Management 600: Practical Research Methods for Managers

Chapters 4, 5 and 6

Chapter 4: Secondary Data Collection in Business Inquiry
Secondary data are data that have been collected by someone else prior to you realizing you needed the data, or even that you knew you would be conducting some bit of research.

The advantage of secondary data is that the data are already available and can often be obtained quickly and maybe even at relatively little expense. It will probably come in a usable form and may have been cleaned of problem responses or observations.

The downsides of secondary data are that they may not exactly suit your purposes and they may not be drawn from the population about which you really care.

Because going out and collecting primary data is difficult, time consuming and potentially very expensive, a lot of research is done using secondary data.

The information gained from secondary sources might include just about anything. Good research depends at least as much upon qualitative and descriptive information, such as descriptions of products, organizations, markets and clients, as it does upon quantitative information such as sales numbers, profits, prices and quantities. A variety of sources will likely be necessary to get all the information, both qualitative and quantitative, that is necessary for the research project.

Information gleaned from secondary sources might be used at any step in the research process. To recount, these are:
1. Problem recognition
2. Information search
3. Problem analysis
4. Alternative evaluation

That is, secondary sources can help you recognize that there is a problem, perhaps one that you had no knowledge of at all. Secondary sources can help to guide your search for still more information, especially if your secondary sources talk about the sources they used. Secondary sources are, of course, a good basis of information for analysis of the problem or opportunity of interest. Finally, secondary sources might provide insights into or offer inspiration for alternative evaluations and courses of action that you might have overlooked.
Several Considerations in Information
The book lists a number of managerial considerations relating to information. These are somewhat obvious, but they warrant some discussion.

First, before using any secondary information, you should be sure that it is appropriate to the questions to be answered. That is, the data should be timely, in that it should have been collected recently enough that it will accurately reflect the current state. How quickly conditions change can be a guide to how recent information must be. The data should be relevant to the question at hand, it should contain the information that is needed to answer the question. The data should be accurate, values given should be measured so that the values recorded are actually meaningful and reflect real conditions. Poorly measured observations may be no better than no data at all. The data should be available at a reasonable cost. Finally, the data should come from a trustworthy source.

Second, before going through the trouble and expense of collecting primary data yourself, you should be sure that no one else has collected the data already. You don’t want to re-invent the wheel if you don’t have to.

Third, there is a make or buy decision to be made with regard to information that is very similar to every other make or buy decision regarding inputs into the production process (including energy, materials, equipment, etc.). If data are proprietary, then an organization should compile its own data sources. If an organization will use the data very frequently and the data are critical to its operations then the organization should compile its own data sources to avoid any sort of hold-up situation, where they are essentially blackmailed over the data. It is worth noting that the make or buy decision for data isn’t really very different from the make or buy decision for other inputs, taking us back to the whole idea of information being one more input into the production process.

<table>
<thead>
<tr>
<th>Data should be compiled and maintained internally if:</th>
<th>Data should be obtained from external suppliers if:</th>
</tr>
</thead>
<tbody>
<tr>
<td>data are proprietary</td>
<td>collection and storage by the organization would be prohibitively expensive to collect, store and maintain</td>
</tr>
<tr>
<td>organization will use the data very frequently and the data are critical to its operations</td>
<td>the data would require extensive attention by organization staff who are better suited to other activities and in which that organization has no competitive advantage</td>
</tr>
<tr>
<td>organization staff have expertise in the matter</td>
<td>the data are only needed infrequently</td>
</tr>
<tr>
<td>the data could be sold to other organizations once it is compiled</td>
<td>the data needed are very general rather than being specific to the organization</td>
</tr>
</tbody>
</table>
| the data can be managed internally                                    | no advantage would be lost if another}
Finally, secondary data from different sources might not be comparable. Different organizations might use different definitions of variables, different measurement techniques or different sampling techniques. As such, making comparisons across different sources of secondary data should be done carefully, if at all, and the results should be viewed with great skepticism. A particularly important example is international data collected by different governments or their agencies. Different governments might collect information regarding income, consumption or unemployment in very different ways. Comparison of, say, household income in one country versus household income in another country based on data collected, processed and reported by the respective governments, may be completely meaningless. Consistent accounting and reporting standards help to assure that this is less of a problem when comparing secondary data from two companies.

Chapter 5: Problem and Proposal Development and Management
“A research proposal is a formalized bid for research… specifying the problem to be studied and the plan of action to solve the stated problem.”

As is said in the book, the research proposal does two things:
1. it specifies the problem to be addressed
2. it outlines the methods and procedures to be used

Problem identification and formulation
Getting answers is relatively easy compared to getting questions. Coming up with the right questions is often the bigger part of the battle.

The book discusses the process of working to a business research question. While this is a very amorphous process, the story told in the text originates with a concern, which might be expanded to also include a real or perceived opportunity. The concern might be seen as the initial version of an issue that might eventually be focused into a management problem and, perhaps, into a research problem.

It might be more helpful to think of the process as being one of general and continuous clarification of a concern or possible opportunity, as managers gradually come to understand the important factors. When these important factors are identified, they can become the focus of management and resource efforts and transform a concern into a management problem and then a research problem.
Yet another way to think about this is that you start with a large, conceptual, question, and try to distill it down to several small, clear, well-identified and researchable questions.

For our purposes in this course, it is worth noting that many student research proposals start out too big and, through a process of investigation, wind up being whittled down to a much smaller question that may be effectively investigated.

A Public Health Example
As an example, imagine that county public health officials are concerned about rising rates of HIV infection among teenagers in their area. This would be a managerial concern. Some research might reveal that these high infection rates are a result of intravenous drug use among teenagers, leading to the managerial problem of how to reduce intravenous drug use among teenagers. The research questions might then be to determine what policy options exist for addressing this problem and which of these options might work best for the county’s teenage population. The final research problem, how to reduce intravenous drug use among teenagers, is different from the original concern, rising rates of HIV infections among area teenagers.

The associated research proposal would probably begin by describing the rising rate of HIV infection and describe the causal relationship between intravenous drug use and HIV. It would then discuss various policy options for reducing intravenous drug use and describe how these options might be compared, perhaps based on experience elsewhere. Additional mention might be made of the possibility of reducing rates of HIV infection without addressing the more difficult problem of reducing rates of intravenous drug use through needle exchanges, and how this option’s effectiveness might be compared with the effectiveness of the other policy options described.

A Business Example

_Golly, there should be a business example of working from a concern to a management problem to a research problem here._

The text offers questions to help guide the process from concern to research problem. They indicate how difficult it can be to even start the process, let alone get to an operational research question. The questions can be divided into groups according to which of the screening steps they address.

Identifying Managerial Concerns
1. What is the current situation of the firm? Are there any undesirable situations that need attention?
2. Are there any conditions, processes, etc., that could be improved?
3. Are any problems foreseeable in the future that could affect the operations of the firm?
4. Are there any potential opportunities in the future that the organization may capitalize on?

Refining into Managerial Problems
5. Is the problem identified really a problem, or is it a symptom of another problem?
6. Does the identification of the problem follow from the available evidence?
7. Is research needed to identify the problem that underlies some undesirable situation?

Why not to do research
There’s an old saying that goes something like, “When the only tool you have is a hammer, everything starts looking like a nail.”

Just because you have research and data analysis skills doesn’t necessarily mean that every managerial decision you face will require a research proposal and an actual research project. The textbook lists these.

The research project may not be cost effective. Its expected benefits may not be greater than its expected costs.

It may not be a priority of the organization. That is, the project either might not be of central importance to the issue of the organization. In the above case of people contracting HIV through intravenous drug use, for example, an interesting question might be as to why people start using illicit drugs, but this would be beyond the mission of public health officials concerned primarily with reducing rates of HIV transfer.

The question may be unresearchable (if that’s even a word). That is, a question might be well focused and clearly defined, but for some reason might be impossible to address through research. This might be because the research would be unethical, would demand data that could not be obtained or might involve variables that could not be reliably measured.

The resources that might be devoted to the research project are otherwise occupied. That is, if the people who would do the research are occupied with other, more valuable projects. If this is the case, however, it might be asked whether outside consultants might be secured to do the job. If, however, the total amount of money available is limited, it might be better allocated elsewhere.

A good example of research not done is the racial profiling issue raised previously. The agency realized that, regardless of the results of any research, there would probably be popular support for racial sensitivity training for police officers, so the research was abandoned, the resources saved and the racial sensitivity training implemented.
Functions of the Research Proposal
A research proposal formally lays out the question to be investigated and the process through which it will be investigated. It should present the larger, underlying question (the spread of HIV among teenagers, for example), the logical link between the larger question and the smaller, more focused question to be addressed, a clear statement of the smaller question to be addressed, what data will be used to address this smaller question and the techniques that will be used in addressing that question, what results are expected and what changes in behavior might be expected as a result.

The research proposal should serve the following functions.
1. It should make it clear to the manager that the researcher understands the manager’s problem. Basically, it should make things sufficiently clear that both the manager and the researcher are on the same page.

2. It should clearly define the researcher’s deliverables, or exactly what is to be contained in the final research report.

3. It should allow a manager with some basic level of analytical skills to assess whether the proposed analytic techniques are appropriate and sufficient for addressing the question.

4. It should help the manager to understand what will be delivered and, with any luck, whether or not it will be of any value in resolving the managerial problem.

5. It should serve as a signal of the quality of the work to be done and the quality of the eventual report. That is, a researcher capable of doing high quality work should be similarly capable of constructing a good proposal whereas a researcher who is incapable of delivering good research will likely be similarly incapable of presenting a high quality proposal.

6. It should force the researcher to think carefully about what question she will address and what techniques will be used to address that question and what managerial recommendation will come from the analysis.

7. It provides the researcher with a clear road map, insuring that she won’t get lost along the way. This is important because once the research process starts it is easy to get lost in the data and the analysis and lose sight of what the original goal was. This assures (or should help to assure) the focus of the project.

8. It should establish a contract between the researcher and the manager so that it will be clear in the end whether or not the results that were promised are actually delivered. It also protects the researcher from mission creep that might lead to the manager making more and more demands on top of the initial agreement without additional compensation.
Types of Research Proposals
The textbook lists several types of research proposals. These differ in terms of the relationship between the manager and the researcher and whether they are both employed by the same organization or not. It is worth noting that the level of detail provided in the proposal is greater when the relationship between the manager and the researcher is more distant. When both are employed by the same organization (when the research is to be done in-house) only a brief proposal may be necessary, perhaps to assure that the researcher understands what the manager is looking for. When the researcher is independent, a more detailed agreement is required to clarify what is demanded and what will be provided.

To the textbook’s list, we add the following types of research proposals.

1. A Business Plan
   (a description of a stereotypical business plan here)

2. A Marketing Plan
   (a description of a stereotypical marketing plan here)

3. A Public Policy Proposal
   A policy-based project or research proposal should start with a general description of some sort of social problem and at least one proposed solution. The managerial question is whether the proposed solution will effectively address the problem, will be ineffective, or will cause more problems than it solves (the so-called law of unintended consequences). The project proposal should describe the data or information that will be used and how it will help to answer at least one question about the proposed policy. It should describe the types of analysis to be used and how the results of that analysis will be used to make a policy decision.

4. A Non-profit Management Plan
   A project proposal built on a non-profit (or not for profit) management problem should begin with a description of the situation and the problem or opportunity. Of particular importance should be a discussion of the organization’s mission or objective, how it relates to the problem and opportunity, and why it is believed that the proposed solution will be effective. As before, the project proposal should describe the data or information that will be used and how it will help to answer at least one question about the proposed policy. It should describe the types of analysis to be used and how the results of that analysis will be used to make a policy decision.

For This Course
A major assignment for this course is a proposal that is consistent with what is described here. The textbook, in Exhibit 5.4, describes the major structural components of a business research proposal. Not all of these will be necessary for the assignment in this course. You should, however, be sure to include the following:
A letter of transmittal, explaining the nature of the proposal, the underlying question and why the study should be done.

The title, which should be descriptive and appropriate.

The background description, explaining the underlying situation.

The statement of problem, explaining the problem or opportunity to which the proposal is related. This may be combined with the background description.

The research objectives, describing what the anticipated results will be and how they relate to the problem or opportunity.

The research strategy and methodology, describing the data and techniques to be used in the analysis. A data description, either of actual data or of fictitious data appropriate to the project, should be part of this.

Chapter 6: Fundamentals of Research Design

The process of research design is how it is decided how the research described in the proposal will be done. When dealing with a familiar problem, research design may be somewhat unimportant if it already well understood how a question can be best approached.

However, when an unfamiliar problem is encountered, research design is critical to avoid wasting a lot of time and driving yourself to the brink of psychosis, because once you start working on a project it is very difficult to remove yourself from the process and critically analyze how it might better be done.

So, putting time and effort into thinking about the approach to a research question, especially when looking at an unfamiliar question, is likely to be well worthwhile.

What could go wrong in research design

1. Planning errors
   (I have no clue how to describe planning errors. Do you have anything, Nancy?)

2. Collection errors
   Data must be collected, either by the person doing the research or by another person earlier in the process. Whether you collect your own data or use data collected by someone else, there are a few things to be concerned with.
First, were the variables that were collected measured accurately? Measuring a person’s height or weight is relatively straightforward, but measuring things like the market value of a house, a country’s GDP, a country’s unemployment rate or the clarity of a television picture are much more subject to measurement errors. That is, they might be measured with some inaccuracy or, in the case of GDP in different countries, may be measured in different ways in different places. If you are collecting your own data, you need to be sure that the data are measured with sufficient accuracy. If you are using someone else’s data, you need to be concerned with their accuracy.

Second, the observations (they may be people, customers, stores, cities, or other things) must be representative of the population you intend to study. The technical terms here are the sample, or the collection of people (for example) that you are looking at, and the population, the collection of all people with whom you are concerned. The sample that you analyze should be representative of the population. That is, it should look like the population that you care about. Statistically, this means that the mean values for variables of the sample should be very close to the mean values for the variables of the population. This can be assured, to some extent, through a good selection process that demonstrates random sampling, meaning that members of the sample should be chosen in such a way that all members of the population are equally likely to be included in the sample.

For example, if you hope to survey employees, but administer the survey to a sample of people who are at work on a particular day, then people who either don’t work that day or people who were sick or out of town or on vacation on that day won’t be included in the sample. Thus, the sample won’t be a random sample.

Similarly, a telephone survey won’t include people who do not have phones.

Third, and this is related to design, the variables that are collected (height, weight, test score, amount spent, etc.) should be the variables that will be needed to answer the research question and guide the managerial decision. For example, if you want to study how house prices have increased in a city, you would probably need to have data that give not only a house’s most recent sale price, but also when it was sold, its previous sale price and the date of the previous sale, and perhaps some characteristics of the house.

Accuracy of data can be at least partially assured by a good plan for the sampling to be done, or how members of the sample will be selected. The proposal should be clear about how variables will be measured and this should be followed in the research process. Accuracy can also be assured by doing a pre-sample, or taking a small sample and collecting data, then doing some preliminary analysis to make sure that the initial results seem reasonable before incurring the expense of the full-blown research project.

Data must be accurate and representative. It is difficult to correct for bad data with good analysis.
3. Analytical errors
Analytical errors occur when the analytical and statistical tools used to analyze the data are either used incorrectly or are inappropriate for the task at hand. The myriad ways in which this might occur are beyond the discussion to be presented here, but a good introduction to them might be found in the excellent and very accessible book, *How to Lie with Statistics*.

The likelihood of analytical errors can be reduced through the correct choice and clear description of analytical techniques to be used in the research. This requires a researcher who is knowledgeable about proper analytical techniques, but the manager for whom the researcher works should know some basic analytic techniques and be able to ask, if more exotic techniques are to be employed, what the justification is.

4. Reporting errors
Reporting errors occur when some importance is incorrectly applied to results of the analysis, or the results of the analysis are incorrectly described. Again, this is somewhat technical and beyond our discussion here, but a fine introduction to this concept may be found in *How to Lie with Statistics*.

A good example of a reporting and analytical error in combination (taken from *How to Lie...*) is the fact that the average temperatures in Oklahoma City and in Seattle are very similar. The conclusion is that since Seattle has a comfortable climate, Oklahoma City’s climate must be similarly comfortable, which it isn’t.

Reducing the likelihood of reporting errors is tricky and basically requires that someone who knows the situation, the data and the analytic tools employed, have a look at the report and determine whether the conclusions drawn and the recommendations made are really supported by the results of the analysis.

For example, a small difference in performance between two groups of employees might not be enough to support a large change in management’s approach to employee management.

**Issues in Research Design**
There’s a whole lot that could be talked about here. We’ll focus on the fundamentals.

First, the usual way of thinking about a research project is to consider a situation in terms of a *dependent variable* and one or more *independent or explanatory variables*. The dependent variable, as its name implies, depends on the explanatory variable. Changes in the values of the explanatory variables influence the value of the dependent variable.

For example, you may think about the price of a house as being the dependent variable and the various characteristics of that house as the explanatory variables. Put somewhat
differently, the characteristics of the house determine its market value. If a house has an
addition added or its kitchen remodeled, the value of the house will rise as a result.

As another example, you might think that the time it takes a person to run a marathon
might depend on the temperature, humidity and the wind speed. Running time is the
dependent variable and temperature, humidity and wind speed are the explanatory
variables. As the temperature rises from, say, 70 degrees Fahrenheit to 80 degrees
Fahrenheit, the amount of time it takes to run a marathon will likely rise.

However, research should be designed with some causal relationship in mind. That is,
the process that relates the explanatory variables to the dependent variable should be well
understood and supported by some theory. If this is not the case, the a relationship
between some variables might be inferred as suggesting some causality when none exists.

For example, if you were to look at the number of professional baseball games played in
the U.S. on different days of the year and the average temperature across the country on
the same days, you would likely find a correlation between baseball games and hot
weather. However, there is no causal relationship here. That is, baseball games do not
cause the weather to be hot.

As another example, imagine that you measure how hard employees work (as if this
could be measured) and employees’ levels of education. A positive correlation between
these two variables wouldn’t necessarily mean that more education makes people work
harder. It may be that people who are inherently more hard working have sought out
educational opportunities whereas those who are less hard working may have found the
prospect of additional education too daunting to pursue.

The important message here is that correlation doesn’t imply causality. If there is a
strong theoretical reason to believe that causality runs in a particular direction then you
might correctly infer causation from correlation, but you need to be very careful about
doing this.

Second, research designs can be grouped according to the degree of control the researcher
has over the subjects of the research.

If subjects are observed in their natural setting, this is, in the terms of the book, *ex post
facto* designs. This is what most of social research is. People are observed in the normal
course of their daily lives and some specific aspects of their actions are the focus of the
study. These studies are relatively convenient because all you have to do is observe
people, either directly or through some sort of survey (either your own survey or one
previously conducted) and make inferences from the data. The problem with these
studies is that they are not well controlled, meaning that there could be outside factors
that you don’t observe that are influencing people. Another problem is that some of the
important explanatory factors might depend on other explanatory factors, greatly
complicating the analysis.
To return to a previous example, an ex post facto study might look at the effect of education on income and, finding that people who have higher levels of education also earn more money conclude that more education causes higher income. However, it might also be the case that people who are smart, hard working and self-motivated tend to seek out more education and, coincidentally, these are the characteristics that would also make them more valuable to employers, so it might be these characteristics that really lead to higher income and not the higher levels of education. In an ex post facto study it is difficult to separate out the effects and determine which really causes higher income.

The alternative to an ex post facto study is an experimental study, in which the researcher exerts more power over the subjects. Whether done in the field or in a lab, this sort of study allows a researcher to randomly assign subjects to a treatment group or a control group and determine what the effect of the treatment is. In essence, this allows the researcher to control for all other factors and isolate the effect of one explanatory variable.

In the case of the education/income question, an experimental study would select a group of people and randomly choose people within that group to receive different levels of education. Some might attend only high school, some would attend college and some would get graduate degrees, and the effect of these different levels of education could be accurately determined because the decision as to whether or not a person would attend college would not have been made on the basis of their intelligence, work ethic or self-motivation.

Of course, it would be difficult to get funding for a study where you would actually prevent some people from obtaining higher education, which is why so many studies have to be done on an ex post facto basis.

**Internal and External Validity**

*Validity* might be expressed as the extent to which the conclusions of a study are to be believed. The distinction is sometimes made between *internal validity* and *external validity*.

Internal validity is the extent to which a study examines the thing that it claims to examine. For example, if you are looking at the relationship between education and income and you get some results, are those results really results about education and income, or is there something else going on? If the study was based on young adult workers, for example, they would be gaining work experience as well as education as they age, so it might be that their increased work experience instead of increased education was really driving a pattern of increased wages. It might also be the case that if the same group of young adults was examined over a period of several years during which there was significant inflation, the inflation, and not their increasing levels of education, might have been driving their wage increases. In each case, the study would
not really be measuring the effect of education on income as there would be some other factor present that could have generated the result.

External validity is the extent to which the results from the sample of people examined can be generalized to the entire population. For example, if a study examined people of a particular age group, people in a particular region of the country or people working in a specific industry, the results from that specific group might be internally valid for the group but might not be externally valid, in that they might not be replicated in samples from other age groups, other regions or other industries.

_Nancy... I was thinking about skipping over the material on Specific Designs and Online Research Issues. Specific designs might be a bit technical for the show and online research issues is a bit trendy for my taste, but that might just be the big stick up my butt talking. What do you think?_

**Managerial Considerations**

The book does a pretty good job of laying out some important considerations that managers need to consider, or at least be aware of, in research design. A few points deserve additional emphasis.

First, there are tradeoffs to be made, as with all managerial decisions. While laboratory experiments are probably more precise than ex post facto studies, they are likely to be more time consuming and expensive, as well as being unethical in many circumstances. Getting slightly less exact information more quickly and at less expense is often a better choice than getting more exact information too late and for a very high price. Further, some questions cannot be ethically studied in an experiment because it would mean risking damage to some of the participating subjects. They key phrase is evaluating the ethics of a study is that the ends cannot be used to justify the means.

Second, the research project should be designed to answer the research question. The design should be the most appropriate to the question of interest and should not necessarily be the design that has been used before (“… but that’s the way we’ve always done it!”).

**Conclusion**

Research design is a somewhat amorphous topic, and if you complete this section thinking that you don’t know exactly what it is, you’re in good company.

Basically, research design is the process by which you choose a method of answering a question. Experience in reviewing approaches others have taken to answering questions is a good tool in trying to design projects to answer your own questions.
It may be the case that a number of approaches might be taken to answering the same question and all of them could generate results of more or less equal validity.

The important thing is to carefully consider the design of any proposed research project and ask questions such as the following:

1. Would this research be ethical?
2. Would this research actually answer the question we’re interested in?
3. Is this research likely to be done in a timely fashion?
4. Will this research be cost effective?