THE ROLE FOR EMPIRICAL RESEARCH IN MANAGEMENT ACCOUNTING*

ROBERT S. KAPLAN
Harvard Business School and Carnegie-Mellon University

Abstract

The dominant research methods in management accounting are a priori reasoning, deductive analysis from well-specified models, and controlled laboratory experiments. None of these methods puts researchers in the field attempting to understand how accounting information is developed or used in actual organizations. As a consequence, the vast majority of management accounting research papers are neither informed by data nor tested on data.

A review of the literature on the process of scientific inquiry reveals that observation and description must be the starting points for scientific research. Observations are the building blocks for subsequent scientific activities, such as classification, measurement and theory building. Observations are also necessary for testing the generalizability and limits of any theory.

Cost accounting and management control procedures function in complex organizational settings. Any empirical research method would need to capture this complexity. Field research methods, including case studies, field studies and field experiments or process tracing studies, provide an opportunity to study management accounting systems in their organizational context.

Initially, case studies can provide the basis for a taxonomy of cost accounting and management control practices. Subsequently, field studies and process-tracing studies can lead to a more informed basis for modeling, theory-building, and hypothesis-formation activities, activities that today occur in the absence of data and observations. Finally, these empirical studies can be used to test the validity and limits of our theories. In a theory testing mode, empirical studies can not only test predictions on the existence of certain practices but also to confirm "how" and "why" these practices have, or have not, been implemented. Therefore, empirical research methods provide a rich, but virtually untapped, research method for the study of management accounting phenomena.

People who have spent their lives observing nature are best qualified to make hypotheses as to the principles that bring great numbers of facts together (Aristotle).

It seems to me that those sciences which are not born of experience . . . and which do not end in known experience are vain and full of error (Leonardo da Vinci).

Several years ago, a conference was held at UCLA on "Accounting in Its Organizational Context." Commenting on the lack of knowledge about how accounting functions in actual organizations, the authors of the first paper concluded:

It would appear that the information-related behavior of managers is an interesting area for accounting research. Unfortunately, the most fruitful approach — field research — has not been a popular one with accounting researchers (Birnberg et al., 1983, p. 126).

Similarly, in his summary of the conference papers and discussion, Hopwood observed:

The paucity of empirical studies of accounting in action was a major constraint on the conference deliberations. Repeatedly it had to be recognized how little was known of the accounting endeavor (Hopwood, 1983, p. 302).

* I have benefited from conversations with Chuck Christenson, who also supplied me with many valuable references for this paper. I also appreciate the comments of my colleagues at the Harvard Business School, the participants at the Stanford Summer Research Conference (July 1984), and at a seminar at the University of Texas, Austin.
The extremely limited number of studies of how accounting information is produced and used in organizations was likely surprising to the social scientists at the UCLA conference—psychologists, organizational behavior researchers, and sociologists—who have extensive experience with studying actual phenomena in organizations. Unlike these other disciplines, accounting phenomena leave an extensive audit trail in the form of frequent cost accounting and management control reports. The outcomes of decisions and actions—such as production and sales, inventory levels, costs and profits, defects and yields, productivity measures, employee-related measures and new product introductions—are also available within many corporations. Thus, collecting accounting data and the consequences of managers' actions would seem easier than the direct participant observation, intervention policies, subjective evaluation, and other qualitative research methods that frequently characterize the data collection process for many social sciences (see for example, Van Maanen, 1983). Indeed, as Hopwood observed:

Our organizational colleagues had assumed that accounting researchers would be applied social scientists, committed to exploring the intricacies of accounting in practice. They had presumed that by now behavioral and organizational enquiries would have resulted in a mass of insights into the ways in which accounting was related to the fabric of organizational life . . . . The simple fact is that accounting research has tended to isolate from accounting in practice, if not accounting practice (Hopwood, 1983, p. 302).

In two recent articles (Kaplan, 1983, 1984a), I too have expressed my concern with the lack of a field-based research strategy in management accounting. Major changes in the organization and technology of a firm's operations may be making the accounting and control systems of corporations obsolete (Kaplan, 1984b). Yet these phenomena will go unobserved and unstudied by accounting researchers unless they undertake studies in actual organizations.

A major impediment to conducting field-based accounting research has been a concern that such research would be viewed as less elegant, less scientific and more time-consuming than the analytic, empirical, laboratory and survey research currently being done by accounting academics. Hopwood observed, about field-based research:

... all such investigations require the very commitment that is missing in accounting research to-date— a commitment to study, analyze and interpret accounting in the contexts in which it operates . . . . To-date the difficulties of access and the inexperience of accounting researchers have been important factors. More significant, however, may be those pressures of the academic reward structure which push for speedy and voluminous research, the very low legitimacy which somewhat paradoxically has been given to grounded empirical inquiry, and the marginal commitment which many academic accountants have to the research endeavor (Hopwood, 1983, p. 303).

In this paper, I will explore the legitimacy of empirical studies in management accounting and attempt to explicate the role they must play in developing a science of management accounting. The next section will summarize the past six decades of management accounting literature to document the scarcity of empirical research in this field. Thereafter consideration is given to the conduct of scientific research, especially the use of observation for studying organizational and management processes. On these bases, the vital role that empirical studies must play for developing theories in the social sciences is then discussed. The paper concludes with a discussion of some specific issues that arise when performing empirical studies in management accounting. While the objective of this paper is to explore the role for empirical research in management accounting, the message may also have relevance for other business school disciplines, including operations research, MIS, marketing, production and microeconomics. The detailed analysis for these disciplines, however, is left to specialists in the respective areas [see, for example, Teece & Winter (1984) for microeconomics, Graham (1984) for operations research, and Bonoma & Wong (1985) for marketing].
Klemstine & Maher (1983) classified management accounting papers that have appeared in the academic accounting journals, primarily The Accounting Review, Journal of Accounting Research, Abacus, Accounting, Organizations and Society, and AAA research studies. Of particular interest is their classification of papers by method of analysis that I have modified slightly for the purposes of this paper:

(1) A priori – conceptual reasoning without an explicit model manipulation, experiment, or empirical analysis.

(2) Modeling/Simulation.

(3) Laboratory/Field Experiment – an experiment with students or practitioners on artificial data or in an artificial setting.

(4) Survey – collection of data from actual organizations, not from direct observation; analysis performed in some studies.

(5) Personal Observation – description of current practices in actual organizations.

(6) Empirical – statistical analysis of data obtained from actual organizations.

(7) Field Study – analysis of impact on accounting information on decisions, actions, or attitudes of managers in organizations.

A summary of management accounting literature by topical category and method of analysis appears in Table 1. A further summary, by method of analysis, is presented in Table 2. These tables provide a striking picture. Seven out of eight (87%) management accounting research papers have neither data from actual organizations nor are tested in actual organizations. Of the remaining papers, 5% (32) are summary descriptions of industry or country practice based on personal observation; two-

<table>
<thead>
<tr>
<th>Topical Category</th>
<th>A priori</th>
<th>Modeling/Simulation</th>
<th>Laboratory or Field Experiments</th>
<th>Survey</th>
<th>Personal Observation</th>
<th>Empirical</th>
<th>Field Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Fundamental concepts</td>
<td>53</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.2 Information economics</td>
<td>5</td>
<td>16</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.3 Firm and industry practices</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.4 Human resource accounting</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5 International comparisons</td>
<td>14</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.6 Systems</td>
<td>14</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.7 Economics of internal organization</td>
<td>9</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.8 Miscellaneous</td>
<td>16</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1 Direct/absorption costing</td>
<td>28</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.2 Distribution costs</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.3 General product costs</td>
<td>19</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Cost allocation</td>
<td>14</td>
<td>30</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Cost estimation</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.1 Cost-volume-profit</td>
<td>16</td>
<td>123</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.2 Capital budgeting</td>
<td>22</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.3 Decision making—behavioral</td>
<td>22</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.4 Decision making—general</td>
<td>23</td>
<td>123</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.1 Budgeting</td>
<td>22</td>
<td>4</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.2 Variance analysis and investigation</td>
<td>16</td>
<td>28</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.3 Transfer pricing</td>
<td>15</td>
<td>7</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.4 Principal—agent</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.5 Planning and Control—general</td>
<td>31</td>
<td>24</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>325</td>
<td>176</td>
<td>51</td>
<td>17</td>
<td>32</td>
<td>10</td>
<td>31</td>
</tr>
</tbody>
</table>
TABLE 2. Summary of research methods in management accounting

<table>
<thead>
<tr>
<th>Method</th>
<th>Number</th>
<th>Percentage</th>
<th>Cumulative Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior A</td>
<td>325</td>
<td>51.5%</td>
<td>52%</td>
</tr>
<tr>
<td>Modeling/Simulation</td>
<td>176</td>
<td>27.8%</td>
<td>79%</td>
</tr>
<tr>
<td>Lab/Field Experiments</td>
<td>51</td>
<td>7.6%</td>
<td>87%</td>
</tr>
<tr>
<td>Survey</td>
<td>17</td>
<td>2.6%</td>
<td>90%</td>
</tr>
<tr>
<td>Personal Observation</td>
<td>32</td>
<td>5.1%</td>
<td>95%</td>
</tr>
<tr>
<td>Empirical</td>
<td>10</td>
<td>0.8%</td>
<td>96%</td>
</tr>
<tr>
<td>Field study</td>
<td>31</td>
<td>4.5%</td>
<td>100%</td>
</tr>
<tr>
<td>Total</td>
<td>642</td>
<td>100.0%</td>
<td>100%</td>
</tr>
</tbody>
</table>

thirds of these appeared before 1950, only one since 1970. Only 31 (less than 5%) papers are actual field studies and these tend to be concentrated in two areas: budgeting (12) and miscellaneous performance evaluation (11); no cost accounting or cost allocation field studies exist.

This analysis of the incidence of field studies in academic accounting journals understates somewhat the actual empirical base for teaching and research in management accounting. Many studies of actual practices have been conducted and published by professional organizations such as the National Association of Accountants (NAA), the Conference Board and the Financial Executives Institute. NAA, in particular, has attempted to document existing practice in cost accounting and many of the studies it sponsored in the 1950s undoubtedly influenced textbook writing from the 1960s to today. The knowledge of management accounting practice for many academic accountants, particularly those of us who never served as cost analysts, cost accountants, or controllers in corporations, derives from the ideas contained in these textbooks. Thus, contemporary academic teaching and research is likely based on the summary descriptions of actual practice that were documented in the NAA studies of the 1950s. It is interesting, however, that few management accounting research papers reference any of these studies indicating that the influence of the studies has been filtered through such secondary sources as cost and management accounting textbooks.

In the social sciences, as in all science, the progress of our understanding is paced by the progress in our accumulation of a reliable and systematic body of factual knowledge. . . . Casual knowledge (does not) provide a satisfactory empirical foundation for general descriptive laws. Without systematic observation, including experimentation where that is possible, our samples will be badly biased, our observations will be severely filtered by our preconceptions, and the phenomena will be altogether too tangled and complex for satisfactory analysis.

An important part of the history of the social sciences over the past 100 years, and of their prospects for the future, can be written in terms of advances in the tools for empirical observation and in the growing bodies of data produced by those tools (H. A. Simon, 1980, p. 72).

Unfortunately, management accounting cannot be included, with other social sciences, as having accumulated a reliable and systematic body of factual knowledge. Nor has it advanced its tools for acquiring empirical observations.

Is it possible that we can develop a systematic theory of management accounting without observation and measurement? Are there precedents or examples of fields where normative models and deductive reasoning are sufficient to develop a substantive body of knowledge? To answer such questions, we examine authoritative writings on scientific research to glean common aspects that are relevant to our topic. We are not attempting to critique these views or comment on their strengths and weaknesses. They are presented to illustrate what thoughtful

1 I am grateful to Germain Boer for suggesting to me the importance of these field surveys on academic thought.
observers believe is essential for scientific advances.

MODELS OF SCIENTIFIC INQUIRY

Many scholars have attempted to describe and to formalize the process of scientific inquiry. Some of these have been motivated primarily by research in the physical and natural sciences; others have attempted extensions to research in the social and managerial sciences. While differences among these various approaches may be obvious and profound to experts in the history and philosophy of science, a relatively uninformed, lay reader, such as myself, is struck more by the agreement on the basic scientific inquiry model than by the differences. In particular, the most striking finding is the contrast between the consentient view of scientific research and the actual research methods used in management accounting (and several other management and management-related disciplines). In this section, I will attempt to summarize the views offered on the scientific research process and contrast these views with the research record in management accounting presented in Tables 1 and 2.

No serious student of science is so presumptuous as to postulate the method for conducting scientific research. One can read in vain attempting to find a nicely ordered set of sequential steps that a researcher or an entire discipline should follow to perform legitimate scientific research. Nevertheless, at the risk of attempting what much wiser observers assiduously avoid, I believe there is general agreement on a broad set of activities that are necessary for a field to be characterized as scientific. I will attempt to identify and describe these activities, presenting them in a precedence ordering that can be easily misinterpreted. The precedence ordering derives from a rough view that activities described first must precede the initiation of subsequent activities. The ordering will be misunderstood if readers believe that the process is sequential, moving seriatim from one stage to the next. In fact, the process is characterized by continuous feedback and iteration with research findings at later stages generating a need for new research activities at earlier stages. Thus, the research process should be viewed as iterative, dynamic, even disorderly.

I will use a formulation of the research process developed by Roethlisberger (1977). I have chosen Roethlisberger's organizing framework because, unlike perhaps better known scholars of the process of scientific inquiry, such as Popper, Kuhn, A. Kaplan, or J. L. Simon, who drew their experiences and examples from the natural and social sciences, Roethlisberger developed his thoughts from a professional lifetime as a clinical researcher of management processes. Roethlisberger formulated the process of scientific inquiry as a "Knowledge Enterprise"; (see Fig. 1).

Skill

The Knowledge Enterprise starts with skill; the way by which man begins to improve his relation to his surroundings. Skillful practitioners, whether in management, medicine, or other professions, are not explicitly concerned with describing phenomena. Rather, they develop their skills to understand the phenomena with which they are dealing and thereby learn to manipulate them to their own advantage. Schön (1984) agrees with Roethlisberger that the study of skilled practitioners is of critical importance:

We should be turning the puzzle of professional knowledge on its head, not seeking only to build up a science applicable to practice but also to reflect on the reflection-in-action already embedded in competent practice.

Description

Given the existence of skilled practitioners, management scientists can enter and attempt to develop systematic knowledge of the phenomena and the means by which skilled practitioners are coping. Such activities move us to the clinical knowledge stage of the Knowledge Enterprise, the stage which represents the start of scientific inquiry. Clinical research aims at description and its dominant research methods are observation and interviewing. At
this stage, definitions are non-operational, measurements are not objective, and hypotheses are actually questions of the form, "What is going on here?"

J. L. Simon (1978, pp. 44–46) concurs that the beginning of scientific inquiry in all disciplines is description:

In the beginning, there is description. When one does not know anything at all about a problem, one must understand it in a general way before beginning to make specific inquiries about specific aspects of the subject. For example, the early explorer in a new land writes a general description of the appearance of the country, its geography, climate, people, flora, fauna, and much else . . . . The early explorer chooses to describe what he thinks to be important and interesting, without any rigid rules of scientific evidence.

Descriptive research in the form of case studies is usually the jumping off point for the study of new areas in the social sciences. As Freud put it, "The true beginning of scientific activity consists . . . in describing phenomena and (only) then in proceeding to group, classify and correlate them."

. . . a descriptive study does not have a set of clearly delineated dependent and independent variables. The absence of a limited number of well-defined variables distinguishes case-study descriptive research from other types of research.

Students should not automatically shy away from descriptive research. Professors often tout students off descriptive projects, however, because they are harder to do well and easier to do atrociously than are other types of research. Descriptive research does not reveal sloppy and brainless work as glaringly as do more "rigorous" types of research. For this reason, other types of research usually make better training exercises than does descriptive research.

Path-breaking descriptive research, such as that of Freud, is especially difficult because one starts with empty hands - no guideposts, no standards, no

---

**Fig. 1. The knowledge enterprise (Roethlisberger, 1977, p. 393). (For the development of knowledge, read from the bottom up; for the practice of knowledge, read from the top down.)**

<table>
<thead>
<tr>
<th>Levels</th>
<th>Characteristic statements (theories)</th>
<th>Methods</th>
<th>Products</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>General propositions</td>
<td>Creative and inductive leap of imagination</td>
<td>Deductive systems</td>
</tr>
<tr>
<td>Analytical (scientific)</td>
<td>Empirical propositions</td>
<td>Operational definitions, rigorous measurement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Elementary concepts</td>
<td>Definition of concepts and variables, elementary measurement</td>
<td></td>
</tr>
<tr>
<td>Clincial</td>
<td>Conceptual schemes</td>
<td>Observation and interviewing</td>
<td>Descriptive cases and syndromes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classification</td>
<td>Taxonomies</td>
</tr>
<tr>
<td>Skill</td>
<td>Knowledge of acquaintance</td>
<td>Practice and reflection</td>
<td>How-to-do-it statements and aphorisms</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The phenomena</td>
<td></td>
</tr>
</tbody>
</table>


yardsticks, no intellectual framework, no categories within which to classify what one sees. (The researcher) must create his own classification and his own guideposts. He must decide what to look at and what to ignore, what to record and what not to record, which clues to follow up and which to drop, what is important and what is valueless. The early descriptive researcher has great freedom, but such great freedom can be terrifying.

As Simon notes, in the above quote, one of the major problems with advocating greater attention to clinical, descriptive research is that this activity is frequently afforded lesser status and value because it may not require the use of mathematics and inferential statistics. Such a danger was recognized well by Henderson (1970, pp. 108-109):

Young men whose bent is intellectual often feel that the only studies worthy of their attention are those like mathematics and physics that deal with high abstractions, or those like philosophy that deal with noble sentiments, or those like politics that deal with great affairs. Such attitudes have no logico-experimental foundation.

Now the study of those features of the interactions between persons that occur very widely or generally in human affairs must evidently be concerned with very common or vulgar phenomena . . . . But obvious things are often important, familiarity is often a cause for overlooking things, nothing, as I have said, is trivial but feeling makes it so, common things are often the uniformities, and in order to make a successful analysis or synthesis it is necessary to take account of the necessary factors and to do so in the necessary way . . . . The fact that somebody thinks it vulgar or trivial or obvious is irrelevant, and unless he is a qualified investigator so also is his opinion that it is unimportant.

We have already noted (in the quote from H. A. Simon in the introductory section of this paper) the requirement that for scientific research, observation must be careful and systematic. This point is also emphasized by A. Kaplan (1964, p. 126):

Scientific observation is deliberate search, carried out with care and forethought, as contrasted with the casual and largely passive perceptions of every day life.

Classification

One very important outcome from careful observation and description is a conceptual or classification scheme for the phenomena. J. L. Simon (1978, pp. 46-49) defines well the classification process:

Classification is the process of sorting out a collection of people or objects and of developing a set of categories among which you divide the collection. No sooner does the scientist see several different examples of a given phenomenon than she begins to say, "This one is like that one and both are different from the other bunch." Then she coins common names for those examples that are like one another. The sorting out may come first and the construction of categories (called "taxonomy") afterward, or the order can be reversed (a priori classification) ....

Classification research is different from other types of research in that one does not usually go out and collect new data for a classification study. Rather one is likely to work with existing data, sorting it into a classification that makes sense of it. Therefore, classification research tends to follow after descriptive research in the sequence of scientific stages ....

... the classification scheme itself can be viewed as one massive variable or as a set of variables . . . . But the variable or variables are not "dependent" or "independent," at least until employed in other research.

A. Kaplan (1964, p. 50) emphasized the important role of the classification stage in developing scientific theory:

The important terms of science - semantics - give science its subject matter; they conceptualize phenomena. Things studied are classified and analyzed. Things are grouped together because they resemble one another; a natural grouping is one which allows the discovery of many more, and more important, resemblances than those originally recognized . . . . Every taxonomy is a provisional and implicit theory; as knowledge grows, concepts change; as concepts fit better, our knowledge grows.

One need only recall Darwin's activities during his travels around the world to understand the vital role that classification research can play in developing scientific theory.

An observant reader may have noticed that we discussed "Classification" after we discussed "Description" whereas, in Roethlisberger's Knowledge Enterprise, "Classification" appears below the "Description" stage of "Observation
and Interviewing.” In part, this reversal is due to the difficulty of establishing a precedence relation for two activities which continually feedback and cycle between each other. That is, careful description is invaluable for modifying a previously developed classification scheme, while a classification scheme is necessary for systematic observation. We can, however, distinguish between two types of Description activities, one which would typically occur before Classification and the other afterwards. The first type consists of the highly exploratory, almost aimless wanderings described by J. L. Simon by which the researcher becomes familiar with existing practice. The researcher is learning the rules of the game and how various people are performing in the game. It would be difficult to be careful, systematic and comprehensive at this stage since the researcher has little guidance on what to look for or what to record.

After obtaining familiarity with contemporary practice, the researcher can develop a classification or taxonomy of what appear to be the critical dimensions in the phenomena. Based on this taxonomy, the researcher can then undertake more careful and systematic observation. The classification scheme guides the researcher about what to look for and how to organize the observations found in the field. Thus, for fields in which some familiarity with the fundamental phenomena already exists, the classification scheme precedes, and provides the basis for, systematic observation and interviewing. This precedence is undoubtedly what Roethlisberger had in mind when he portrayed Classification occurring at a lower level of the Knowledge Enterprise than Description. ²

**Measurement**

The next stage up the Knowledge Enterprise and the start of analytic knowledge occurs when the researcher becomes able to measure one or more aspects of the phenomenon being studied:

Measurement research seeks to establish the size of a phenomenon on one or more of its dimensions. . . . Measurement differs from case-study description in these ways: Measurement research focuses on one or a few dimensions, and measures them systematically and in relatively great detail; case study gathers information on many dimensions of the phenomenon, with or without numerical description, and in a more ad hoc fashion. Deciding what to measure and how to draw the definitional boundary lines around the quantity to be measured – translating the theoretical (hypothetical) concept into empirical terms – is a crucial decision in measurement research (J. L. Simon, 1978, p. 49).

A. Kaplan (1964, pp. 173–175) also comments on the desirability of measurement in social science research:

Systematic study can be carried on in the social sciences as elsewhere by many devices which are less precise than strict quantitative measurement but nonetheless far better than unaided individual judgment . . . . There is a direct line of logical continuity from qualitative classification to the most rigorous forms of measurement, by way of intermediate devices of systematic ratings, ranking scales, multidimensional classifications, typologies, and simple quantitative indices.

Nevertheless, when measurement is possible, it is a device for standardization, by which we are assured of equivalences among objects of diverse origin . . . . Also, it makes possible more subtle discriminations and corresponding more precise descriptions . . . . When descriptions give way to measurement, calculation replaces debate . . . . More important, measurement makes it possible to apply to inquiry available mathematical techniques for verification, prediction or explanation (Kaplan, 1964, pp. 173–175).

**Discovering relationships**

Our first set of scientific activities – observation, description, classification and measurement – attempt to reveal what phenomena are related to one another. We start to examine whether there is association between phenomena:

An investigation into whether there is a relationship between two occurrences or variables is an attempt to find out whether two (or more) phenomena are part of the same scheme of things, that is, whether they are closely associated with each other in nature’s cobweb . . . . If two particles are entrapped close to each other in a cobweb,
and if one of them moves, the other will move in close agreement with it. But if the particles are much farther from each other, movement in one will not be as closely accompanied by movement in the other. Furthermore...

movement in particle A and in particle B can be related even if neither A nor B but rather C initiates the motion (J. L. Simon, 1978, p. 53).

While the most desirable state of affairs is to be able to infer cause and effect, relationships between or among variables can still be useful even without being able to determine causality. If the relationship is strong and consistent, we can use one variable (or phenomenon) to predict the occurrence of the second variable or phenomenon. Second, we can use one phenomenon to serve as a proxy measurement of the other, perhaps costlier to measure, phenomenon. Finally, we may be interested in the relationship between two proxy variables, such as the correlation between alternative stock market indices, or scores on different aptitude tests.

Vancil (1978) provides an interesting description of how a researcher with twenty years of clinical experience in management control issues moved one stage up the "Knowledge Enterprise" to obtain analytic knowledge (measurement and association) through an exploratory empirical study:

I did not delineate a set of hypotheses to be tested by the data. Rather, I viewed my task as one of trying to invent new instruments which might measure some of the important dimensions that account for some of the differences in the ways that decentralized firms are managed..... I was seeking to discover relationships that I could rationalize and accept, and seeking to meld the set of such relationships into a cohesive pattern. Data reduction was almost mandatory, and I tried a variety of methods to create summary variables for the elementary concepts of authority, responsibility, and autonomy (Vancil, 1978, p. 138.).

In summary, analytical knowledge starts with more precise definitions of concepts and variables. Then new methods for measuring these concepts are developed and, when successful, the researcher can attempt to discover correlations among variables or phenomena. Association can help us to discover certain uniformities in the phenomena, but the findings may be confounded with the circumstances under which they were obtained. Therefore, they cannot be generalized to other situations where the conditions may be different. For scientific explanation, we need to move toward explaining how variable x varies with variable y under certain given conditions.

This next stage, of producing empirical propositions, attempts to discover and verify hypotheses. Empirical propositions require operational definitions, rigorous measurement, experimental designs and sophisticated statistical procedures. Such propositions can be developed in the natural sciences and occasionally in the social sciences when investigating phenomena in which data are relatively easy to collect and analyze. The extensive research on capital market behaviour comes to mind as an excellent example of where we have been able to formulate and test empirical propositions. This development has been possible because of the ready and inexpensive availability of data bases on the information used by capital market participants as well as information on the exchange prices in these markets. For much research on social and managerial processes, however, such information and cardinal measurements may not be available. Consequently, formulating and testing empirical propositions in these fields could be extremely difficult. For such fields, our level of knowledge may therefore be restricted to knowledge about association and correlation; we may not be able to reach the Empirical and General Proposition levels of the Knowledge Enterprise.

Theory building: general propositions

The final stage of the Knowledge Enterprise, and one that may be exceptionally hard to achieve in the social/managerial sciences, is to obtain general propositions that explain the empirical propositions discovered and validated at the previous stage. The goal is to show how a variety of empirical generalizations follow logically from a small number of general propositions under certain conditions.

A. Kaplan (1964, p. 302) indicates that a theory is necessary
... to make sense of what would otherwise be inscrutable or unmeaning empirical findings. A theory is more than a synopsis of the moves that have been played in the game of nature; it also sets forth some idea of the rules of the game, by which the moves become intelligible.

Henderson (1970, pp. 67–68) drew upon his medical background to describe the intimate interaction between observation and theory formation:

... both theory and practice are necessary conditions of understanding, and the method of Hippocrates is the only method that has ever succeeded widely and generally. The first element of that method is hard, persistent, intelligent, responsible, unremitting labor in the sick room, not in the library (emphasis added) ... The second element of that method is accurate observation of things and events, selection, guided by judgment born of familiarity and experience, of the salient and the recurrent phenomena, and their classification and methodical exploitation. The third element of that method is the judicious construction of a theory ... and the use thereof. All this may be summed up in a word: the physician must have, first, intimate, habitual, intuitive familiarity with things, secondly systematic knowledge of things; and thirdly, an effective way of thinking about things. His intuitive familiarity must embrace his systematic knowledge and his way of thinking as well as the things he studies .... Competent men of science are no more logico-experimental than other men when they step outside the field in which they have acquired intuitive familiarity with things.

There is general agreement that theory formation is much more than a synopsis or simple extrapolation of empirical propositions. Christenson (1976, pp. 640–641) critiques the views of some empiricists who believe that theoretical knowledge can accumulate with experience and observation:

The picture suggested by this point of view is that of theoretical knowledge accumulating with experience. Like a drop of water added to a bucket, each new observation statement makes its own individual contribution to the sum total of theoretical knowledge.

Christenson argues, from the logico-experimental framework, that there can be:

... no generalization without theory. Theory is like the bucket, rather than like the water in the bucket. A person has to have a bucket before he can collect the water, and he has to have a theory – a framework for measurement – before he can collect "the facts." No matter how much water he collects, it will never accumulate to a bucket; no matter how many facts he collects, they will never accumulate to a theory.

In these remarks Christenson is emphasizing that we cannot wander around aimlessly looking for observations in the world. Even in the early stages of research, we must be guided by some conceptual or classification scheme. Thus, Darwin, with his theological background, was guided by the biblical model of creation of animal species. Initial studies in management accounting would be influenced by current classification schemes such as fixed vs variable costs, responsibility, controlability, traceability, decision influencing, risk sharing and incentive compatibility. Scientists in collecting observations will, implicitly or explicitly, subject their prior conceptual schemes to continual tests. When attempting to understand these observations, the scientist must, with a creative leap of imagination, construct a theory (the bucket) to organize and explain "the facts".

A. Kaplan (1964, pp. 308–309) has described this process:

Arriving at workable theories calls for the exercise of creative imagination, as has been emphasized by countless working scientists ... We may perhaps speak of "discovering" laws, but theories must be said to be "invented" or "constructed" ... The formation of a theory is not just the facts, of organizing and representing them ... A theory must somehow fit God's world, but in an important sense it creates a world of its own.

Popper (1959) has also emphasized the creative process of theory formation:

Significant advances in knowledge occur when the inquirer goes beyond data; performs a conceptual leap of imagination to consider analogies, metaphors, models, myths as a way to explain the data. Such leaps are called conjectures. Conjectures are distinguished from myths in that tests can be devised for conjectures that could potentially falsify them. Conjectures gain in scientific stature the more they survive such tests.

Stephen Gould, in an essay in The Panda's
Thumb, provided a vivid illustration of the importance of analogy and metaphor when constructing theories. It is well known that Darwin delayed for many years publishing his theory of natural selection and survival of the fittest. Only with the emergence of a rival, Wallace, was Darwin motivated to overcome his caution to produce *The Origin of Species*. Gould recounts that an important factor in Darwin's decision to finally go public was his reading of Adam Smith's *The Wealth of Nations*. The mechanism of the invisible hand in market economies and the optimality of an economically competitive system provided a close analogy to natural selection and competition in biology, and this correspondence increased Darwin's confidence in his own theory. This is an interesting instance of how a social science (economics) may actually have influenced the development of a natural science. (The anecdote may also serve to remind us that narrow training in a discipline does not encourage creative thinking and bold theory-building activities.)

Kuhn (1970a) agrees with Popper that theories must be invented to explain facts:

Sir Karl (Popper) and I are united in opposition to a number of classical positivism's most characteristic theses. We both emphasize, for example, the intimate and inevitable entanglement of scientific observation with scientific theory; and we both insist that scientists may properly aim to invent theories that explain observed phenomena and that do so in terms of real objects . . . . Neither Sir Karl nor I is an inductivist. We do not believe that there are rules for inducing correct theories from facts, or even that theories, correct or incorrect, are induced at all. Instead we view them as imaginary posits, invented in one piece for application to nature.

Theory development, as characterized by A. Kaplan, Henderson, Christenson, Popper and Kuhn seems notably absent in management accounting. Upon reflection, this is not a surprising finding since all of the above authors stress the role of theory as a unifying framework to organize and explain observations. With virtually no observations (recall Tables 1 and 2), there has been little demand for an organizing, explanatory, refutable theory. Following Christenson's metaphor, without water we don't need a bucket. Yet we do observe that the great majority of management accounting research articles are "theoretical", reasoning logically or deductively from a set of premises, stated either explicitly or implied by the author.

**Theory and mathematics**

Many research papers in management accounting (and other disciplines such as economics, operations research, finance, and operations management) use mathematics extensively. Such papers are frequently referred to as "theoretical," perhaps to distinguish them from empirical papers which analyze data. While it is true that many theories are formulated and developed using mathematics, it is not true that every paper using mathematics and proving theorems is theoretical. This point has been made recently by Michael Jensen when indicating his reservations about mathematical representations of organizational phenomena:

Attempts to use (mathematics) at . . . an early stage in the development of an area are often counterproductive because authors are led to assume the problem away or to define sterile "toy" problems that are mathematically tractable . . . . Mathematical is not the same as rigorous nor is it the same as analytical or theoretical (Jensen, 1983, pp. 335–336).

Mathematical models can only be considered theoretical when they are representing real phenomena. Models based on assumptions made for analytic convenience, rather than because they correspond closely to real-world phenomena, may be fine mathematics but they are not theoretical because they do not help to explain observations or empirical phenomena. This distinction between mathematical models and theory may not be well understood; yet the distinction is extremely important. A. Kaplan deals extensively with this difference, especially the shortcomings of model building that is uninformed from observation. Rather than intersperse his observations here, however, I have excerpted the relevant sections in an Appendix to this paper and the interested reader can pick up the arguments there.

For our purposes, we can conclude with J. L.
Simon's observation (1978, p. 217) on the value and limitations of deductive reasoning from well-specified and well-grounded models:

If the premises are sound, deductive reasoning can be more accurate than empirical research. But often a scintillating chain of reasoning can be founded on incorrect premises.

**Theory testing and falsification**

Extensive literature exists on theory testing for the natural and social sciences. The writings of philosophers such as Popper, Kuhn, A. Kaplan and Feyerabend are centrally concerned with issues of falsification, normal science testing, crucial or nomological testing, and boundary testing (designed to identify the range of application of the posited theory). Given the lack of theory in management accounting, much less the lack of competing theories, a discussion on how to test theories in this field seems premature at this time. Perhaps, the theories we develop in management accounting will never be amenable to the crucial testing that characterizes natural science phenomena. That is, we will not have analogies of evolution replacing “creation” theories, continental drift/tectonic plate theory making obsolete previous geological theories, or quantum theory overthrowing Newtonian mechanics. Therefore, I prefer to have interested readers consult Kaplan (1964), Popper (1959), Kuhn (1970b), and Lakatos & Musgrave (1970) on theory testing and falsification and move on, in this paper, to a more detailed discussion of the role for clinical research and field studies for knowledge creation in management accounting.

**CLINICAL VS ANALYTIC RESEARCH**

Roethlisberger (1977, p. 390) observed that the clinical approach of acquiring knowledge by describing and classifying the activities of skilled practitioners, is frequently by-passed by researchers because clinical research is considered less elegant than deductive research:

Going from the bottom up involves sweat, tears, toil, a great deal of imagination, and little deductive logic. However, one goes from the top to the lower levels by pure logic. So the tree, when looked at from the bottom up, is indeed less elegant than when looked at from the top down.

Roethlisberger (1977, p. 392) also warns of the dangers of attempting to use the tools and methods of one level to obtain knowledge that can only be obtained by working at a different level. Examples of such mistakes are:

The clinician who plasters descriptions of situations with diagnostic concepts and thinks that he is thereby explaining them; the conceptual logician who thinks that he is explaining phenomena by tying concepts into neat logical bundles; the correlation seeker who thinks his significant correlations explain something; the causal hypothesis seeker who thinks he has the tools for getting hold of propositions . . . ; the general proposition maker who is so enamored with scientific explanations that he ceases to understand the phenomena with which he deals.

Mitroff et al. (1974) have also criticized conventional forms of scientific activity, particularly in the managerial sciences, both for not starting from the reality of actual problem situations, and for not returning to reality for testing proposed solutions in actual settings. Instead, researchers using either a formal deductive approach (with axioms, models, and theorem proving) or a formal inductive approach (with well-specified rules for data gathering and hypothesis testing) alternate between ever more complex model building and model solution activities without either questioning initial starting assumptions or attempting to implement and test their solutions. In terms of the Knowledge Enterprise of Fig. 1, conventional scientific researchers in the managerial sciences seem to cycle within the upper reaches of the tree never getting near the phenomena they are attempting to understand or influence.

These views have also been articulated by Schön (1984) in describing the dilemma between rigor and relevance in all professional schools:

Rigorous professional practice is conceived as essentially technical. Its rigor depends on the use of describable,
testable, replicable techniques derived from scientific research, based on knowledge that is objective, consensual, cumulative, and convergent.

Unfortunately, as Schön observes:

In the varied typography of professional practice, there is a high, hard ground which overlooks a swamp. On the high ground, manageable problems lend themselves to solution through the use of research-based theory and technique. In the swampy lowlands, problems are messy and confusing and incapable of technical solution. The irony of this situation is that the problems of the high ground tend to be relatively unimportant to individuals or to society at large, however great their technical interest may be, while in the swamp lie the problems of greatest human concern. The practitioner is confronted with a choice. Shall he remain on the high ground where he can solve relatively unimportant problems according to his standards of rigor, or shall he descend to the swamp of important problems and nonrigorous inquiry?

Some researchers have continued to develop formal models for use in problems of high complexity and uncertainty, quite undeterred by the troubles incurred whenever a serious attempt is made to put such models into practice. They pursue an agenda driven by evolving questions of modeling theory and techniques, increasingly divergent from the contexts of actual practice.

A professional who really tried to confine his practice to the rigorous applications of research-based techniques would find not only that he could not work on the most important problems but that he could not practice in the real world at all.

Thomas & Tymon (1982) describe five problems that arise when researchers of management processes, driven by their desire to use "rigorous" research methods, appropriate only for the higher branches of the Knowledge Tree, attempt to study phenomena for which little clinical knowledge or awareness of practitioner skills exist. First the models or data will not be representative of the phenomena in an organizational setting. Because of a bias towards internal validity, the researcher will use research methods, such as controlled laboratory experiments, that constrain the data and phenomena so severely that the conclusions have limited external validity in actual organizational settings.

Second, in the choice of goals, the researcher may choose dependent variables that diverge considerably from practitioner concerns; variables that arise from the researcher’s own values or a priori notions of managerial goals. Third, the researcher may focus on independent (causal) variables that the practitioner may be unable to manipulate. Fourth, the isolated researcher may find it difficult to produce a theory which meets or exceeds the complexity of common sense theory already used by a practitioner. That is, as researchers attempt to move from their deductive systems down the right-hand side of the knowledge tree to affect practice, they may find it difficult for their recommendations to improve upon the rules-of-thumb and considerable abilities of the skilled practitioner. Fifth, phenomena in organizations may change faster than scientists can come to grips with. Thus, many managerial scientists will find themselves working on problems that existed decades ago but not the critical problems facing managers of contemporary or future organizations.

In summary, the scientific inquiry of management processes must start and end with clinical research; observing skilled practitioners in actual organizations and observing how proposed improvements to practice actually affect real organizations. The methods used for clinical research will be less formal and less rigorous than those used at subsequent stages of the knowledge creation process. But accurate descriptions of the environment in which practitioners operate (the ecological context of decision-making) and the active involvement of practitioners with the research process are just as necessary, if not more so, than our traditional deductive and inductive research activities. It is in this area where empirical studies can play a decisive role.

Note that the absence of external validity for laboratory experiments is a major difference between research in the natural sciences and research on management processes. Physical phenomena observed in the laboratory will be representative of phenomena occurring outside laboratories also. But experiments on students or even practitioners in artificial laboratory settings will give virtually no insight into the practice of management in its complex organizational setting.
EMPIRICAL RESEARCH METHODS

One can envision three types of research programs being carried out by systematic observation in cooperative organizations: case studies, field studies, and field experiments. Case studies are characterized by the intense examination of a single entity. The case study provides a rich description of an actual situation such as a management or organizational setting. The original case study, of course, is the detailed record arising from a physician’s clinical examination. Case study data are frequently collected by multiple means. These include qualitative methods, such as personal interviews, personal observation and qualitative descriptions of a company, its markets, products, competitors, technology, systems, etc. But quantitative data sources, such as financial data, market performance data, and quantities of equipment personnel, materials and products should also be incorporated. The case study would also describe the context in which a management system operates and in which managers’ decisions and actions occur. In general, however, no attempt is made by the researcher to manipulate or control for independent variables within the organization. Thus, for the most part, case studies tend to be used more for hypothesis-generation than for hypothesis-testing (an exception will be noted below). The principal benefit from a case study is to develop an in-depth description and understanding of a particular managerial problem. The primary drawback is that conclusions from case studies can be subject to numerous interpretations because of the possibility of confounding factors in the entity being studied and the methods by which case study data are collected, aggregated, and analyzed.

Field studies can be viewed as cross-sectional case studies. That is, the researcher would measure independent, dependent, and possibly confounding variables in an unobtrusive way as possible across a number of organizations. The research would be nonexperimental; based on extensive visits and data collection at organizations, the researcher would attempt to develop classification schemes for phenomena, perhaps measure association by correlating variables across organizations, and even tentatively produce a theory to explain the observed phenomena. Recent excellent examples include Garvin’s (1983) study comparing the quality of U.S. and Japanese air-conditioning companies, and Donaldson’s (1984) study of the financial policies of twelve U.S. corporations. Field studies require long times for completion because of the need to obtain extensive and detailed information from many organizations. Time periods of three to four years would be typical, or even minimal, for the study of an interesting management phenomenon.

Field experiments involve the experimental manipulation of one or more variables within an actual organization and the subsequent measurement of the impact of the manipulation on one or more dependent variables. Field experiments where the researcher has some control over the type and domain of the manipulation seem less likely for studying management phenomena; they have been used more in other social sciences interested in studying the effects of government programs such as Head Start (education) and the negative income tax (welfare). It would be unusual for profit seeking (or even most non-profit) organizations to agree to a major change in their internal accounting and performance measurement systems in order to advance our understanding of management accounting phenomena.

But a fortunate researcher may become involved with an organization which, for its own reasons, undertakes a major change in its measurement and reporting system. With such an

---

4 The definitions and descriptions of these three methods have been adapted from Stone (1978, ch. 7). I have deliberately omitted three other types of empirical research methods described by Stone, laboratory experiments, surveys, and simulation, because I believe their external validity to be far too low for understanding management accounting phenomena. Chapter 7 in Stone (1978) contains an excellent discussion of the strengths and weaknesses of all these empirical research methods.
occurrence, the researcher would then be in a position to document the impact of the change. Such a study would be especially valuable if the researcher had previously performed a case study in the organization so that an extensive description of the former system already existed. With this strategy, a researcher could embark on a program of field research by performing an initial series of case studies. These would establish the researcher's credentials and reputation in the cooperating organizations. Then, should some of these organizations institute changes in their internal accounting systems, the researcher could return to study the impact of these changes.

The advantage of such a longitudinal study, which we may call a "process-tracing" study, is that confounding variables, which vary across organizations, do not arise when studying individual organizations over time. The disadvantage is that in such a nonexperimental longitudinal study, the change in the accounting system can occur simultaneous with or even be caused by more fundamental changes in the organization's strategy, technology and competitive environment. In this case, the researcher may find it difficult to distinguish impacts caused by changes in the management accounting system from changes also occurring in the firm's organization, strategy, technology and markets.

These three types of field-based research studies can be used across the spectrum of the knowledge enterprise; from description and classification through theory formulation and testing. Bonoma and Wong (1985) have devised a four-stage model to describe how field-based studies can be used throughout the research process:

1. *Drift.* Explore and clarify the nature of the project's scope; learn the issues and hypotheses (accepted wisdom) in the field.

In the drift stage, the researcher writes case studies, attempting (a) to learn the concepts, locale, and jargon of the phenomenon as it is practiced in the field, and (b) to integrate the existing literature, prior concepts of the phenomenon, and what is observed in practice. The researcher must be open to new concepts and possibilities; management contexts are observed, informed practitioners are interviewed with an open-ended, flexible agenda in an attempt to glean the state-of-art of practitioner knowledge. Each case constructed at this stage can both be studied as an interesting problem in its own right and to stimulate further thought on possible general rules and classification schemes.

2. *Design.* Collect observations at field sites to complete cells in a classification scheme.

The drift stage can evolve into design when the researcher develops a complete classification scheme and perhaps develops a tentative model or theory to explain the variety of observations collected to date. For example, if one discovers that cost accounting schemes vary by the organization of production operations (ranging from general purpose job order shops, through batch assembly and process operations), and by product type (frequent vs infrequent changes in product characteristics), then we will probably wish to have at least one or two observations for each combination of classifying variables.

3. *Prediction.* Test tentative theories or models in new field sites.

At the prediction/test stage, the researcher has a model and a classification scheme that enable predictions to be made about the phenomena being studied. The researcher can then develop new cases at different sites from those used to develop the conceptual model. In effect, these new sites serve as a test or "hold out" sample for the researcher, but they can also serve to extend the theory to industries or settings that have not been examined in the prior two stages. These extensions permit an exploration of whether industry or setting-specific variables are relevant.

4. *Disconfirmation.* Boundary experiments at extreme cases to determine the limits of the conceptual model.

The fourth stage attempts to test the emerging theory in extreme situations, where the
generalizations from the theory might be expected to break down. For example, one might look at much smaller or newly emerging organizations, or explore service rather than manufacturing entities.

It may be implausible for us to expect theories of management processes ever to reach the stage of those in the natural sciences where a small number of experiments can disprove or overthrow a theory. Because of the complexity and ambiguity of social/managerial phenomena, it seems more likely that we will continue to have competing theories co-existing, as a function of researchers' varying perceptions of the critical issues in the phenomena they observe. Probably the best we can hope for is to develop theories that are applicable in different domains and circumstances. The purpose, therefore, of this fourth stage of case study research is to determine and delineate the domain of applicability of theories developed to explain phenomena observed in prior field research.

In summary the Bonoma-Wong proposal attempts to capture the full range of scientific activities by a sequence of field research projects. These projects are developed for: (1) Description; (2) Classification; (3) Theory Development and Limited Testing; and (4) Testing Limits of Theory.

Other social science researchers have also described the role of field research not just for their obvious roles in description and classification activities, but for their role in hypothesis and theory formation activities. Mohr (1985) describes conditions where detailed case study analysis, especially in a "before-after" design (a "process-tracing" study), can be more effective for theory building than traditional large sample research methods. For investigating the impact of an intervention (e.g. a change in production methods or a change in accounting procedures), a detailed process-tracing study, with a careful understanding of the underlying process, can provide a high level of internal validity. For associational research, Mohr believes:

When the aim is understanding or explanation, there is a tradeoff between case-study research with [detailed diagnosis and process tracing] and large-sample research, typically rendered in correlations and regressions. The former emphasizes understanding how and why X and Y are related, on the average, across many subjects.

George & McKeown (1985) also argue that case studies, especially process-tracing, detailed within-case analysis, can provide valuable evidence, beyond that obtained from traditional "quasi-experimental" methods:

The case study method differs from quasi-experimental approaches in its heavy reliance on *within-case* analysis as a way of evaluating claims about causal processes . . . .

A second distinction between quasi-experimental and case study approaches lies in their treatment of the distinction between theory formation and theory testing. The conventional quasi-experimental position is founded on a clear-cut distinction between these two activities. Theory formation is a subject which is normally ruled "out of bounds" — statistical analysis, for example, is about testing hypotheses not about generating them. By contrast, the case study techniques often entail viewing hypothesis formation as an objective of the study. In case studies, a clear demarcation between hypothesis formation and hypothesis testing often is absent, because the research process often involves iterative cycling in which hypotheses are successively "fitted" to observations. Problems with the fit then lead to revision of the hypotheses and may also incite searches for additional data.

Yin (1981a, 1981b) also disputes the common stereotype that case study research:

(1) should be used only at the exploratory stage,
(2) leads to unconfirmable conclusions, and
(3) is a research method of last resort.

Yin (1981b) argues that case studies are especially relevant for studying knowledge utilization because it is characterized by:

- a series of decisions that occur over a long period of time, with no clear beginning or end points;
- outcomes whose direct and indirect implications are too complex for single factor theories;
- a large number of relevant participants; and
- situations that are special in terms of (organizational) context, historical moment in time, and other key elements.

... if one is desirous of answering "how" and "why" questions instead of or in addition to questions of frequency, case studies are the more appropriate strategy.
Thus, were researchers in management accounting to undertake field research, they could draw upon the experiences and knowledge acquired by other social science researchers to guide their efforts. In particular, the study of knowledge utilization described by Yin seems especially relevant for studying the use of management accounting information in organizations.

SUMMARY AND CONCLUSIONS

We have surveyed and quoted, in previous sections, from many observers of the scientific research process. Despite the differing backgrounds and points of view of these observers, one can hardly avoid noticing that all agree on the need for scientific activity to start with observation and description of phenomena. Given the paucity of detailed, systematic observation and description in the management accounting literature, it seems compelling that we must initiate activities to overcome this serious defect.

Cost accounting and management control procedures function in complex organizational settings. Therefore, our initial effort to observe and describe management accounting practices must capture the richness of the organizational environment. Initially case studies would seem to provide the ideal vehicle for communicating these deep, rich slices of organizational life. The multiple data sources used to prepare case studies permit a variety of factors to be captured in the descriptive material. Clearly, the field researcher will need to understand more than just cost accounting issues. The phenomena being accounted for — production, marketing, R&D, administrative activities — must be described in some detail and patterns of reporting relationships, decision making, and control must also be captured if we are to understand the context within which the accounting system is functioning.

A continual stream of excellent, descriptive case studies can provide at least three significant benefits for management accounting academics. First, the case studies provide the basis for other forms of research activities. They will be the crucial building blocks for classifying cost accounting and management control practices. Our current schemes were derived 80 or more years ago; it would be useful to determine their validity and usefulness for contemporary organizations. The case studies will also provide a firmer basis for our modeling, theory-building, and hypothesis-formation activities, activities that today occur in the absence of data and observations. As a consequence, we tend to work on problems already in the literature or found in the literature of other disciplines, rather than on problems that emerge from the decision and control activities of managers in actual organizations.

A second benefit from descriptive case studies arises from the discipline they provide to seek out interesting organizations or interesting practices. Unlike the natural sciences where the underlying phenomena seem quite stationary (it is unlikely that gravitational constants, the speed of light, or the structure of the benzene molecule have varied during the past several hundred million years), organizational and management phenomena must continually adapt to changing circumstances. The factors critical to the success of global corporations in 1985 can be very different from the success factors for Adam Smith's pin factories, 1850 textile mills, 1880 steel plants and 1920 automobile companies. In order that academic research does not fall too far behind changes in the competitive environment and the new procedures being introduced by innovative organizations, researchers need to be informed about contemporary practice and the evolving skill of the best practitioners. Because of these changes in the phenomena of interest, there is even more of a need for continual observation and description in management-related disciplines than in the natural sciences (where, of course, observation and measurement are long established and accepted research practices).

When selecting organizations for study, the researcher should not be interested in obtaining a randomly selected or "representative" set of
firms. The goal for this research is not to document "average" practice. Rather, it is to learn about innovative, leading-edge practice. We want to anticipate in which directions best practice is developing. When interested in forecasting which way the thundering herd is going, we don't ask the bull in the middle. We wish to seek out the bulls or firms who are leading the pack.

A third benefit of case studies accrues more to our teaching than to our research activities. Having rich descriptions of cost accounting and management control practices in actual organizations will enhance greatly our ability to communicate with students about the strengths and limitations of alternative schemes. At present, with only a few exceptions, we use extremely simplified settings to describe management accounting concepts. Students graduating from our programs will likely encounter, in their first job, an organizational setting and a cost accounting system that are orders of magnitude more complex than anything they have been trained on in school. This is not an ideal situation for implementing new ideas. Currently, we train our students in the mechanics of simple systems operating in simple production environments. Students are not trained to assess whether a standard or actual cost system, a job-order or process costing system, will function better (or even what it means, in an operational sense, to "function better") in their company's situation. Thus, our failure to have rich descriptions of management accounting situations in contemporary organizations is diminishing the effectiveness of our teaching, as well as our research activities.

As we move from isolated case studies to more systematic field and process-tracing studies, we can begin to develop analytic knowledge: to perform measurement, association and theory formation and testing activities. In fact, it is hard to imagine how to test theories in management accounting if the testing is not performed within actual organizational contexts. Also, if a sufficient number of complex cases are produced and become publicly available to researchers, they can be used as a data base for quantitative, statistical analysis (see Miller & Friesen, 1980, 1982).

Testing a limited set of hypotheses could probably occur at present. For example, one could perform "before—after" studies to determine whether the introduction of measures of quality, inventory levels, or productivity caused performance improvements in the attributes being measured. Perhaps, the particular way that an attribute was measured would affect the actions and decisions taken relative to that attribute. Such studies may find it difficult to isolate the unique contribution of changes in the management accounting system since such changes would likely be made in conjunction with organizational actions to increase quality, reduce inventory, or improve productivity. It may be possible, however, to find a situation where real changes were attempted throughout an organization but measurement changes were implemented only in certain sections of the organization.

For examples drawn from management control, we may be interested in the reasons why firms do (or do not) adopt inflation-adjusted statements (see Hertenstein, 1984), or relative-performance evaluations (see Antle & Smith, 1984) for internal motivation and evaluation of division managers. Initially, one would study a few organizations in considerable detail to learn, not just whether or not, but why they did (or did not) adopt either of the above performance evaluation schemes. From this initial set of studies, the researcher may formulate hypotheses on where the scheme would be expected to be adopted and where not. These hypotheses would then serve both to guide the sample selection for the next set of companies and to be the targets for the testing/disconfirmation procedures. When examining the management control procedures for the new set of companies, the researcher would be interested both in prediction (whether adoption of inflation-adjusted internal statements or of relative performance evaluation occurred where expected and whether they did not occur where they were not expected) and also in whether the reasons for adoption and non-adoption were consistent with the reasons given in the researcher's
hypotheses. In other words, field studies in a theory testing mode can serve not only to document the existence or non-existence of practices but also to confirm *how* and *why* certain practices have, or have not, been implemented. Therefore, they provide a rich research method for the study of management accounting issues.

Given the potential for field research to illuminate the field of management accounting, one might reasonably wonder why a stronger tradition for this research method does not already exist. The omission undoubtedly arises from an impression that field research activities lack legitimacy as scientific research; that scientific activity consists of either deductive reasoning, model-building and theorem proving, or the statistical analysis of large data bases. With this "model" of scientific activity, there is no role for rich description or for testing in relatively few organizations. The goal of this paper has been to document the role that observation, description and classification must play as the precursor for informed theory creation and model-building, and the potential for testing theories about accounting in organizations by studying the accounting system of actual organizations. Stating the case somewhat stronger, it now seems implausible that research in a field can claim to be "scientific" when its models and theories are neither inspired from nor tested by actual observations.

There are a host of subsidiary reasons that also serve to inhibit clinical research activity. Among the inhibiting factors are that the research takes longer to accomplish, it requires additional resources for traveling to organizations, it may be difficult to gain access to cooperating organizations, the standards to distinguish "good" from "bad" field research are not yet established, and prestigious journals do not typically publish field studies. Also, the necessary research skills for interviewing practitioners, gleaning information from operating organizations, and observing and listening carefully to learn about the key success factors in corporations are skills not typically taught in doctoral programs. But, these real concerns need not be decisive if it can be generally agreed that empirical research is essential to the future development of management accounting. If a commitment can be made to value the output of field research and field researchers, we can overcome all the other obstacles that may currently be serving to inhibit empirical research.

One additional concern, however, may be more difficult to deal with. At present research on many interesting financial accounting questions can be performed by researchers in their offices, at their computer terminals, and in their libraries. The availability of stock price data, periodic financial statements, and important news releases on computer data bases plus access to documents such as annual reports, proxy statements, lending agreements and prospectuses in libraries makes it possible for financial accounting researchers to conduct a full range of scientific activities without engaging in costly and time-consuming field studies. Therefore, if we advocate that management accounting research must start and eventually end with observations from actual organizations whereas financial accounting research can be performed without leaving universities, we could well be encouraging a further shift from managerial to financial accounting research.

Some differences between the way research can be conducted in financial and management accounting may be inevitable. A corollary of a proposition quoted by A. Kaplan in the Appendix, "Nature does not recognize mathematical difficulties", is that "Nature does not guarantee that research in all disciplines requires the same amount of time and effort". Even conceding such differences, however, prices and demand for accounting researchers may adjust to compensate, at least in part, for shifts in the supply of academics willing to do research in financial vs management accounting. Also, financial

---

*The issue of standards for field research and the lack of such research in academic journals is aggravated by having most management-related field research published in the *Harvard Business Review* or monographs (e.g. Garvin, 1983; Donaldson, 1984), whose intended audience is experienced practitioners, not other academics.*
accounting researchers may soon discover the value of clinical research to supplement and perhaps substitute for their statistical analyses. As these researchers move beyond associational studies of capital market processes and attempt to understand managerial behaviour (how managers choose their accounting procedures, how they disclose information relevant to the firm, why disclosures are made at different times, and how accounting data are used to compensate senior executives, to name just a few topics currently occupying the attention of financial accounting researchers), they may eventually realize that these "how" and "why" questions are best answered by speaking directly with executives and collecting data within the organizations making these decisions rather than by attempting to induce causality from reduced form regression models with low explanatory power. Thus clinical research methods, once accepted and underway, may actually spread from managerial to financial accounting.

To conclude, field research seemingly provides the only mechanism by which management accounting can become a scientific field of inquiry. There is no disagreement among scholars of the scientific inquiry process that formal, systematic observation must be part of any science:

> It is in the empirical component that science is differentiated from fantasy. An inner coherence, even strict self-consistency, may mark a delusional system as well as a scientific one (A. Kaplan, 1964, p. 34).

Because management accounting practices can only be described and understood within the context of actual, ongoing organizations, the domain of empirical work in management accounting must be in these organizations.

The current lack of theories in management accounting also can likely be traced to our failure to observe, firsthand, how management accounting functions in organizations. Recall Henderson's (1970) observation that effective medical research starts with

> ... hard, persistent, intelligent, responsible, unremitting labor in the sick room, not in the library.

Daft, in a provocative essay on the "craft" of research on organizations, also emphasizes the role for on-site observation by researchers:

> The difficulty that many authors have developing interesting and insightful theories about organization(s) probably is explained by the lack of experience with organizations. G. R. Grice admonished his students who were trying to understand animal learning: "No matter how many assistants you may hire, always handle your own rat" (Daft, 1983, p. 543).

For too long now, management accounting researchers' attitude to empirical research has had all the enthusiasm of most people for handling their own rats. This distaste for conducting research within organizations must be overcome if we are to gain the intuition, familiarity and experience that are necessary for management accounting research to progress.

BIBLIOGRAPHY


THE ROLE FOR EMPIRICAL RESEARCH IN MANAGEMENT ACCOUNTING


APPENDIX. A. KAPLAN ON MODELS VS THEORIES

The role of models in the entire process of scientific inquiry has been treated in great breadth and depth by A. Kaplan (1964, pp. 258–293). Kaplan dislikes using the word "model" to be synonymous with "theory." Instead:

A more defensible usage views as models only those theories which explicitly direct attention to certain resemblances between the theoretical entities and the real subject-matter. With this usage in mind, models have been defined as "scientific metaphors."... No two things in the world are wholly alike, so that every analogy, however close, can be pushed too far; on the other hand, no two things are wholly dissimilar, so that there is always an analogy to be drawn, if we choose to do so. The question to be considered in every case is whether or not there is something else to be learned from the analogy if we do choose to draw it ...

We... distinguish five different senses in the confusing and often confused usage of the term "model":
(1) any theory more strictly formulated than is characteristic of the literary, academic, or eristic cognitive styles, one presented with some degree of mathematical exactness and logical rigor; (2) a semantic model, presenting a conceptual analogue to some subject-matter; (3) a physical model; (4) a formal model, a model of a theory ... a structure of uninterpreted symbols; (5) an interpretive model, providing an interpretation for a formal theory.

Kaplan identifies two important functions for models:

Data organization
Without a theory there is only a miscellany of observations ... But the more ill defined the theory, the more vague and uncertain is the significance it confers on data. Models have this merit, that they do not allow us to comfort ourselves with the notion that we are following up an "idea" when we are only moving from one observation to the next in the hope that something will turn up ... Models are ... conscious, explicit, and definite.

Cognitive Styles
The model allows the scientist to make clear to others just what he has in mind ... The literary style ... provides valuable materials for scientific treatment but (is) not ... scientific in itself ... Though the symbolic style also has its own notations and idioms, their meaning is likely to be much more fully and explicitly specified.

Secondly, the eristic and especially the symbolic styles distinguish between definitions and empirical propositions ... It is the postulational style that has recently excited so much interest and attention among behavioural scientists.
First, ... the postulational method provides us with a way of handling complex phenomena ... We may lay down separate postulates for each of the various factors of the situation and formulate explicitly whatever theories we might have as to how they interact. It then becomes a matter of deduction, or even of calculation, to determine the outcome in specific circumstances.
Second, the construction of postulational systems reveals gaps in our knowledge. It allows us to identify the propositions that are needed for us to be able to derive the conclusions in which we are interested: gaps in knowledge are revealed by the gaps in "proofs". This is one of the specific ways in which a carefully formulated theory can direct the search for data. For example, as economic inferences were made more rigorous, economists became increasingly interested in data bearing on motivations operative in the economy, in connection with such matters as attitudes toward risk, the value of gratifications lying in the future, or utilities not reflected in money prices.

The broadening of the conception of measurement to include the application of scales that are not fully quantitative has also been closely associated with the use of the postulational style ... The whole development of the theory of games and much of utility theory illustrate this feature of the postulational style.

Perhaps the most important advantage conferred by the postulational style lies in its deductive fertility. It allows us to process our information so that we can squeeze out of our data a great deal of content not otherwise available to us ... A postulate system is the more valuable the richer it is in unanticipated consequences, for these show us how much more we know than we thought we did.
Of course, the conclusions deduced do not constitute knowledge unless and until the postulates from which they are derived are themselves known to be true. The system itself must be separately justified. The postulational method simplifies and shortens the process (of verification of our hypotheses) in that a verification of the postulates constitutes at once a verification of all the theorems that flow from them. Conversely, if any of the theorems is shown to be false, the postulate set as a whole is thereby falsified. Thus we can take advantage of very remote consequences to provide tests of our ideas.

Finally, the postulational style makes possible an economical summary of our actual or anticipated findings. The postulates logically contain within themselves all the information embodied in the whole set of theorems which they entail.

Kaplan is very specific in the reservations he has about concentrating too greatly on model building to the exclusion of other forms of scientific inquiry:

Those scientists who speak of their work as “model building” often give the impression that such an endeavor is the only true begetter of scientific knowledge, and the construction and testing of models is itself the very model of modern scientific activity. In short, models — to play on another meaning of the word — are much in fashion, though to say so is by no means to prejudge their scientific significance and worth.

An entire section is devoted to identifying the shortcomings of models:

1. **Overemphasis on symbols**
   I find it hard to resist the impression that part of the appeal of the model lies in an unconscious belief in the magic of symbols. Unfortunately, in behavioral science, it is not uncommon that the symbolic style is only a mode of expression and not of thought. Elaborate notations often only codify the obvious. What they achieve is hardly worth the trouble. By the use of symbols, a proposition is given the form of a scientifically useful statement, but not always the content. The essence of mathematics is not its symbolism but its methods of deduction. Models [of the symbolic style enable us to] learn from them only what we already knew without their help:

2. **Overemphasis on form**
   What limits the usefulness of models is not always an inadequacy in our knowledge of mathematics or logic but rather an inadequacy in our knowledge of the subject-matter. The requirements of a model then impose a premature closure on our ideas. We tinker with the model when we might be better occupied with the subject-matter itself. Our knowledge may be on the level of folk wisdom; incorporating it in a model does not automatically give such knowledge scientific status.

   Rationalism is more concerned with logical form than with empirical content. The model itself then becomes the object of interest, as means so often usurp the importance of the ends they are meant to serve. The failing is the tendency to engage in model building for its own sake. Whether models help is not something that can always be taken for granted.

3. **Oversimplifications**
   Science always simplifies; its aim is to formulate what is essential for understanding, prediction, or control. It is one thing to ignore certain features of a complex reality (if what is being neglected is not essential, at least to a first approximation. It is quite another thing to follow the principle of “the drunkard’s search” and to simplify in a particular way because then the model would be so much more elegant or so much easier to work with. (“Nature doesn’t care about mathematical difficulties,” Fresnel.)

   In short, a crude but more realistic set of hypotheses may serve the purposes of inquiry in particular cases far better than a refined but oversimplified model. This condition is especially likely when either we do not know what factors can safely be neglected, or when we cannot treat by the mathematics available to us for the model some factors which we already know to be important in that context of inquiry.

4. **Overemphasis on rigor**
   Another failing of models — or of model builders — consists in an undue emphasis on exactness and rigor. Models are often improperly exact: they call for measures that we cannot in fact obtain, or that we would not know how to use if we did obtain them.
Many (formal postulational) models, for all their rigor, have no remarkable deductive fertility: the ratio of significant theorems to the number of postulates and definitions is often disappointingly low. Students of man and society (Veblen, Weber, Freud) are universally recognized to have been men who had ideas, even though their work was not conspicuous for the rigor of its deductions. Even in physics, Faraday has made the most extensive additions to human knowledge without passing beyond common arithmetic. Careful observation and shrewd even if unformalized inferences have by no means outlived their day.

5. Map reading

A fifth shortcoming in the use of models is the failure to realize that the model is a particular mode of representation, so that not all its features correspond to some characteristic of its subject-matter. It is the obverse of oversimplification: instead of leaving out of the model something which should be in it, we read the model as containing something which in fact is not a part of it. If a computer can serve as a model for a theory of the human brain, this has nothing to do with the fact that they both work by electric impulses, nor does it in the least justify the assertion that the brain is nothing other than a computer or that the artifact is nothing other than a thinking machine.

6. Pictorial realism

We may misconceive (the properties of the model) as constituting an image or likeness of what is being modeled. That the model can effectively perform its functions in inquiry is regarded as being due to, and hence a sign of, its being an accurate representation. The mistake of pictorial realism is forgetting that the similarity exists only in a given perspective, that it depends upon a particular mode of representation.

If a model is a scientific metaphor, pictorial realism takes the metaphor to be a literal statement. We think "that's what it is" instead of "that's what it's like."

In a summary section (33) on the limitations of models and model building in the behavioral sciences, Kaplan again warns of the dangers of isolating models from observations:

Mathematical social science is, first and foremost, a social science. If it is bad social science (i.e. empirically false), the fact that it is good mathematics (i.e. logically consistent) should provide little comfort. A strong emphasis on empirical materials is important to counteract the rationalist interest in form which is intrinsic to model building. I have often had the impression that what the behavioral scientist needs is more concern with his subject-matter and less concern with whether he is building a "Science".

Many students of man and his ways seem to have very little interest in—people! Most scientific achievements were attained by men who were fascinated by a particular subject-matter and were prepared to use anything or to do anything that might satisfy their interest in the subject.

The division of labor may allow a theoretician to pursue his ends without himself providing the data by which the theory is shaped and tested. But the empirical materials must nevertheless constantly be in his mind. As attention focuses on the properties of a model, the data which it is to fit tend to be of only secondary importance, and at last are brushed aside as ugly facts capable of ruining a beautiful theory. The economist is often disdainful of the concreta of business administration.

... an assumption need not be true to do its work effectively as an assumption. The point is that what is only affirmed as an assumption in the model may come to be asserted as a truth in the theory that emerges from the model, and mathematical needs come to be projected onto the facts as empirical discoveries.