economics, anthropology and psychology.

Social scientist: a professional practitioner of one of the social sciences.

Support for a hypothesis: see 'credibility of a hypothesis'.

Subjective question: see 'Question, subjective'.

Survey research: a research technique that asks questions of a sample of respondents with a questionnaire or an interview.

Table: an arrangement of words and numbers, usually in organized columns, to exhibit a set of facts or relations.

Testing a hypothesis: to attempt to determine the validity or genuineness of a certain proposed relationship.

Theory: a set of interrelated hypotheses or propositions concerning some phenomenon. Most social science research is designed to test theory and the best research questions are often those derived from theory.

Theory section: the section of the research paper which describes the theoretical ideas underlying a given hypothesis or series of hypotheses.

Validity: the degree to which an operationalization actually measures the variable it is attempting to measure.

Variable: a trait which can be measured in continuous degrees, as the number of years of schooling, years

of marriage, amount of income, etc.

Variable, dependent: the variable observed by the
          researcher to change in response to change in the
          independent variable.

Variable, independent: the phenomenon that is used to
          explain the dependent variable.

Writing up data: see ´Data analysis´.

References

American Psychological Association. (1984). <u>Publication</u>
<u>manuel of the American Psychological Association, 3rd</u>
<u>edition</u>. Lancaster, Penn.: Lancaster Press.

Berger, Peter. (1963). <u>Invitation to Sociology</u>. Garden
City, N.Y.: Doubleday & Co., Inc.

Campbell, D. & Stanley, J. (19). <u>Experimental and</u>
<u>Quasi-Experimental Designs for Research</u>. Chicago:
Rand McNally College Publishing Co.

Chadwick, B., Bahr H., & Albrecht, S. (1984). <u>Social</u>
<u>Science Research Methods</u>. Englewood Cliffs, N. J.:
Prentice-Hall, Inc.

Durkheim, E. (1951). <u>Suicide</u>. New York: The Free Press

Elkin, F. & Handel, G. (1984). <u>The Child and Society, 4th</u>
<u>ed.</u> New York: Random House.

Fowler H. W. & Fowler F. G. (Eds.). (1964). <u>The Concise</u>
<u>Oxford Dictionary, 5th ed.</u> Oxford: At the Clarendon
Press.

Goode, W. J. & Hatt, P. K. (1952). <u>Methods in Social</u>
<u>Research</u>. New York: McGraw-Hill Book Company.

Ontario Ministry of Education (1987). <u>Family Studies</u>
<u>Curriculum Guideline</u>. Toronto: Government of
Ontario.

Pineo, P. & Porter, J. (1967). Occupational Prestige in
    Canada. In J. Curtis, & W. Scott, (Eds.), Social
    Stratification: Canada. Scarborough: Prentice-Hall
    of Canada Ltd., 1973.

Porter, J. (1971). The Vertical Mosaic. Toronto:
    University of Toronto Press.

Reading, H. (1977). A Dictionary of the Social Sciences.
    Boston: Routledge & Kegan Paul.

Selltiz, C. Jahoda, M., Deutsch, M., & Cook, S. W. (1966).
    Research Methods in Social Relations. New York: Holt,
    Rinehart and Winston.

Squibb, P. G. (1975). Education and Class. In H. R. Stub
    (Ed.), Educational Research (pp. 142-163). Homewood,
    Ill.: The Dorsey Press.

Stein, J. (Ed.). (1969). The Random House Dictionary of
    the English Language. New York: Random House.

Stinchcombe, Arthur L. (1968). Constructing Social
    Theories. New York: Harcourt, Brace & World, Inc.

Teevan, J. Basic Sociology. Scarborough, Ont.: Prentice-
    Hall Canada Inc.

Zadrozny, J. (1959). Dictionary of Social Science.
    Washington, D. C.: Public Affairs Press.