INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of “sectioning” the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.
Selling, Thomas Ira

COGNITIVE PROCESSES IN INFORMATION SYSTEM CHOICE

The Ohio State University

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106

Ph.D. 1982
COGNITIVE PROCESSES IN INFORMATION SYSTEM CHOICE

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Thomas Ira Selling, B.S., M.B.A.

****

The Ohio State University

1982

Reading Committee:

Thomas J. Burns, Chairman
R. Bhaskar
Gordon Clark
Mari Jones
Ray G. Stephens

Approved By

Thomas J. Burns
Advisor
Faculty of Accounting
ACKNOWLEDGEMENTS

This dissertation could not have been written without the support of others. It is indeed a pleasant task to take this opportunity to thank them all.

My committee members were Professors R. Bhaskar (Accounting), Thomas Burns (Accounting), Gordon Clark (Industrial Systems Engineering), Mari Jones (Psychology) and Raymond Stephens (Accounting). I am grateful to all for their support and guidance. Professor Jones also taught a course in cognitive psychology which I found enjoyable and very instructive. Professor Clark was kind enough to fill a void when Professor William Morris left the university. Professor Morris's courses in decision theory influenced me greatly, and the initial ideas for this dissertation were formulated while writing a paper for one of his courses.

I derived significant benefit from the research experiences of Professors Stephens and Bhaskar in verbal protocol analysis. Professor Stephens also provided much help in transforming the preliminary drafts into a final product. Both were of great help in making the many design decisions involved in conducting the study. Professor Ernest Hicks provided assistance through practical suggestions which were incorporated in the experimental task. James Noel, formerly a peer at OSU, and now Assistant
Professor at Carnegie-Mellon University, was also of great help in the formulation of the dissertation proposal. The idea to represent the information evaluator's problem as a matrix (Figure 5) was his. Professor John Shank provided constructive suggestions as well as moral and financial support. The financial assistance was used to have the protocols transcribed, and was derived from the donations of Arthur Young and Company to OSU.

I am particularly grateful to Professor Burns, chairman of my committee, and my advisor since I came to Ohio State. There were many painful occasions when his criticisms provided much needed momentum to my efforts, but I will always remember most clearly his words of comfort in numerous difficult situations.

Eileen Brubeck transcribed the protocols with alacrity and accuracy. It was the kind of task I was very glad I did not have to perform myself, and I am thankful that Eileen was there to be relied upon so completely. Rodger Holland, now Assistant Professor at Florida State University, introduced me to the SCRIPT text processing language which was used to produce this dissertation. Use of this language in conjunction with some of the programs Rodger wrote saved me many hours of work, and I am grateful for his help and continual patience.

I owe a large debt of gratitude to family and friends. Dennis Ferguson and Gary Gabaccia are not only lifelong
companions, but have been models of comportment and achievement for me. My parents, Martin and Hilde Selling, have always been my greatest inspiration. My father, in typical fashion, also proofread the entire dissertation, and provided numerous helpful suggestions for improving its readability.

Finally, I dedicate this dissertation to the memory of my grandfather, Adolf Kaufmann, and my uncle, Leo Selling. It is my greatest sorrow that they are not here to witness the completion of this dissertation.
BIOGRAPHICAL INFORMATION

July 3, 1951

Bora - Rahway, New Jersey

1973

B.S., Hotel Administration, Cornell University, Ithaca, New York

1973-1975

Lecturer, School of Hotel Administration of Puerto Rico, Auspices of Cornell University, San Juan, Puerto Rico

1975-1977

Teaching and Research Assistant, Cornell University School of Hotel Administration

1977

M.B.A. Cornell University

1977-1981

Teaching Associate, Department of Accounting, The Ohio State University, Columbus, Ohio

PUBLICATIONS


FIELDS OF STUDY

Major Field: Accounting

Minor Fields: Decision Theory and Statistics
### CONTENTS

**ACKNOWLEDGEMENTS** ........................................ ii

**BIOGRAPHICAL INFORMATION** ............................... V

**LIST OF TABLES** ........................................... viii

**LIST OF FIGURES** .......................................... ix

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Motivation</td>
<td>1</td>
</tr>
<tr>
<td>Overview of the Normative Solution</td>
<td>5</td>
</tr>
<tr>
<td>Overview of Findings</td>
<td>12</td>
</tr>
<tr>
<td>Organization of Remaining Chapters</td>
<td>14</td>
</tr>
<tr>
<td>II. LITERATURE REVIEW</td>
<td>15</td>
</tr>
<tr>
<td>Internal Control Evaluation</td>
<td>16</td>
</tr>
<tr>
<td>Decision Process Methodologies</td>
<td>22</td>
</tr>
<tr>
<td>Lens Model</td>
<td>23</td>
</tr>
<tr>
<td>Process Tracing Models</td>
<td>26</td>
</tr>
<tr>
<td>Expert Judgment and Optimality</td>
<td>38</td>
</tr>
<tr>
<td>Decision Maker Algorithms</td>
<td>41</td>
</tr>
<tr>
<td>Heuristics</td>
<td>42</td>
</tr>
<tr>
<td>Mathematical Models with Predictive Ability</td>
<td>46</td>
</tr>
<tr>
<td>Related Studies in Accounting</td>
<td>51</td>
</tr>
<tr>
<td>Studies Utilizing Protocol Analysis</td>
<td>51</td>
</tr>
<tr>
<td>Empirical Studies in Information System Choice</td>
<td>66</td>
</tr>
<tr>
<td>Conclusion</td>
<td>73</td>
</tr>
<tr>
<td>III. RESEARCH DESIGN</td>
<td>74</td>
</tr>
<tr>
<td>Experimental Task</td>
<td>74</td>
</tr>
<tr>
<td>Description and Background</td>
<td>74</td>
</tr>
<tr>
<td>Subjects</td>
<td>81</td>
</tr>
<tr>
<td>Statistical Test for Trend</td>
<td>83</td>
</tr>
<tr>
<td>Statement of Research Question</td>
<td>85</td>
</tr>
<tr>
<td>Procedure</td>
<td>85</td>
</tr>
<tr>
<td>Protocol Analysis</td>
<td>86</td>
</tr>
<tr>
<td>Statement of Research Questions</td>
<td>87</td>
</tr>
<tr>
<td>Procedure</td>
<td>88</td>
</tr>
<tr>
<td>Computer Simulation</td>
<td>91</td>
</tr>
<tr>
<td>Statement of Research Questions</td>
<td>91</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Procedure</td>
<td>91</td>
</tr>
<tr>
<td>IV. EXPERIMENTAL RESULTS</td>
<td>96</td>
</tr>
<tr>
<td>Test for Trend</td>
<td>96</td>
</tr>
<tr>
<td>Processing Consistent with the Normative Model</td>
<td>98</td>
</tr>
<tr>
<td>Other Information Sources</td>
<td>104</td>
</tr>
<tr>
<td>Anchoring and Adjustment</td>
<td>111</td>
</tr>
<tr>
<td>Simulation of Anchoring and Adjustment</td>
<td>116</td>
</tr>
<tr>
<td>V. FINAL REMARKS</td>
<td>119</td>
</tr>
<tr>
<td>Appendix</td>
<td></td>
</tr>
<tr>
<td>A. NOTATION SYSTEM FOR EQUATIONS</td>
<td>124</td>
</tr>
<tr>
<td>B. INSTRUCTIONS TO SUBJECTS AND CASE MATERIALS</td>
<td>126</td>
</tr>
<tr>
<td>C. DESCRIPTION OF COMPUTER PROGRAMS</td>
<td>135</td>
</tr>
<tr>
<td>D. SUBJECT PROTOCOLS</td>
<td>143</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>262</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Subject Demographic Data.</td>
<td>83</td>
</tr>
<tr>
<td>2. Results of Tests for Downward Trend</td>
<td>97</td>
</tr>
<tr>
<td>3. Operator Totals by Subject.</td>
<td>100</td>
</tr>
<tr>
<td>4. Binomial Tests for Use of the Anchoring and Adjustment Heuristic.</td>
<td>113</td>
</tr>
<tr>
<td>5. Protocol Lines Supporting the Anchoring and Adjustment Heuristic.</td>
<td>115</td>
</tr>
<tr>
<td>6. Results of Tests for Trend on Simulated Subjects</td>
<td>117</td>
</tr>
<tr>
<td>7. Simulation Tests for Trend Counted by Varied Parameters</td>
<td>118</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>1. Overview of the Audit Engagement</td>
<td>21</td>
</tr>
<tr>
<td>2. The Brunswik Lens Model</td>
<td>24</td>
</tr>
<tr>
<td>3. Representation and Mapping</td>
<td>29</td>
</tr>
<tr>
<td>4. Bhaskar and Dillard's (1978) Representation of Accounting Knowledge</td>
<td>54</td>
</tr>
<tr>
<td>5. Choice Matrix for Simulated Information Evaluator</td>
<td>92</td>
</tr>
<tr>
<td>6. Flowchart of Anchoring and Adjustment Heuristic</td>
<td>94</td>
</tr>
</tbody>
</table>
Chapter I

INTRODUCTION

Task environments are a major determinant of the behavior of problem solvers (Newell and Simon, 1972). Understanding behavior better by examining specific problem solving domains may make it possible to develop modifications to the decision making processes of experts, leading to better decisions. Determination of both decision processes and decision making quality in specific domains may help by evaluating more accurately future costs and benefits from developing new decision methods or by improving the training of decision makers to use existing techniques. The task domain examined was the selection of information systems by accountants. The experimental task was the choice of sample sizes for tests of compliance with stated company internal control procedures which constitutes a part of the external audit process.

1.1 MOTIVATION

Prior empirical research in information system choice measured the performance of student subjects making sample size decisions in an abstract task. Subject decisions made over time were compared with the optimal decisions prescribed by a normative model. These studies found that
subjects performed poorly relative to a criteria of expected payoff maximization. The significance of these results rely on the supposition that the experimental task was relatively simple; i.e., the relevant features of the problem were presented parsimoniously and free from extraneous information. If subjects perform poorly in a simple task, it is doubtful that they could perform better in more complex tasks involving more alternatives where costs and benefits are more difficult to evaluate and compare.

These results, however, could have been easily anticipated because the task was unfamiliar, and therefore complex to the subjects. If any suboptimal, systematic decision rules were attempted, they were developed on the spur of the moment, and they were bound to be no better (and possibly worse) than simple trial and error information evaluation. The information processing requirements for finding the optimal solution (even when the type of decision maker is known with certainty) is on the order of minutes of computation time for a digital computer. A human subject in an experimental setting probably could not perform this task in a reasonable time unless, by random chance, the optimal solution was selected on an early trial.

Since accountants are characterized as information evaluators, an alternative to previous research designs is to examine the performance of professional accountants in problems with which they are familiar; i.e., they encounter
routinely. Application of the normative model is still impossible for the subjects. But a subject who is performing a familiar task is more likely to use decision rules believed to be efficacious either (1) because he/she was taught the rules and believed in their validity because of source authority and reliability, or (2) because he/she has accumulated enough previous experience to be confident that the decision rule is the best.

This dissertation also improves upon previous research designs by attempting to determine what data subjects utilize and how that data is processed in making the IS choice decision in addition to determining the quality of subjects' performance of the task. If subjects show significant convergence on the optimal sample size over time, why would such findings differ from the results of previous research? Perhaps certain types of redundant or "irrelevant" data are a necessary input for a successful heuristic even though not utilized by the normative model. If subjects perform poorly relative to the normative criteria, as might be expected from prior research, information pertaining to subjects' cognitive processes can be used to estimate the long-run opportunity costs of using suboptimal decision rules instead of the normative model. If these costs are large, new variations of the rules might decrease the opportunity costs of using suboptimal rules.
The main purpose of this research is to gain some understanding of the behavioral aspects of information system choice which are not adequately captured by mathematical formulations of the problem. The primary area of interest is the cognitive processes of accountants making an information system choice decision; the subjects' decisions may also be compared to the solution prescribed by a normative model. No model is developed, however, to predict subject behavior.

A laboratory experiment consistent with these purposes was constructed where subjects were asked to make information system selections over fifty trials. A relatively simple decision problem was chosen for which the solution can be computed using information economics formulas. The task in the laboratory experiment, however, was stated to depict as nearly as possible a realistic auditing problem.

Protocols were collected from the subjects during the performance of the experimental task. The protocols were analyzed to test several hypotheses of interest. Some tests were formal, statistical procedures; others were not statistical. Both yielded informative results. The hypotheses tested were primarily concerned with the following three issues:

1. Was the subject able to converge on the optimal sample size over time?

2. What information was used by the subject?
3. How does the subject process the information?

Tversky and Kahneman (1974) hypothesize that subjects utilize heuristics to shorten and simplify the processing requirements of complex judgment tasks. The application of these heuristics could lead to extremely biased judgments in some tasks while they may be reasonable approximations of the optimal solution in other tasks. For this reason, the experimental task was also simulated to investigate the efficacy of a "pure" anchoring and adjustment heuristic in locating the optimal sample size.

The remainder of this chapter overviews the normative solution of the information system (IS) choice problem, the major findings of the study, and the organization of the remaining chapters.

1.2 OVERVIEW OF THE NORMATIVE SOLUTION

A discussion of normative models is important for at least two reasons. First, examination of the normative model motivates the investigation of several behavioral issues. Second, a quantitative abstraction of the problem is necessary for simulation purposes, and relatedly, in order to determine an optimal information system for both the simulations and the experimental task.

Man, when making economic decisions, should be an expected utility maximizer (Harschak and Radner, 1972). An optimal choice among alternative courses of action then is
found by computing the expected utility for each possible action and choosing the action with the highest expected utility.¹ The following equation describes the choice situation for an expected utility maximizer.

\[(1) \quad E(U|a*) = \max \{ \sum (U(a_s) \cdot P(s)) \} \]

This computation requires the decision maker (DM) to specify all of the following:

1. An exhaustive list of all possible actions, \(a \in A\)
2. A partition of all possible states of the world, \(s \in S\)
3. An outcome value for each action/state pair in \(A \times S\)
4. A utility function on outcomes, \(U(a_s)\)
5. A subjective probability distribution, \(P(s)\), over all states, \(s \in S\).

The DM often has the opportunity to purchase an information system before choosing the expected utility maximizing action, \(a^*\). An information system is a set of signals; the receipt of one of the signals will cause the DM to revise his probability distribution over the possible states of the world, \(S\), which may in turn cause him to choose an action different from the action which would have been chosen without the information system. Since numerous information systems may be available at different levels of

¹ See Appendix A for a description of notational conventions used throughout the paper.
cost and informativeness, a formidable problem of selecting the optimal course of action without information becomes even more complex when the opportunity to purchase an information system is available. An optimal information system must now be chosen before an optimal action is taken.

Information system choice is a decision within a decision. Since two decisions are made in sequence, possibly one party can choose the information system, and a second party can choose the action upon receipt of the signal from the chosen information system. In this organizational setting, the accountant's role is to select an information system among a set of systems varying in informativeness and cost, and to report the signal from the system to a decision maker (Sundem, 1979).

Insights into the decision making process and issues involved in the design of an experiment in the domain of accounting and the task environment of information system choice may be gained through the solution of a simple problem. Uecker (1978) used such a simple problem in his laboratory experiment, described more fully below.

In Uecker's experiment, each of 10 urns contain 100 black and white marbles in four varied proportions. Two urns contain 90 black and 10 white marbles, four urns contain 70 black and 30 white marbles, three urns contain 50 black and 50 white marbles, and one urn contains 30 black and 70 white marbles.
There are two parties to the experiment: Q and R. For each trial, Q chooses a sample size from zero to three. An urn is chosen at random and is sampled from without replacement. Both Q and R are told the number of black balls in the sample and R must guess which urn the sample was drawn from. If R guesses correctly, both Q and R receive a reward of 3 cents minus a charge of one cent per ball sampled. If R is incorrect, both Q and R lose 3 cents in addition to the sampling charge. The experiment is repeated for as many trials as desired. What is the sample size which maximizes expected payoff to Q (and consequently to R)?

To develop the solution to this problem, Q and R are first treated as one (call him or her QR), with no opportunity to sample. The solution to QR's problem is given by:

\[
E(\$|a^*) = \max_{a,s} \text{SUM}(\$*P(s))
\]

where:

- \(a\) is the urn type guessed
- \(s\) is the urn type which obtains
- \$ is the net monetary payoff (inclusive of sampling costs as applicable below).

Next, assume QR is given a sample size, \(n\), and is told the number of black balls in the sample, \(y\). His problem is now:
To select an optimal information system, $n^*$, $\mathcal{A}$ determines the best action (similar to equation (2)) for each possible $n$ and $y$ combination. For each $n$, the expected value given the best action is computed. The sample size which has the highest expected value is chosen. This requires solution to the following:

$$E(s|a^*) = \max \left( \sum_{a} \max \left( \sum_{s} \sum_{y} \mathbb{P}(s) \mathbb{P}(y|s,n) \right) \right)$$

Since:

$$P(s|y,n) = \frac{P(y|s,n) \mathbb{P}(s)}{P(y|n)}$$

by Bayes' rule equation (4) can be restated as:

$$E(s|n^*) = \max \left( \sum_{n} \max \left( \sum_{y} \sum_{a} \mathbb{P}(s) \mathbb{P}(y|s,n) \right) \right)$$

Finally, the assumption that $\mathcal{Q}$ and $\mathcal{A}$ are one person is eliminated. Since $\mathcal{Q}$ will choose the information system, the one chosen will presumably maximize $\mathcal{Q}$'s expected value to the problem. This requires that Equation (6) be augmented by $\mathcal{Q}$'s probability distribution over $\mathcal{B}$'s possible urn type choices:

$$E(s|n^*) = \max \left( \sum_{n} \max \left( \sum_{y} \sum_{a} \mathbb{P}(s) \mathbb{P}(y|s,n) \mathbb{P}(a|y,n) \right) \right)$$
Since the separation of information system choice from decision maker action choice in an accounting context is the primary motivation, the components of the urn ball problem can be related to more complex problems in accounting. Q, as the accountant, provides the decision maker, R, with information in the form of reports, y. These reports may be lengthy, formal financial statements with accompanying notes or simply a short exception report of variances between budgeted costs and actual costs incurred. The decision, a, is made by R, and the actual state of the world, s, is to be observed sometime in the future.

The information system, n, may represent numerous dimensions such as a period of time over which data is collected, the number of transactions sampled, or the application of an accounting principle. For instance, in measuring the assets of the firm, the accountant utilizes verifiable historical cost information and ignores other information such as changing market conditions which may be more relevant to decision making, but more costly to obtain. Theoretically, the basic structure of all accounting problems where the accountant and decision maker are separate entities may be described by the mathematical model.

An example which conveniently lends itself to experimental testing and which can be most directly related to the urn-ball experiment is sample size selection for
compliance testing in auditing. The mathematical model, equation (7), can be applied directly with minimal sacrifice in the realism of the problem statement.

Generally accepted auditing standards require that compliance testing should help determine the level of substantive testing to be conducted (AICPA, 1973). The stronger the perceived level of internal control, the less extensive need be the substantive tests (SAS No. 1, Sec. 315.06). How extensive a test of internal control procedures should be undertaken to determine the effectiveness of the firm's internal control procedures? For the experiment, a specific aspect of compliance testing could be investigated, for example, the tracing of cash disbursements through the purchases/payables internal control system. A junior accountant (Q in the example) samples from the past invoices, and his supervisor (R) must decide on the "true" level of the internal control procedures based on the number of sampled transactions which did not comply with internal control procedures.

The urn-ball experiment and equation (7) suggest a number of issues which lend themselves to empirical testing. In an experiment where y and Y, s and P(s), n and H are all well specified, a number of issues remain:

1. Can the accountant assess P(y|s,n) properly? That is, does he assess the probability of the sample result?
2. Can the accountant assess $P(\alpha | y, n)$ properly? That is, does he know the decision maker's model?

3. Does the accountant maximize expected value? That is, (a) does he act as though he had a linear utility function and maximize it, (b) maximize a non-linear utility function, or (c) merely "satisfice"?

The literature review in the following chapter suggests some tentative answers to these questions and provides motivation for the proposed experiment.

1.3 **OVERVIEW OF FINDINGS**

This dissertation makes two methodological contributions to the accounting literature. First, the experimental task is the first to approach an actual task an accountant might perform, while still allowing for the calculation of optimal behavior. Second, the analysis of the data obtained from the experiments relies heavily on both protocol analysis and statistical methods. This type of analysis has appeared infrequently in the accounting literature.

The results of the field experiment corroborated previous information system choice research in that most subjects were unable to converge on the normative solution over time. This occurred in spite of the enhanced external validity of the laboratory experiment in relation to the
task environment. In addition, the cognitive processes of subjects, as represented by the verbal protocols, bear almost no resemblance to the processing required to find the optimal solution using the normative model.

There is strong evidence that at least three of the seven subjects applied the anchoring and adjustment heuristic. However, the computer simulation of this heuristic did not yield information with respect to its efficacy for the particular problem used when parameters of the heuristic or decision makers were varied.

Analysis of the protocols provided insights into auditors' usage of financial statements and flowcharts of internal control systems. For financial statements, the basic types of processing involved identification of significant changes in a financial statement item from one year to the next, and the evaluation of a financial statement item based on the subject's conception of what a prototypical firm would disclose in its financial statements. This result was expected given Bouwman's (1980, 1981) research on financial statement analysis with professional analysts as subjects. The subjects utilized the flowcharts to gain overall familiarity with internal control procedures and to identify strengths and weaknesses in the system.
1.4 **Organizational of Remaining Chapters**

Chapter II is a review of previous empirical research in decision making in the behavioral psychology and accounting literatures. In addition, two prevalent methods of studying decision processes, the Lens model and verbal protocol analysis, are compared for the purpose of explaining why verbal protocol analysis was judged a superior method of analysis given the motivation of this study. The discussion of protocol analysis includes a review of the information processing theory upon which it is based and the empirical evidence pertaining to the validity of the technique. The literature review concludes with a description of the problem of internal control evaluation by auditors; the appropriate standards promulgated by the American Institute of Certified Public Accountants (AICPA) are cited and internal control evaluation procedures as described in popular auditing textbooks are reviewed.

The research design is discussed in Chapter III. This includes a detailed description of the procedures used, an exposition of all hypotheses, and a discussion of the numerous design decisions made with supporting rationales.

Chapter IV contains the results of the experiments from both computer simulation and human subject performance of the auditing task and the interpretations of these results. Concluding remarks are made in Chapter V.
Chapter II
LITERATURE REVIEW

Decision processes and decision quality in the area of information system choice were investigated in this dissertation. An experimental task in the domain of auditing, specifically sample size selection for compliance testing of internal control procedures, was investigated. To interpret the data from such an experiment properly requires acquaintance with both the task of internal control evaluation and possible methods of designing an experiment therein.

The first section of the literature review has the purpose of acquainting the reader with the internal control evaluation function. In the next section, two methods of assessing decision rule characteristics are examined: the lens model and process tracing. The latter methodology is better suited to answering the questions posed in this research. Expert judgment research and the issue of optimality criteria in evaluating human performance is discussed in the third section. Literature on the use of heuristics as decision rules will be examined in the fourth section since evidence is available that certain heuristics are utilized by subjects in both general probabilistic judgment tasks as well as accounting tasks. To conclude the
second section, a method which models conservative probability revision with high predictive ability, but which does not explain actual processes, will be discussed. This will prove useful in the computer simulation of the laboratory experiment in addition to heuristics. The fifth section will review studies in accounting which are either directly related to the topic of information system choice or utilize verbal protocol analysis as a methodology.

2.1 INTERNAL CONTROL EVALUATION

Normative models of the audit process are available from two major sources: mathematical representations of the problem, and official pronouncements of the American Institute of Certified Public Accountants (AICPA). Mathematical representations have already been discussed in Chapter I. Their role in this dissertation is to provide a benchmark for subjects' performance of the experimental task.

Official AICPA pronouncements are published as a series; each member of the series is called a Statement on Auditing Standards (SAS). Together, they comprise Generally Accepted Auditing Standards (GAAS). Since an auditor's training involves the learning of procedures which are designed to meet GAAS, a knowledge of the relevant GAAS and audit procedures is important in analyzing auditor's actual behavior in the experimental task (Biggs and Mock, 1979).
Summaries of actual audit procedures used may be found in auditing textbooks and scholarly research on audit procedures.

The first two standards of field work state:

1. The work is to be adequately planned and assistants, if any, are to be properly supervised.

2. There is to be a proper study and evaluation of the existing internal control as a basis for reliance thereon and for the determination of the resultant extent of the tests to which auditing procedures are to be restricted. (SAS No. 1, Para 150.02)

Both standards are directly relevant to the experimental task. The first standard refers to planning, and this can be taken to imply a number of functions. In the mathematical model, all the possible payoffs and relevant probability distributions are prespecified before the actual sample size is selected. If a statistical sampling plan is to be used, the auditor must prespecify a confidence level, an upper confidence bound, and develop a prior estimate of the population error rate. The choice of the sample size itself is part of the planning function.

The second standard relates to the development of an internal control flowchart for a particular firm (and the experimental case), generation of sample outcomes by the computer, and the evaluation of the sample outcome by the decision maker. The purpose of this evaluation "...is to establish a basis for reliance..." on the system of
internal controls "...in determining the nature, extent and timing of audit tests to be applied in ... examination of the financial statements." (SAS No. 1, Para. 320.06).

The process of study and evaluation of a system of internal control is accomplished in two phases: review of the system and tests of compliance. (SAS No. 1, Para. 320.50). This two-stage process is outlined below.

Arens and Loebbecke (1980, Chapter 7) suggest that evaluation of a client's internal control system begin by dividing the overall system into a series of transaction cycles. A typical taxonomy of transaction cycles would include the following major cycles:

1. Sales and collections
2. Payroll
3. Acquisition and payment of goods and services
4. Inventory and warehousing
5. Capital acquisition and payment

Arens and Loebbecke also state that an internal control system has eight objectives of recorded transactions relevant to the audit process. These are: (1) reasonable, (2) valid, (3) properly authorized, (4) include all existing transactions, (5) properly valued, (6) properly classified, (7) recorded at the proper time, and (8) properly included in subsidiary records and correctly summarized. (p. 216)

These objectives are met when the following six elements are present in each major cycle in the internal control system:
1. Competent, trustworthy personnel with clear lines of authority and responsibility

2. Adequate segregation of duties

3. Proper procedures for authorization

4. Adequate documents and records

5. Physical control over assets and records

6. Independent checks on performance

An internal control system must first be summarized in order that the system may be evaluated using the above criteria. This summary may take the form of a narrative, a flowchart, or scoring of a questionnaire.

Compliance testing is the process of testing each control that is to be relied upon to determine if the control is effective, i.e., it is one way of making an independent check on performance. Using the results, an auditor then infers a degree of effectiveness for the system.

Sampling past transactions is one method of performing tests of compliance. As officially defined, "Audit sampling is the application of an audit procedure to less than 100 percent of the items within an account balance or class of transactions for the purpose of evaluating some characteristic of the balance or class" (SAS No. 39, Para. 1). The use of a sample as opposed to a procedure which examines 100 percent of a class of items exposes the auditor
to "sampling risk", a component of "ultimate risk". Ultimate risk exposure arises from the possibility that material misstatements of account balances and other required disclosures exist, and these errors have not been discovered and corrected by the auditor (SAS 39, Para. 8). The sampling risk component of ultimate risk "...arises from the possibility that, when a compliance or a substantive test is restricted to a sample, the auditor's conclusions may be different from the conclusions he would reach if the test were applied in the same way to all items in the account balance or class of transactions" (SAS 39, Para. 10).

Tests of compliance may be applied on either a subjective or on a "statistical" basis (SAS No. 39, Para 3). Statistical formulas may be used for the determination of sample sizes and/or the testing of formal hypotheses.

Figure 1, from Arens and Loebbecke, illustrates the placement of the study and evaluation of internal control in the overall audit process.
Figure 1—Overview of the audit engagement (Arens and Loebbecke, 1980)
2.2 **DECISION PROCESS METHODOLOGIES**

The objective of this research was to investigate decision processes in the area of information system choice. Libby and Lewis (1977) classify possible process variables of interest in their review of human information processing in accounting as:

A. Judge Characteristics
   1. Human-Mechanical
   2. Number of judges
   3. Personal characteristics
      a. Intellectual ability
      b. Personality
      c. Cognitive structure
      d. Attitudes
      e. Demographics (e.g., age, sex)
   4. Task related characteristics
      a. Prior experience–stored information
      b. Interest and involvement

B. Characteristics of Decision Rule
   1. Form (linear, configural, compensatory, etc.)
   2. Cue usage (weighting)
   3. Stablility (change-learning)
   4. Heuristics

All subtopics under "Characteristics of Decision Rule" are germane to this research, and therefore are included in this section of the literature review. The issue of cognitive structure will also be examined, but not in the same sense as Libby and Lewis intended. The psychology literature treats cognitive structure as a logical modelling of human intelligence common across all individuals (similar to human physiological traits), not as a set of personal characteristics which varies across humans. The validity of certain theories of cognitive structure are crucial to a
justification of protocol analysis as a research tool because it is necessary to establish that certain classes of verbalizations represent at least a partial trace of thought processes.

2.2.1 Lens Model

Numerous descriptions of Brunswik's lens model can be found in both the accounting and psychology literatures (e.g., Slovic and Lichtenstein, 1971; Libby and Lewis, 1977; Amey, 1979). To minimize redundancy, its main features will be described briefly.

Brunswik viewed the individual as one who functions in an environment rich in stimuli (cues). Behavior was seen as vicarious functioning in response to noisy cues (random error exists). Through his model, one seeks to measure the statistical relationships between the individual, the cues, and the environment. Actual cognitive processes are considered to be too complex for direct analysis and "lie hidden within an inaccessible 'black box'." (Hogarth, 1974, p. 298) The interrelationships of cues to subjects and the environment are usually depicted as shown in Figure 2:
In a lens model study, decision makers make a large number of decisions and CLS models are constructed as paramorphic representations. Several models have been used:

\[ Y^s(e) = b(1s)^*C(1) + b(2s)^*C(2) + \ldots + b(ns)^*C(n) \]

where \( Y^s(e) \) is the regression estimate of the true environment or criterion value, and subscripts are denoted by parentheses.

\[ Y^s(s) = b(1s)^*C(1) + b(2s)^*C(2) + \ldots + b(ns)^*C(n) \]
where $Y'(s)$ is the decision maker's predicted response. The following correlations are also computed:

$r(Y(e), Y'(e))$ - a measure of the predictability of the environment

$r(Y(s), Y'(s))$ - a measure of subject's consistency

$r(Y(e), Y(s))$ - a measure of subject's achievement

$r(Y'(e), Y'(s))$ - a measure of task knowledge

The advantages of using the lens model as a method for interpreting experimental data lie in its ability (1) to concisely and objectively summarize the characteristics of the decision process using statistical relationships, and (2) to predict subjects' responses with a parsimonious, quantitative model. The disadvantages are the loss of model validity when (1) cues are combined by the decision maker in a different way than the statistical model is constructed or (2) cues are utilized by the subject which were not included in the model. Specifically, a linear regression model on the cues requires the use of an additive compensatory decision rule, where all the available cues from a parsimonious cue set are considered for each decision.

A description of the lens model has been included because it represents a significant methodology which we chose to eschew for use in this study, even though it has been a widely used tool by accounting researchers (Libby and Lewis, 1977). Psychologists have accumulated a significant body of evidence to support the use of other decision rules as well. First, the goals of this dissertation do not
include (1) the development of a model predicting subjects' decisions or (2) a cue set for the experimental task which can be parsimoniously specified. Second, a description of auditor cognitive processes at a more detailed level than can be attained through the application of the lens model is desired.

2.2.2 **Process Tracing Models**

Process tracing research uses methods which allow collection of data on cues attended to by subjects during the performance of the experimental task. Process tracing research also provides data which allows comparison of the human cognitive processes during performance of the task with algorithms used in computer simulations of the task. One process tracing method is verbal protocols. A verbal protocol is a record of verbalizations. The subject's utterances while he/she performs the task are tape recorded and then transcribed into a series of consecutively numbered, short phrases. Protocol analysis is the process of data with the goal of determining the subject's cognitive processes.

In order to interpret verbal protocol data, a behavioral theory of human problem solving is necessary. Newell and Simon (1972) present a general theory of problem solving. Three types of problems, cryptarithmetic, logic and chess were used to empirically test the applicability of
their theory in specific task environments. The analyses of the three tasks serve as prototypes for investigation of other task environments.

The entire theory is extremely detailed, as evidenced by the fact that Newell and Simon's (NS) development of the theory consumes 900 pages, more than two-thirds of which are devoted to a description and analysis of the three specific task environments mentioned above. The three assertions below cover a substantial portion of the theory derived from this analysis:

1. Humans are examples of information processing systems (IPS).

2. Problem solving takes place in a problem space.

3. The IPS solves problems by evoking programs from memory to transform information currently part of the problem space to a new state of knowledge.

Humans as IPS's - "An IPS is a system consisting of a memory containing symbol structures, a processor, effectors and receptors" (NS, p. 20). As an instance of an IPS, a human IPS contains all these elements common to all IPS's. Additionally, there exist characteristics of a human IPS which are more or less invariant across humans.
The human IPS is a serial processor in the sense that it acts as if it can only execute one elementary operation at a time. The processing rate of the human is dependent on the rate of access to memories as well as the speed of execution of an elementary operation.

Information can be stored in one or more of three types of memory. Short-term memory (STM) has a very small capacity (between five and nine "chunks" of information (NS, p. 795)), and information is combined with elementary processes in short-term memory to produce new information. As a result, most information has to be retrieved from long-term memory (LTM) before it can be processed.

The capacity of LTM is very large, and "there is no evidence that the human LTM is fillable in a lifetime" (NS, p. 792). Access times, or "read" times, are considerably longer for items stored in LTM than for STM. STM retrieval time is almost instantaneous (NS, p. 795) and LTM retrieval can range from a few hundred milliseconds to one second, depending on the task (NS, p. 793). External memory (EM) consists of information displays outside of LTM and STM such as paper, a chessboard, or a hand-held calculator. Consideration of external memory is important to the study of problem solving for at least two reasons. First, STM can be redefined to be internal STM plus the portion of EM which is in foveal view (NS, p. 801). Second, subjects given the use of EM behave quite differently than subjects who must
solve a problem without the aid of EM. For example, the problem of multiplying two four-digit numbers without pencil and paper takes approximately 100 times longer than if it were allowed to be used (NS, p. 800).

Perhaps the most important consideration to this research is the relationship of the knowledge possessed by the information system to the task environment, or state of the world. Figure 3, from Bobrow (1975), summarizes much of the issues of representation discussed in Newell and Simon.

![Diagram](image)

**Figure 3 - Representation and Mapping (Bobrow, 1975)**

Through a selective mapping, M, a representation of the world is constructed. New representations are constructed by applying operations ("programs" to Newell and Simon) to the current representation (knowledge state) as the state of the world is altered through actions.

NS develop very rigorous and formal representations of the tasks they investigated, but accounting researchers give little attention to this aspect of the theory. One notable exception is Bhaskar and Dillard (1978) who propose and test
the adequacy of their representation of accounting knowledge. Other studies do not discuss the subject's representation of the task environment in great detail because a useful representation in tasks as well-defined as those studied by Newell and Simon or by Bhaskar and Dillard is difficult to specify.

**The Problem Space** - All subjects over the range of tasks examined by NS utilized a problem space. NS contend that sufficient evidence is available to postulate that the use of a problem space in problem solving is a major invariant across individuals for complex, non-automatic, problem solving tasks. The problem space is simply the space where problem solving takes place, and it contains all the possible solutions a problem solver might consider. Formally, a problem space consists of:

1. A **set of elements**, \( U \), which are symbol structures, each representing a state of knowledge about the task.

2. A **set of operators**, \( Q \), which are information processes, each producing new states of knowledge from existing states of knowledge.

3. An **initial state of knowledge**, \( u(0) \), which is the knowledge about the task that the problem solver has at the start of problem solving.

4. A **problem**, which is posed by specifying a set of final, desired states \( G \), to be reached by applying operators from \( Q \).

5. The **total knowledge available** to a problem solver when he is in a given knowledge state, which includes (ordered from most transient to most stable):
a. Temporary dynamic information created and used exclusively within a single knowledge state.

b. The knowledge state itself—the dynamic information about the task.

c. Access information to the additional symbol structures held in LTM or EM (the extended knowledge state).

d. Path information about how a given knowledge state was arrived at and what other actions were taken in this state if it has already been visited on prior occasions.

e. Access information to other knowledge states that have been reached previously and are now held in LTM or EM.

f. Reference information that is constant over the course of problem solving, available in LTM or EM. (p. 810)

The problem space for the experimental task used in this study is specified in the next chapter.

Programs — A program is a collection of elementary processes and is applied to elements of the problem space for the purpose of achieving a new knowledge state. This concept of a program suggests that problem solving is achieved by setting and attaining goals. Therefore, analysis of problem solving behavior can take place on two levels: (1) identification of the goals and sub-goals set by the subject (macro-level) and (2) identification of the operators or programs used to attain those goals (micro-level). A researcher's ability to access the elementary operations is limited to what the subjects verbalize as previously noted. Highly familiar programs (for example, multiplying two single digits) are automatic
to the subject and such a program usually is not verbalized by an adult subject.

Einhorn, Kleinmuntz and Kleinmuntz (EKK) (1979) show how the linear regression and the process tracing approaches to modelling judgment can be combined in a single study to yield new insights through comparison. Although the models are outwardly dissimilar (one is statistical in nature and the other attempts to focus directly on the actual cognitive processes), EKK claim that the important differences are the models' "emphasis and descriptive level of detail". To support this assertion, EKK demonstrate that process tracing can be generated from a general additive rule. The results of two experiments are presented which combined both models of judgment. Two experiments were conducted in which the predictive ability of the process model and the linear regression models were compared. In the first experiment, a LRM of the process outpredicted the computer model developed from the protocols. To explain this result, EKK took into account the high familiarity of the task to the subject and hypothesized that certain processes were not verbalized by the subject because they occurred automatically. A second experiment was constructed with a less familiar judgment task—evaluating the nutritional value of breakfast cereals—and both models made similar predictions with the computer model being slightly more accurate.
Linear regression models (LRM) of judgment are successful for prediction, EKK state, because they capture the basic properties of "vicarious functioning." This ability to make tradeoffs and to use cues interchangeably is one of the foundations of human behavior.

LRMs capture the basic process of vicarious functioning in at least four important ways. First, the additive combination function implies a fully compensatory system. Second, the degree to which cues trade-off, which is determined by the beta weights in the regression equation, is a function of the specific environment in which the judgment is made; that is, the beta weights are determined by considering all of the cues and their particular levels in the situation. Third, cue redundancy is incorporated in the model since the beta weights are determined by the correlational structure of the cues (and the correlation of cues with the judgment). Moreover, the indeterminancy in estimating weights when cues are correlated parallels the organism's difficulty in this matter. Also, redundancy is defined with respect to the particular environment being studied. Fourth, the inconsistency and random error in judgment, resulting from the lack of cognitive control in executing one's strategy is explicitly defined and measured within regression procedures. (EKK, pp. 467-468)

However, this alone does not explain why many experiments have found judgments to be linearly predictable even when the environment is unpredictable. EKK hypothesize that the use of two heuristics may explain this phenomenon. In one heuristic, the sign relationship of cues to judgment is assumed to be independent of the level of the other cues. This is consistent with many observed physical rules learned early in life such as volume with length, height with width. The second rule of thumb is that the tradeoff of cues is
treated as independent of the level of other cues despite cue redundancy (correlation of cues) and that tradeoff rates may be a function of different levels of other cues.

Although the two heuristics of conditional monotonicity and trade-off independence imply an additive model, it is important to emphasize that vicarious functioning as we have described it is a cognitively complex and sophisticated strategy. Therefore, criticisms that LBMs of judgment are too simplistic to capture the highly interactive and contingent methods that are believed to be used in judgment rest on the erroneous belief that the mathematical form of a model is isomorphic with the process it is supposed to represent. Ironically, from our perspective, a highly interactive and contingent process is already captured in LBMs. (EKK, p. 469)

Two experiments were conducted in which the predictive ability of the process model and the linear regression models were compared. In the first experiment, a LBM of the process outpredicted the computer model developed from the protocols. To explain this result, EKK took into account the high familiarity of the task to the subject and hypothesized that certain processes were not verbalized by the subject because they occurred automatically. A second experiment was constructed with a less familiar judgment task—evaluating the nutritional value of breakfast cereals—and both models made similar predictions with the computer model being slightly more accurate.

A similar type of complementary analysis would be desirable in the present study. Both of EKK's two experiments, however, were conducted in domains where the cue sets are limited and well-specified; EKK's goal was to
compare the predictive abilities of the two classes of models. As will become apparent in the next chapter, there are too many cues to include in a regression model, and the hypotheses tested are not concerned with the prediction of individual subject responses, the most powerful aspect of linear regression modelling.

Verbalization of mental behavior for the purpose of studying cognitive processes has frequently been scrutinized. Nisbett and Wilson (1977) reviewed numerous behavioral experiments in addition to conducting several new experiments to determine subjects' ability to post-experimentally report their higher order cognitive processes. Results of these experiments led Nisbett and Wilson to the following three conclusions:

1. People cannot often report accurately on the effects of stimuli on their decision processes "the accuracy of subjective reports is so poor as to suggest that any introspective access that may exist is not sufficient to produce generally correct or reliable reports" (p. 233).

2. "When reporting on the effects of stimuli, people may not interrogate a memory of the cognitive processes that operated on the stimuli, instead, they may base their reports on implicit, a priori theories about the causal connection between stimulus and response." (p. 233)

3. The correct reports of cognitive processes are most often due to the coincidental application by the subjects of their a priori causal theory which explains their behavior and are not due to introspective awareness.

Although their work is not principally directed toward "methodological issues", Nisbett and Wilson advise that
techniques such as protocol analysis may be quite misleading based on their evidence.

Ericsson and Simon (1980) point out, however, that Wisbett and Wilson's criticisms of protocol analysis are too broad since their review is only concerned with the informativeness of post-experimental verbalizations. Concurrent verbalizations represent a significantly different methodology from post-experimental verbalizations. The concurrent/post-experimental dichotomy of verbalization is still not sufficient for Ericsson and Simon; they further subdivide concurrent verbalizations into two classes and cite three types of criticisms of concurrent verbal protocol analysis:

1. **Effect of verbalization** - The act of verbalization alters thought processes.

2. **Incompleteness** - The subject may not verbalize a significant amount of his thoughts, and the protocol "will fail to track the actual path of the activity as revealed by other observations, or inferred from theory." (p. 4)

3. **Irrelevance** - Verbalizations are a separate cognitive process which occur independently of the activity of interest.

Ericsson and Simon studied the validity of the first criticism, because it is the only criticism which would render the methodology useless for researching cognitive processes.

Three levels of verbalization occurring in laboratory experiments are defined. In Level 1 verbalization, the
subject is asked to reveal only information from short-term memory (STM) which has been previously orally encoded. Level 2 verbalization is also concerned with STM, but all information is spoken, whether orally encoded (a number, for example) or not (an odor, for example). Level 3 verbalization relies primarily on long-term memory. Subjects are requested to explain the motivation for their thoughts. This type of verbalization may occur during the performance of the task or after task completion. Nisbett and Wilson's research was limited to experiments where the subject was requested to be introspective after the task was completed. Thus,

The experiments reviewed by Nisbett and Wilson provide valuable information on incidental memory (or lack of it) for heeded information, but have no bearing at all on the interpretation of protocols obtained under standard thinking aloud instruction... if a methodological lesson is to be drawn from the Nisbett-Wilson paper, it is that one shouldn't ask subjects why [emphasis in original], but should simply use verbalization instructions to discover what information is attended to or is stored [emphasis in original]. (Ericsson and Simon, p. 43)

Ericsson and Simon hypothesize that Level 2 verbalization has no effect on cognitive processes and Level 3 verbalization does, both by construction of models of the verbalization processes and from a review of empirical evidence. The set of empirical studies cited seems to confirm their hypothesis. There are some effects from verbalization; for example, subjects who are instructed to verbalize take significantly longer to solve the same
problem, with however, no effect on performance measures. The instructions in the experimental task of this dissertation had the objective of eliciting Level 2 verbalization from the subjects.

2.3 EXPERT JUDGMENT AND OPTIMALITY

The experiment utilizes experienced auditors to perform a task with which they have had some experience. The task is modified in that some variables (payoffs and prior probabilities) are explicitly stated. This moderate change in the environment enables evaluation of subjects' performance relative to the normative model, and determination of the extent that subjects make use of the new variables. Consequently, literature dealing with judgment by experts and discussing the reasonableness of comparing subjects with normative criteria is appropriate.

Einhorn (1974) provides some necessary, if not sufficient, conditions for defining expertise:

1. Experts should tend to cluster variables in the same way when reducing the dimensionality of the cue set.
2. There should be high intrajudge reliability.
3. The cues utilized by experts "should have status as explanatory concepts" (p. 563).
4. Expert judges should show low amounts of biases.
5. Experts should weight and combine cues in similar ways.
The application of these concepts was demonstrated in an experiment where three medical pathologists were asked to measure specific cues and to form a diagnostic judgment of the degree of severity of Hodgkin's disease in 193 cases. The design of the experiment and the methodologies applied to analyze the data were consistent with the lens model of human information processing. Einhorn recognized that a completely different type of experimental setting (such as this dissertation) may examine significantly different aspects of expertise. For example, a high performing judge may discern or form cues that no other subject has ever seen before; or, he may consider actions that other judges overlook. The true expert may also be distinguished by the efficiency and method of his information search processes. These variables cannot be incorporated in a lens model study. A process tracing methodology can address such questions.

A major criticism of the conditions for expertise stated in the present context may be that too much stress has been put on agreement. This agreement has been with respect to clustering, measuring, and weighting cues. As is well known, the history of science is replete with oddballs who did not agree with anyone, yet, were proved to be correct by subsequent events. This criticism has value, yet it conceals the fact that eventually some relevant criterion other than agreement became available... We can say that in a highly probabilistic world, there may be many routes to the same goal and that there may be more than one way to perform the cognitive tasks involved in judgment (pp. 570-571).
Einhorn and Hogarth (1981) state that if behavior is goal directed, it is reasonable to assume that some ways of attaining a goal are better than others. The evaluation of discrepancies between normative (the "best" way of attaining a goal) and human behavior has raised questions concerning the nature and causes of such discrepancies. When considering the related factors of limited human processing capabilities and misspecification of the normative model used as a benchmark, three possibilities could exist:

1. Both models could inadequately represent the task, but in different ways; 2. the optimal model is a more adequate representation than that of the person. Indeed, this is the assumption upon which most decision research is predicated; and 3. the person's model is more appropriate than the optimal model... (Einhorn and Hogarth, p. 3)

Optimality refers to the maximization or minimization of an explicit, measurable criterion, conditional on specified assumptions about the environment. Two major limitations can be identified for consideration when comparing actual to prescribed behavior for audit sample selection in this dissertation. First, optimal behavior assumes that maximization of fictional cash flows is an adequate representation of the subject's goals in the experiment. Subject judgments and choices may be based on a significantly different single criterion, or on multiple criteria. Second, optimal behavior assumes that the prior probability distribution is an adequate representation of the uncertainty faced by the auditor.
A base rate can only be defined conditional on some population (or sample space). Whereas many might agree that the base rates defined by experimenters in laboratory tasks make the sample space clear, the definition of the population against which judgments should be normalized in the natural ecology is unclear. Consider an inference concerning whether someone has a particular propensity to heart disease. What is the relevant population to which this person should be compared? The population of people in the same age group? The population of the United States? Of Mexico? There is no generally accepted normative way of defining the appropriate population. Thus, for naturally occurring phenomena it is neither clear whether people do or do not ignore base rates, nor whether they should (Einhorn and Hogarth, p. 12).

2.4 DECISION MAKER ALGORITHMS

Having discussed the task domains of auditing and information system choice, and methodologies used to investigate problem solving in a domain, various types of general strategies, or algorithms, humans will invoke to solve a problem can now be considered. One very important type of algorithm is a heuristic. The problem solution resulting from the use of a heuristic may be viewed as the solution to a similar but less complex problem than the problem the subject is asked to solve. Therefore, the power of a heuristic to solve, or to provide reasonable approximations to the solution, of complex problems is non-generalizable.
2.4.1 Heuristics

Tversky and Kahneman (1974) describe three heuristic principles on which their subjects relied during decision tasks involving probability inferences: representativeness, availability, and anchoring and adjustment. Heuristics can reduce the complexity of decision making tasks; they can also lead to large errors in judgment (Tversky and Kahneman, p. 1124).

In applying the representativeness heuristic, "probabilities are evaluated by the degree to which A resembles B" (p. 1124). Representativeness is not influenced by several factors which would enter into judgments concerning the probability that A is drawn from B. These are described by Tversky and Kahneman as follows:

1. Insensitivity to prior probability of outcomes - In assessing P(B|A), P(B) (the prior probability) is ignored when the representativeness heuristic is used.

2. Insensitivity to sample size - When calculating a sample statistic such as the sample proportion, the "judged probability... of that statistic's occurrence will be essentially independent of sample size" (p. 1125).

3. Misconceptions of chance - "People expect that a sequence of events generated by a random process will represent the essential characteristics of that process even when the sequence is short. In considering tosses of a coin for heads or tails, for example, people regard the sequence H-T-H-T-T-H to be more likely than the sequence H-H-H-T-T-T, which does not appear random, and also more likely than the sequence H-H-H-H-T-H, which does not represent the fairness of the coin... Another consequence of the belief in local representativeness is the well-known gambler's fallacy. After observing a long run of red on the roulette wheel, for example, most people
erroneously believe that black is now due, presumably because the occurrence of black will result in a more representative sequence than the occurrence of an individual red" (p. 1125).

4. Insensitivity to predictability - Lack of measurement accuracy in sample information is not given enough weight. For example, a highly favorable report is likely to lead to high earnings prediction irrespective of the report's accuracy.

5. The illusion of validity - Even if information is incomplete, unreliable or outdated, people will use it to make confident predictions when there is a high correlation between the information and the predicted outcome.

6. Misconceptions of regression - "Suppose a large group of children has been examined on two equivalent versions of an aptitude test. If one selects ten children from among those who did best on one of the two versions, he will usually find their performance on the second version to be somewhat disappointing. Conversely, if one selects ten children from among those who did worst on one version, they will be found, on the average, to do somewhat better on the other version. More generally, consider two variables X and Y which have the same distribution. If one selects individuals whose average X score deviates from the mean of X by k units, then the average of the Y scores will usually deviate from the mean of X by less than k units. These observations illustrate a general phenomenon known as regression toward the mean, which was first documented by Galton more than 100 years ago" (p. 1126).

In assessing the probability of a given event, a person may attempt to recall the number of times the event has occurred in the past. Tversky and Kahneman termed this the availability heuristic. Biased probability estimation can occur for a number of reasons when this heuristic is employed. Most important to this research is the recognition that instances which are more easily retrieved from subjects' memory are judged to occur more frequently.
Anchoring and adjustment is a process of making probability revisions by selecting an initial value (via some process) and making adjustments to that value as information and experience is accumulated. Use of this heuristic could lead to three types of biases:

1. Insufficient adjustment - Probability revisions are less than the amount prescribed by a normative revision model.

2. Biases in the evaluation of conjunctive and disjunctive events - People tend to overestimate the probability of conjunctive events (for example, rolling 2 sixes on a pair of dice is a conjunctive event) and underestimate the probability of disjunctive events (for example, rolling at least one six on a throw of two dice).

3. Anchoring in the assessment of subjective probability distributions - People tend to state overly narrow confidence limits than is justified by sample information.

The anchoring and adjustment heuristic was adopted as the algorithm to simulate information evaluators in the simulation portion of the experiment. Details of the algorithm are given in Chapter 3, Section 5.

One way in which the decision maker in the urn-ball problem described earlier could be simulated would be to place reliance solely on the sample proportion from the sample selected in making inferences about urn the sample was drawn from (the representativeness heuristic). Swieringa et al. (1976) used six experiments on graduate business students to test for evidence of bias due to insensitivity to prior probabilities, insensitivity to
sample size, and the illusion of validity (source reliability). While some experiments were replications of those previously performed by Tversky and Kahneman, two experiments used accounting information. The accounting problems were used as a basis for comparison with problems identical in their basic structure, but presented in a different context.

Overall, the experiment yielded evidence that subjects were susceptible to errors associated with the representativeness heuristic although there were some differences from Tversky and Kahneman's earlier findings. In the experiments designed to test insensitivity to sample size, the number of subjects giving incorrect answers to questions was significantly different for the same problem posed in two different contexts (accounts receivable versus urn-ball).

The representativeness heuristic was programmed and could have been applied in both the experimental task and the simulation. It was not utilized because of a feeling that the algorithm was not accurate enough to satisfactorily reflect judgments by audit managers. The results of the study by Swieringa et al. also contributed to the determination during preliminary testing of the interactive program used in the experimental task that the representativeness heuristic was inappropriate for this research.
2.4.2 Mathematical Models with Predictive Ability

As previously described, the objectives of this study do not include prediction of responses. A discussion of prediction is important, however, because of prior research on information system choice (Decker, 1978). The particular model reviewed below was considered as a way to simulate decision makers, but was rejected due to a need to limit the experiment to manageable proportions. Two other models of decision making, the Bayesian model and a variation of the representativeness heuristic found in textbooks on auditing, were deemed more desirable for use in this study because the former has normative appeal and the latter has high intuitive appeal given the type of university and on-the-job training auditors typically receive (as reflected by textbook topics and questions asked on the Uniform CPA Examination).

Phillips and Edwards (1966) analyzed the responses of student subjects in three "container and object" experiments where the subjects were asked to give their estimate of posterior probabilities. The tendency of the subjects in all experiments was to revise probabilities conservatively: that is, the absolute differences between the subjects' prior and posterior probability assessment were less than the amount prescribed by Bayes' theorem. Although no attempt was made to generalize the findings beyond the specific experiments, Phillips and Edwards also observed:
1. Subjects performed in a less Bayesian manner when the difference between red and blue poker chips in the bag was greater.

2. Probability revision became slightly more Bayesian as the data became more diagnostic.

3. Probability revision was unaffected as the sample sizes increased.

4. Conservatism was mitigated when payoff systems were added to the experiment due possibly to their instructional value (comparison of actual payoff to the payoff which would have occurred if the alternative action were chosen) as well as their motivational value. Also, subjects exhibited varying degrees of conservatism depending on the payoff system used.

5. Less between subject variation was observed when a payoff system was used. This was also attributed to the instructional value of payoffs.

In addition to a description and analysis of their findings, Phillip and Edwards proposed a one-parameter model which they feel should capture most of the observed behavior on an aggregated basis:

\[ Z'(1) = (L^c) \cdot Z(0) \]

where:
- \( L \) is the likelihood ratio
- \( Z(0) \) is the prior odds
- \( Z(1) \) is the posterior Bayesian odds
- \( c \) is the conservatism parameter (0 < c < 1)
- \( Z'(1) \) is the posterior conservative odds

Note that when the conservatism parameter, \( c \), is equal to one, Equation (10) models Bayesian revision. The same
algorithm could have been used to simulate decision making based on both bayesian and conservative probability revisions (see Appendix C).

Wright (1978) attempts to extend psychologists' findings from abstract probability inference tasks by examining a similar problem in the domain of accounting. Eleven graduate students in business administration were each given 15 independent cases. Each case consisted of one or two financial statistics for a firm, and each statistic, or cue, was set at 3 levels. The fifteen cases consisted of 6 individual cue observations and 9 (3 x 3) joint cue observations. Each subject provided 75 probability estimates since each posterior distribution was massed to 5 possible values for beta.

The task of the subject was to generate discrete, subjective probability distributions for the firms' beta coefficients. A beta coefficient is a quantitative expression for the sensitivity of a firm's value to general, non-firm specific economic events. The criterion, or environmental, values for the posterior distributions were determined on a relative frequency basis, using a sample of approximately 450 firms. The three levels for each cue were determined in a similar manner.

The analysis of the elicited posterior probability distributions was done on both an aggregate and an individual level. Both analyses considered all 75
observations together, as well as the 30 single cue and 45 joint cues separately.

In the aggregate analysis, Spearman rank correlations between the median posterior probability estimates and the corresponding criterion values were computed for the single cue cases, joint cue cases and the pooled observations. All three computations yielded correlation coefficients significantly greater than zero. Simple linear regression and corresponding plots of the two variables for the pooled median values and the two subsets demonstrated that aggregate subject accuracy was greatest for the single cue observations.

The aggregate data was also studied with respect to the amount of revision of the prior probabilities. The computations suggested that revisions were generally made in the proper direction, but were insufficient, or "conservative".

An investigation of the individual subject results began by first computing the average accuracy of each subject's responses. Only 2 of 11 subjects were more accurate than the average accuracy over all subjects. The subjects were also more accurate for the single cue observations than the joint cue observations. The mean of the joint cue errors was greater than the mean of the single cue errors for all but one subject.
One of the most intriguing results of the study was the poor performance of subjects on joint cue cases relative to their performance on single cue cases. Wright suggests that this may be due to two reasons:

1. Simultaneous evaluations of two cues had to be integrated over a joint cue distribution.

2. The amount of revision required for the joint cue cases was generally greater than for the single cue cases.

Johnson's (1978) primary criticism of Wright's paper is that Wright only analyzed subject responses, and he did not do enough to determine how these responses were generated.

We are becoming increasingly aware that the difference between those who are competent or proficient in a domain of tasks (e.g., chess, medical diagnosis, security investments) and the rest of us, is not their capacity for processing information, but essentially what information they process (p. 14).

Huesman (1978) also believes that less research effort should be devoted to issues such as conservatism in probability inference. Instead, effort should be concentrated on the study of the underlying processes which subjects employ. In addition, he does not agree that the subjects' probability revisions in Wright's experiment could necessarily be classified as conservative:

One can categorize revision into four classes: extreme, conservative, accurate, and wrong. "Accurate" is a revision that is exactly correct; "Wrong" is a revision completely in the wrong direction... "Extrema" is a larger change in the desired direction than needed; and "Conservative" is a smaller change in the desired direction than called for. One can see that 38% of the revisions were extreme; 37% of the revision were conservative;
20% of the revisions were wrong; and 4% were accurate. This is hardly compelling evidence for a general conservatism (pp. 23-24).

The above criticisms of the Wright study are relevant to this dissertation in that they seem to call for a process orientation to problem solving issues in accounting.

2.5 RELATED STUDIES IN ACCOUNTING

The overall research strategy of this dissertation is based on an attempt to integrate the two types of research described in this section. The first subsection presents most of the research in accounting that has employed protocol analysis. As such, the various ways protocol analysis has been applied is more important than the details of the task environments themselves. In the second subsection, studies in the domain of information system choice are reviewed. Here, the priorities are reversed: it is demonstrated that the research methods applied to this particular task environment are not able to explain cognitive processes of an accountant performing the information evaluation function.

2.5.1 Studies Utilizing Protocol Analysis

Most accounting research utilizing protocol analysis as a methodology has studied some aspect of the behavior of subjects while analyzing financial statements. However, studies have also been conducted where protocol analysis has
been used to investigate other accounting related tasks such as auditing and cost variance investigation. The primary objective of this section of the literature review is to describe how protocol analysis has been used by accounting researchers because much of the methodology in this dissertation is influenced by the way verbal protocols have been used in these earlier efforts.

In Stephens (1979), ten bank lending officers each made a lending decision based on financial and supplemental data from one of two fictitious cases, and protocols from the decision process were collected. The protocols included observations by an observer (for example, noting what ratios were computed) as well as the subjects' verbalizations. The two cases were identical (except for a size factor) in all respects except for the variables under investigation. These were:

1. Inventory method (LIFO v. FIFO)
2. Depreciation method (straight line v. double declining balance)
3. Primary line of business (retail v. wholesaling)
4. Secondary business line (real estate v. manufacturing)

The hypotheses tested pertained to differences in loan processing style due to the above treatment effects.

Analysis of the protocols showed that the subjects were primarily concerned with the computation of ratios, identification of significant trends in the ratios, and use
of the ratios for attention directing to non-financial statement sources of information. Stephens noted that there was insufficient evidence to conclude that the median number of ratios computed by each group differed, although "working capital" computations were prevalent for the wholesale company, and were not performed for the retail company.

A stage of the protocol analysis in Stephens, Shank and Bhaskar (1980) involved classifying each phrase of the protocol into one of twelve operators (i.e. "read, calculation, query, etc."). A simple program was written consistent with a single pass, compensatory decision model and was compared with one of the protocols:

"This program works by extracting a component from the component set (in some specific order) and applying operators to this component (i.e., a single pass in some order). When this is done, the next component is extracted and processed, and so on until the component set is exhausted. The result from the total processing of all components is a set of inputs into a decision model which computes the final loan decision" (p. 21).

The analysis showed that the protocol traced a process quite dissimilar from the program.

Bhaskar and Dillard (1978) also argue that a Brunswik lens type of approach to the study of human information processing in accounting addresses issues pertaining to "processing" (weighting of cues and their relationship to criteria values), but not enough is known about how accounting information is represented, what data is used, and how various subprograms act on accounting data. The
particular way chosen by a subject depends on the way the relevant knowledge is represented by the subject" (Bhaskar and Dillard, p. 327).

The first task of their paper was to present a proposed representation for accounting knowledge. The representation consists of five basic elements, as shown in Figure 4:

```
MLI
Pointers

MLI

AVA (1)  AVA (2)  AVA (n)

BP (1)  BP (2)  BP (n)

Statement
Preparation
Program
```

Figure 4—Bhaskar and Dillard's (1978) representation of accounting knowledge

Main Line Items (MLI) are schema level units of knowledge such as "accounts receivable", "inventory", etc. They correspond both to items on a balance sheet as well as to the chapters of a typical intermediate accounting textbook. "A schema is described by the programs it contains together with a collection of rules for recognizing it." (p. 328) MLI pointers are list level items which interface language with the accounting domain. Asset Valuation Algorithms (AVA) are
programs associated with each MLI, and Bookkeeping Programs (BP) are programs which enable journal entries to be made after appropriate values are calculated in the corresponding AVA.

The second task of Bhaskar and Dillard was to select four accounting problems from an intermediate accounting textbook and to construct a task analysis from the general model for each problem. The task analyses purported to serve two purposes:

First, it will tell us whether the framework is generally valid. That is, if it is impossible to analyze a problem using this framework then we are forced to reject it. Second, it will provide specific statements about the requirements of the problem which can then be tested by protocol analysis. (p. 332)

Since an MLI pointer, MLI(s), AVA(s) and BP(s) could be named and specified for each of the four problems, the authors had found no evidence to suggest that their framework was not generally valid.

The task analysis was validated with respect to the second objective, above, through the experimental task. Two subjects, an accounting professor and a student enrolled in an intermediate accounting course, were instructed to solve the four accounting problems previously analyzed, and to verbalize their thoughts while solving the problems.

In general, it was found that both the student and professor executed the problems by first specifying an MLI, and then, an AVA. However, some differences were found in
the execution of the AVA's. The student undertook many unnecessary operations to solve the problem, whereas the professor recognized immediately how to solve the problem in fewer steps. Further work by Stephens, Bhaskar and Dillard (1980) supports this structure of accounting knowledge.

Considerable effort in the area of cost variance investigations has resulted in the identification of optimal solutions to simple cost variance problems. Computer simulations have been employed to approximate the opportunity costs of suboptimal decision rules for these problems. Lewis Shields and Young (1981) attempt to identify the decision rules used by ten MBA student subjects and to quantify the long-run opportunity costs (relative to a formal model) of a particular decision rule. The decision rule used by the subjects was determined via protocol analysis.

Each subject was provided with a job description and was assigned the role of a departmental supervisor with the responsibility of minimizing costs of producing a precision-tooled steel product made with a new experimental tooling machine. The description stated that the machine behavior could be described by a two-state process which, when operating normally, produced a product with a mean of 36 ounces. When the machine malfunctioned, the mean product weight was 40.5 ounces. In either state, the process was normally distributed with a standard deviation of 3.0 ounces. The subject was also advised of the prior probability that a process which was in-control would go out-of-control. Costly investigations, ordered by a subject, would correct any malfunctions, but there would still be a known probability of malfunction the following day. If the machine went out-of-control, it would remain that way unless and until an investigation was ordered. The costs associated with the
consequences of the investigation decision were also provided (Lewis, Shields and Young, p. 9).

All weight observations ("production reports") given to the subject were "nontrivial"; that is, between 36 to 40.5 ounces.

Lewis, Shields and Young determined that nine of the ten subjects used a control chart strategy. A control chart strategy provides for the investigation of all variances that are a specified number of standard deviations away from the in-control mean.

Each version of the case was simulated using a one standard deviation control chart strategy. The control chart strategy was found to be 9X and 4S less efficient than the optimal strategy for the two cases over 5000 simulated trials for each case.

Lewis et. al. has a similar objective to this dissertation in that (1) subjects or simulated subjects' behavior is being compared to the optimal solution to a "realistic" accounting problem, (2) the analysis of decision processes is oriented toward the identification of well-known heuristics and (3) the simulation of a psychological construct is undertaken. The problem statement for the case used by Lewis et. al., however, does not recall a "cost variance" situation since only weight variances—not cost variances—are reported to the subject. The Lewis et. al. study exemplifies the difficulty of devising an accounting problem where an optimal solution is available.
Biggs (1978) researched the use of four types of decision rules by financial analysts: additive compensatory (AC), conjunctive (C), additive difference (AD), and elimination by aspects (EBA). The analysts were given five sets of financial statements and were instructed to determine which company possessed the greatest future earning power. The information processes of the analysts were inferred from an analysis of verbal protocols and the answers the analysts gave to a post-experimental questionnaire. This review is limited to the segment of the research pertaining to protocol analysis, given the shortcomings of the post-experimental questionnaire technique (Nisbett and Wilson, 1977).

All four decision rules have been hypothesized in the psychology literature as possible explanations of behavior in complex tasks. Previous empirical verification has concentrated on testing for the fit of experimental data to mathematical representation of the models.

While this is an important type of empirical test of the models, it does not directly test the information processing assumptions associated with the models. That is, we cannot know from this type of evidence if subjects actually processed information as the model implied. (Biggs, p. 36)

Biggs' study contains a discussion of the mechanics of the four models.

Protocol analysis was conducted on two levels. On the more general level, each subject's behavior was described in
terms of a sequence of episodes. An episode was deemed to have begun when there was either an explicit statement of a goal by the subject, or the existence of a goal could be inferred by examination of the activities the subject performed. The episode ended when a new goal was set or the task was completed. An important aspect of an episode is that it allows for distinguishing information gathering segments from choice making segments of the protocol.

The more detailed levels of analysis concerned the identification of operators in the protocols. The most important operators to the research were evaluation operators. Independent evaluation operators were assigned when the subject made a "... judgement about one or more of a company's financial aspects without regard to how that company compares with another company" (Biggs, p. 79). Dependent evaluation operators were assigned when an evaluation was made across two or more companies.

A subject's overall model was inferred from the existence of a predominance of use of dependent versus independent evaluation operators and evidence of compensatory or non-compensatory processing. The four possible combinations yielded inferences as follows:

1. Dependencies/Compensatory --AD
2. Dependencies/Non-compensatory --EBA
3. No dependencies/Compensatory --AC
4. No dependencies/Non-Compensatory--C
Nine of the eleven subjects were identified as predominantly using one of the choice models. The remaining two were judged to use a hybrid form of two of the models. The most important finding was evidence of use of all four models.

Two papers by Biggs and Hock (1979a, 1979b) used protocol analysis to investigate the decision processes of auditors evaluating internal controls and making audit scope decisions. Concurrent review is used since both studies utilize the same subjects and experimental task. Five specific research questions were addressed:

1. Are there discriminable patterns of behavior which are descriptive of the subject's overall decision process and if so, were there similarities in these patterns across auditors?

2. What specific items of information did subjects search for and use in making their decisions and were there similarities of information usage across auditors?

3. What proportion of available information did subjects use in making their decisions?

4. How did subjects evaluate the information? That is, what criteria did they use in their evaluation of information?

5. How did subjects make their scope decisions? Of particular concern were questions such as what alternatives did subjects consider when making their decisions? And what line of reasoning provided the basis for their decisions (p. 4)?

The experiment utilized four auditors from the same large CPA firm. Two auditors had two years of experience and two had four years of experience.

The auditors were told to assume that they had just taken over an audit engagement because the previous audit senior had resigned... In essence
the task was to examine the audit material and decide whether the tentative scopes for four audit procedures established by the previous audit senior were appropriate in the circumstances. If not, they were to make specific scope recommendations. (Biggs and Mock, 1979b, pp. 8-9)

Within each experience group, one subject was given a case with "fair" improvements in internal control from one year to the next, and the other with "strong" improvements.

The analysis of the protocols was characteristic of the methods used by Biggs (1978). A micro-level of analysis consisted of coding the protocols on the basis of operators. Operators were divided into four major categories: task structuring (gaining understanding of the task and setting goals), information search, analysis ("... the subjects' processes as they evaluated the information in terms of the assumptions and judgments they were required to make." (Biggs and Mock, 1979b, p. 13)), and choice (generating alternatives and final determination of scope for the required audit procedures). The coding rules for the operators within the four categories are presented below since the protocol analysis of this dissertation proceeds along similar lines.

I. Task Structuring
   A. Set Goal (SG)—Assigned when the subject specifies a goal to be accomplished in performing the sample size decision. The SG operator usually signifies the beginning of an Episode or Subepisode.

II. Information Search
   A. Information Search (IS)—Assigned when the subject searches the case materials for specific pieces of information (directed search) or searches following
some systematic pattern (usually sequential search). A piece of information is defined as all the words contained under one label (section) in the case materials.

B. Algebraic Calculation (AC)—Assigned when subject makes a mathematical calculation in order to obtain new information about the task.

C. Information Retrieval (IR)—Assigned when subject retrieves a previously stored piece of information from external memory (i.e., notes, calculations) or internal memory.

III. Analytical
A. Assumption (AS)—Assigned when subject generates an arbitrary (unspecified) fact about the case.

B. Conjecture (CJ)—Assigned when subject makes an if-then or hypothetical statement.

C. Comparison (CN)—Assigned when subject makes a judgment based upon a comparative process (i.e., two alternatives, the current and prior year's program).

D. Evaluation (E)—Assigned when the subject makes a teleological judgment about the task based upon some explicit or implicit criterion.

E. Generate Alternative (GA)—Assigned when subject states, in a tentative form, an alternative sample size, audit procedure or other task related action.

F. Generate Query (GQ)—Assigned when subject raises a question about the task.

IV. Decision Process
A. Decision Rule (DR)—Assigned when subject specifies a method (including heuristics) of determining a sample size or parameters (i.e., stratification) directly related to sample size decision.

B. Sample Size (SS)—Assigned when subject finalizes sample decision (SS) or specifies a temporary sample size (TSS) which is ultimately revised.

C. Other Decisions (OD)—Assigned when subject recommends other actions to be taken (i.e., "must consult with manager," or recommends additional audit procedure).
Two types of overall solution strategies were employed by the auditor subjects. A directed strategy involved "... selection of a particular audit step and then a directed search for and evaluation of information relevant to that audit step alone. Once that this scope decision was made, a similar process was employed on the next audit step" (Biggs and Mock, 1979b, p. 18). A systematic strategy involved a consideration of all audit decisions simultaneously. The two most experienced subjects used the directed strategy and the least experienced subjects used the systematic strategy.

There was considerable variability observed in the amount of information used by subjects. Biggs and Mock suggest that the subjects utilized general classes of information in relatively similar proportions, even though the actual amount of information used in each category differed. The two least experienced subjects used a much greater proportion of the available information than the two most experienced subjects.

Bouwman (1980,1981) views the financial statement analysis task as a diagnostic process. That is, the analyst searches for symptoms and from these symptoms, he attempts to discover the actual state of the system. A large portion of the experimental task in this dissertation can be viewed as a problem in diagnosis similar to that of Bouwman's subjects in another task; the auditor uses information to discover the state of a system of internal controls.
The first step in diagnosis, according to Bouwman, is the examination of information. This serves two functions: to screen information and to translate information into convenient terms. The examination process begins with the selection of an item of information. Three types of strategies influencing the order of item selection, or the search for information, are identified:

1. A "checklist" approach may be employed where the information processor has a predetermined set of items to examine. The "checklist" may be partially conditional on certain findings.

2. "In-depth search" involves beginning with an important item, and branching out to its components.

3. "Sequential search" is simply the inspection of all items in the order of presentation.

The translation process involves the use of operators to translate the item into qualitative terms. The result is then evaluated for its potential significance. If the fact is judged to be potentially significant, it is retained in long-term memory.

The diagnosis may be formulated through either the recognition of a pattern of symptoms or through a process of reasoning. One would expect the former strategy to be employed by experienced persons only.

The analytical techniques used by Bouwman are quite similar to the Biggs (1978) and Biggs and Mock (1980) studies discussed previously in that the protocols are examined on an episodic level and on a task level. The task
level operators, or "activity codes", used by Bouwman are shown below. Like those of Biggs and Mock, they aid in the development of operators for the analysis of the experimental task in this dissertation.

**GENERAL CODES**

1. read **R**  
   read information

2. paraphrase **PAR**  
   paraphrase what has been read

3. question **Q**  
   formulates a question

4. comment **COM**  
   comment re task content

5. meta comment **MC**  
   comment re problem solving process

**GOAL CODES**

6. state goal **SG**  
   states a goal

7. get report **GR**  
   selects a specific report

8. get item **GI**  
   selects a specific item

9. future goal **PG**  
   states potential future goal

**EXAMINATION CODES**

10. computation **COMP**  
    performs mathematical calculation

11. compute trend **TREND**  
    computes a trend over time

12. compare **C**  
    compare two items

13. internal norm **CI**  
    compare with internal norm

**MEMORY CODES**

14. significant fact **SF**  
    stresses a specific observation

15. retrieve **RET**  
    retrieve information from memory

**REASONING CODES**

16. assumption **AS**  
    makes an assumption

17. summarize **SUM**  
    summarize evaluations

18. inference **INF**  
    infers causes/consequences

19. hypothesis **HYP**  
    hypothesizes a formal cause of problems

20. explanation **EXPL**  
    (fact 1) explains (fact 2)

21. confirmation **CONF**  
    confirmation
2.5.2 Empirical Studies in Information System Choice

The first published empirical test of the information economics model with consideration of the accountant and decision maker as separate persons was performed by Uecker (1979). In his experiment, forty junior and senior accounting students were instructed to imagine an urn-ball experiment similar to the one described in Chapter 1 (Section 2). The task of the subject was to select a sample with replacement of not greater than 50 balls from an urn selected at random such that a computer simulated, hypothetical decision maker could choose the urn from which the sample was taken. Cost and rewards were equal for both the subject and the decision maker. The cost of sampling was one cent per ball, and if the decision maker selected the correct urn type, a reward of fifty cents less the cost of sampling would be credited to both.

The student was given no information concerning the decision algorithm employed other than the decision maker's choices. In fact, there were two types of simulated decision makers: a Bayesian decision maker and a conservative decision maker (which applied an Edwards and Phillips exponential function to the likelihood ratio, see section 2.4.2). The particular power function used was selected from a range of .05 to .50 such that the difference between the optimal sample sizes for the two decision makers was maximized while maintaining an overall positive expected
value for the experiment. Given the posteriors in either form, the objective of the decision maker was that of expected payoff maximization.

Each subject performed eighty trials, forty with each decision maker. Half the subjects were paired with the Bayesian decision maker first, and the conservative decision maker second. The order was reversed for the remainder of the subjects. All subjects were informed that the decision maker had been changed after the fortieth trial.

The data obtained were analyzed first with respect to the ability of the subjects to converge on the optimal sample size (computed from equation (6)) and then for differences in performance due to the two different decision makers. The absolute deviation between a subject's chosen sample size and the optimal sample size for the appropriate decision maker was computed for each trial as a measure of subject performance. A Spearman rank correlation coefficient on this value and the ordinal position of the trial (1 to 40) served as the statistic by which convergence on the optimal sample size was measured. Subjects in the sequence with Bayesian decision maker first showed some convergence, but not enough to be significant at the .10 level.

A one-tailed t-test was applied to the mean correlation coefficient of each treatment group to test for significant convergence on the optimal sample size. This procedure was used because the mean correlation coefficient should be less than zero if subjects learned over time; the dispersion around the optimal sample size would decrease as subjects' knowledge of the decision maker increased.
level. The second sequence showed no convergence whatsoever; in fact the mean correlation coefficient was positive, indicating some possible divergence away from the optimal sample size.

The salient features of the experimental task in the dissertation are very similar (albeit more complex) to the features of the urn-ball experiment used by Uecker. The differences in the tasks presented to the subjects and the choice of the subjects themselves are explained by differences in objectives. Uecker was concerned with measuring the ability of subjects to learn the optimal information system over time. As a first attempt at this type of study, he was not concerned with the realism of the task. Uecker states that poor subject performance in such a relatively simple task casts doubt on the ability of subjects to perform well in more complex contexts of essentially the same task.

The effect on human performance of task environments and the process of information search is recognized. Although the goal of evaluating performance remains, a major emphasis is placed on identifying problem solving processes. Consequently, there is a need for a larger quantity of data (verbal protocols) and the use of experts as subjects.

In a second experiment, Uecker (1980) tested for (1) effects of knowledge of the information user's decision model and (2) the use of the anchoring and adjustment
heuristic on subjects' choice of sample sizes. The experimental task was very similar to that of his previous experiment on information system choice, except for the following:

1. The number of trials was increased to fifty.
2. Payoffs were changed.
3. The simulated DM chose urns based purely on the representativeness heuristic instead of the Bayesian model (or a variation of it).
4. The primary measure of subjects' performance was changed from the absolute difference of actual versus optimal sample size to the expected value of the subject's chosen sample size. The change was made because the new metric was thought to be more consistent with the subjects' instructions to maximize expected value and because subjects may not have incentives to find optimal sample sizes due to small differences in expected payoffs for sample sizes around the optimal. Also, due to the erratic nature of the expected value function over the range of sample sizes in the experiment, a sample which was very far from optimal could have an expected value closest to that of the optimal sample size.

Based on a 2 x 2 factorial design, subjects were placed in one of four treatment groups prior to performing the experiment. They were either ignorant of the DM's method of choosing urns from sample information or they were given a sufficient description of the representativeness heuristic to enable them to predict the DM's choices for any sample size and sample outcome. Subjects were initially anchored at a sample size of 4 or 40. The anchoring was accomplished by asking the subjects to guess whether the actual optimal sample size was above or below the anchor (4 or 40) prior to the start of the experimental task.
The results of tests concerning knowledge of the DM's decision model were largely negative. An ANOVA on subjects' estimated sample sizes prior to beginning the experiment yielded no significant main effects. Neither knowledge treatment group exhibited significant performance improvement over the fifty replications. It is also interesting to note that nearly all of the "no-knowledge" subjects were able to learn the DM's algorithm sufficiently well to predict the urn choices with at least a 70% hit rate. However, this result may not have occurred if the DM's algorithm was not so simplistic. The representativeness heuristic may be a very good approximation of how the subjects themselves might have acted as the DM, and their choices may have been highly influenced by their a priori beliefs about how a DM should act.

The analysis focused on actual sample sizes chosen as the dependent variable in testing for the anchoring effect. There was a significant anchoring treatment throughout the experiment. However, the evidence is weak that subjects were using anchoring and adjustment because it was only shown that the sample sizes chosen were significantly different for the two groups. Statistical tests would have yielded the same results if subjects randomly chose sample sizes centered around their anchor, which does not correspond to an anchoring and adjustment heuristic since adjustment should be a systematic process.
Both of Uecker's experiments surrogated expected dollar payoff computations for expected utility. Uecker maintains that utility elicitation techniques are not very accurate, and when used in other experiments they have yielded results not very different from expected payoff maximization. However, students' utility functions over small monetary values such as those used in this experiment may not lend themselves to accurate linear approximation.

The incorporation of expected utility in Uecker's study could make the computational requirements burdensome since utility functions would have to be elicited for each subject. Assuming that the simulated decision maker had an identical utility function, each subject would still have a unique optimal sample size which would have to be computed. Interpretation of results would also be more difficult since the marginal benefit for performance improvement would vary among subjects. For instance, there might be a very low marginal increase in utility for one subject to expend the effort to search for a better sample size. For another subject who chose, on average, the same sample size as the first subject, the incentive to search for a better solution might be greater. This might inhibit comparisons among subjects as well as increase the number of computations required in the experiment. A simulated experiment will "impose" a linear utility function on the decision maker and thus eliminate this problem until the laboratory experiment stage of the research.
Subjects had little incentive to find optimal sample sizes in the first experiment because of small differences in expected payoffs for sample sizes around the optimal. The expected value of the optimal size minus the expected value of the chosen sample size could be used as another measure of convergence for this reason. Absolute values are not necessary since the results of this computation are always non-negative. This is important since absolute deviation measures can yield the same convergence score for a number of differing selection strategies, some of which are more successful than others. One subject might consistently underestimate the optimal sample size by choosing constant sample sizes for each trial, and another subject may try sample sizes above and below the optimal. It is conceivable that their scores as measured in the experiment could be identical even though their strategies are quite different. Yet, the two strategies could yield large differences in actual payoffs, the more realistic long-term measure of success.

Finally, one should be hesitant to draw inferences concerning the behavior of accountants from an experiment which surrogates accountants with students and which surrogates an actual accounting problem with a problem in mathematics.
2.6 CONCLUSION

Previous research in the area of information system choice has been the primary motivation for this dissertation. The notion of studying behavior in the task domain of information system choice is relatively new, and protocol analysis, a powerful method of tracing cognitive processes has not been applied. Only decision inputs and outputs have been examined to date, and it is still not known how subjects will perform in a task where the problem statement more nearly recalls an actual accounting task. The literature review has established the general validity of verbal protocol analysis as a method of tracing subjects' problem solving behavior in complex tasks, and it has described how protocol analysis has already been applied to other important accounting research questions. The final three chapters describe how the methodologies of verbal protocol analysis, computer simulation and statistical inference were applied to an information system choice problem.
Chapter III

RESEARCH DESIGN

This chapter contains a summary of the experimental task and the methods used to analyze subject behavior. The three methods of analysis consist of (1) a statistical test for trend, (2) protocol analysis, and (3) computer simulation of the anchoring and adjustment heuristic. A section on each analysis will begin with a statement of the research question(s) to be addressed before the procedure itself is described. These questions are not considered only as formal hypotheses which can be accepted or rejected, primarily because questions of interest addressed do not lend themselves well to experimental designs involving control groups. The main purpose is to obtain data about each individual auditor's performance of the task with as little reference as possible to the larger population of auditors.

3.1 EXPERIMENTAL TASK

3.1.1 Description and Background

An experimental task which allows computation of a normative solution and is realistic enough to recall the type of task and accountant might perform was necessary in
order to conduct an empirical study of performance and process simultaneously. Previous research shows that the normative model is easily applied to sample size selection problems. Since auditors are faced with sample size selection problems in numerous phases of the audit, an experimental task in the domain of auditing was considered to be realistic.

The subjects' task involved the function of compliance testing of internal control procedures for purchases and payables. The task of the subject (information evaluator) is to choose the number of transactions sampled. The experimental problem states that each transaction sampled can be unambiguously classified as fully meeting internal control standards or as deficient in meeting standards. The following language was used:

You have just completed a review of the internal control procedures over purchases and payables, and the results of your inquiries and observations are summarized in the flowcharts and the memorandum contained in Exhibit E. Your manager has examined all of the above information and he has instructed you to choose a random sample of cash disbursements and to trace these transactions back through the system of internal control and to classify each transaction as either complying or not complying with the internal control requirements for that transaction.

The results of the sample (number of transactions sub-standard) will be made known to the subject and to his supervisor (decision maker). The supervisor will then select the level of substantive testing to be conducted.
based on his estimate of the true level of internal control in the company. The supervisor's decision is one of the following:

1. Internal controls are tight, and only minimum substantive testing is necessary.

2. Internal controls are satisfactory, except for some minor departures, and substantive testing will be at a higher level.

3. There are significant departures from internal control standards, and extensive substantive testing is necessary.

4. Internal control procedures cannot be relied upon, and the maximum amount of substantive testing will be employed.

The experimental problem was repeated a maximum of 50 times for each subject, and at the end of each trial, the subject was told (1) the true state of internal controls for that trial, (2) the action his supervisor took, and (3) the net payoff to the subject resulting from the sample size selection and the manager's action.

The reports to the subjects were generated through the use of an interactive computer program operated by the experiment monitor. The program, described in detail in Appendix C, simulated states of internal control, sample

---

3 One research design alternative available is to allow time to elapse between each experimental trial. The purpose of this strategy is to ensure sufficient time is available for the subject to learn from the last trial, and to control for any fatigue which might exist. Although this is a desirable safeguard, this alternative design was not used because of the desire to utilize the time donated by the subjects (practicing CPAs) as efficiently as possible. The strength of any conclusions or inferences made on the basis of the experimental results should be considered in light of this possible limitation.
outcomes, decisions, and recorded data for each trial on magnetic disk for later reference.

Audit sample size decisions are often made with the aid of statistical sample tables. To use the tables, the auditor must specify a confidence ("alpha") level, an estimate of the population error rate, and a reliability level (upper confidence bound for a one tailed test). A sample size program modeled after the sample size tables an auditor would normally use was built into the interactive program to accommodate subjects who felt it appropriate to consult a sample size table at any time during the experiment. If the subject requested sample size tables, he was supplied with instructions for using the sample size program.

Each subject was exposed to the same sequence of internal control states to guard against differences in subjects' behavior due to different random sequences of states. Sample outcomes could not be controlled as tightly because subjects were able to choose their own sequence of sample sizes. However, the sequence of random numbers associated with each trial was the same for each subject. To illustrate the effect of this technique on sample outcomes, assume that on trial "i", subject A chose a sample size of 50 transactions and observed a sample outcome of 3 defective transactions. If subject B chose a sample size of 70 on the same trial during his experiment, his sample
outcome would consist of two components: the 3 defective transactions from the first 50 transactions sampled plus the defective transactions observed from the incremental 20 transactions sampled.

Two types of decision makers were simulated. The first four subjects were paired with a decision maker who tested a series of hypotheses to infer the state of internal controls. The tests used were the binomial test and the related normal approximation to the binomial when the sample size was greater than 30 transactions (Hollander and Wolfe, 1973). Motivation for simulating the DM in this manner is derived from auditing texts which describe how auditors can apply statistical techniques to sample outcomes (Arens and Loebbecke, 1980). The last three subjects were paired with a bayesian decision maker.

Appendix B contains the case given the subjects together with the supplementary instructions for using the sample size selection program. The information contained in the case may be classified into the following basic components:

1. General background of company
2. Financial statements
3. Internal control flowchart and memo
4. Instructions on how to perform the task, and to think aloud during all phases of the task.
5. Information not commonly made explicit in an auditing situation, but necessary for computation of an optimal sample size.
The flowchart of the internal control procedures and the financial statements and accompanying notes were taken from the working papers of a small client of a Big Eight CPA firm in Columbus, Ohio. All other information regarding the client was withheld (except for the volume of cash disbursements), and therefore, fictionalized for the purpose of constructing the case.

It is the practice of this CPA firm to make an overall qualitative judgment of the effectiveness of the internal control system after reviewing the system and observing the results of the tests of compliance. The four judgment levels in the case reflect this practice. Population error rate intervals were chosen to correspond to these levels. The reasonableness of these intervals was supported by consultation with a former partner of a Big Eight firm who previously served as chairman of AUDSEC.

Information contained in the case which is not commonly made explicit in an audit setting is the prior probability distribution of error rates and payoffs. Two general types of prior probability distributions were given consideration: triangular and lognormal. Both were candidates because of their asymmetry.

The triangular distribution was considered desirable primarily because of its simplicity. Most importantly, a specific distribution from the general class of triangular distributions is determined by three easily understood
parameters: the mode (the apex of the triangle, and the most likely value to occur), and the two end points of the triangle, which determine the two values above and below which the probability of occurrence is zero. Other desirable properties are that the distribution has no infinite tails, and it is possible for a digital computer to integrate a given area under the triangle exactly, and in a negligible amount of time.

Although the triangular distribution was used during much of the early computer program development, it was later discarded in favor of the lognormal distribution because no way could be found to manipulate the parameters of the distribution to produce a distribution reflecting properties of error rate distributions in an internal control setting. In particular, the left tail of the distribution should be very fat and the right tail should be very thin, producing a distribution with high probabilities of low error rates, and low probabilities of high error rates.

The lognormal distribution is determined by two parameters, a mean and variance. Unlike the triangular distribution, it has no definite integral, but the right tail is easily manipulated. However, since the right tail is infinite (the left tail is finite), a truncated form of the lognormal must be employed since there is a stated upper bound of error rate occurrence in the case.
There is no evidence to show that auditors are familiar with the properties of either the triangular or lognormal distribution. Therefore, the subjects were not informed of the exact prior probability distribution. Instead, they were supplied with 100 random numbers generated from the underlying prior probability distribution. The procedure for generating these numbers is described in Appendix C.

Payoffs were chosen with two objectives: (1) to afford the subject with a wide range of sample size alternatives and (2) to impose a large penalty for gross errors in inference made by the decision maker. The first objective was necessary to study whether subjects would recognize the rapidly declining marginal utility (or diagnosticity) of large sample sizes. The second objective was an attempt to simulate the risk of large losses from litigation to which auditor's expose themselves to if they "underaudit".

3.1.2 Subjects

Nine subjects were utilized in the experiment: two in a pilot study and seven in the main experiment. The purpose of the pilot study was to determine if there were aspects of the experimental task which would cause the subject to have difficulty in completing the task. The two subjects used for the pilot study were faculty members at The Ohio State University. One was a professor with considerable experience in teaching auditing, and the other was a
lecturer who had no experience in teaching auditing, but had recent experience as an auditor. Both subjects were CPAs.

Only one major change was made in the experimental task as a result of the pilot study. Initially, the plan was that all subjects would repeat the experiment fifty times. Based on the pilot study, the experiment could be terminated whenever the subject appeared fatigued or thought it pointless to undergo further analysis of the data. One of the pilot study subjects took nearly one and a half hours to complete only twelve repetitions. The cognitive strain on the subject was evident; he indicated that he had analyzed the data as much as he could and was satisfied that he had refined his sample size selection as much as possible.

The seven subjects utilized in the main experiment were senior auditors drawn from four of the seven Big Eight CPA firms with a Columbus, Ohio office. The firm which contributed the case materials was not solicited for help because of the possibility that a subject might recognize the client. Of the remaining six, four volunteered a total of ten subjects. Only seven auditors were used in the experiment due to scheduling differences. Demographic data on the subjects is presented in Table 1, below. Subjects A and B were the pilot study subjects and are not included in any subsequent analyses.
### TABLE 1

**SUBJECT DEMOGRAPHIC DATA**

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>28</td>
<td>25</td>
<td>26</td>
<td>24</td>
<td>23</td>
<td>29</td>
<td>26</td>
</tr>
<tr>
<td>Years in Public Accounting</td>
<td>4.5</td>
<td>4</td>
<td>5</td>
<td>2.5</td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Job Title</td>
<td>Senior</td>
<td>Senior</td>
<td>Super.</td>
<td>Senior</td>
<td>Senior</td>
<td>Senior</td>
<td>Senior</td>
</tr>
<tr>
<td>Years at Present Position</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2 mos.</td>
<td>1 mo.</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Highest Degree Earned</td>
<td>B.S.</td>
<td>B.S.</td>
<td>B.S.</td>
<td>B.S.</td>
<td>B.S.</td>
<td>B.S.</td>
<td>B.S.</td>
</tr>
</tbody>
</table>

### 3.2 STATISTICAL TEST FOR TREND

The question of convergence on optimal criteria is significant in this research because (1) it is an extension of previous behavioral research in the area of information system choice and (2) it is an objective measure of subject performance. This section describes the non-parametric test for trend to be used in testing for convergence. The application of this statistical method is a departure from previous research. It is believed to be a methodological improvement since the time series of sample size selections do not satisfy the independence assumption necessary for parametric test procedures. This assumption is not necessary for the test described below.
Before proceeding to a description of the test used to determine convergence on the normative solution, it should be emphasized that the validity of the normative solution to the experimental task is conditional on the assumption that subjects attempt to optimize the criteria they are instructed to optimize (Einhorn and Hogarth, 1981). The instructions in the experimental task state that they are to attempt to maximize the expected payoffs to their firm. An alternative design to the experiment would entail the estimation of subjects' utility functions on the payoffs. This was not done, and expected payoff maximization was assumed for the following reasons:

1. Estimation of utility functions would require that each subject be asked numerous additional questions which would lengthen the experiment considerably and exacerbate any fatigue factor which could exist. In addition (a) there is no assurance that the utility function estimated from this procedure would be a reasonably accurate measurement of the subject's risk preferences, and (b) the additional questions might induce behavior which would not have otherwise been manifested.

2. A unique utility function for each subject forces the calculation of optimal solutions to the task for each one. Since the calculation of optimal solutions already consumed significant time and resources, this was not viewed as desirable.

3. Maximization of expected payoffs of the firm is arguably criteria would maximize the wealth of the owners.

4. Since fictional payoffs were used, there is little reason, a priori, to expect that subjects would not subsume the goal of expected payoff maximization when instructed to do so.
3.2.1 **Statement of Research Question**

Q1: Will the subjects exhibit statistically significant convergence on the optimal sample size or the optimal expected payoff over time?

3.2.2 **Procedure**

For each subject, and for each trial, the following two metrics were calculated:

\[
Y(i,j) = |N(i,j) - N^*| \\
Z(i,j) = E(\$|N^*) - (E(\$|N(i,j))
\]

where

- \( N^* \) is the optimal sample size
- \( N(i,j) \) is the sample size decision on the \( i \)th trial by the \( j \)th subject
- \( \$ \) is the monetary payoff
- \( E \) is the expectation operator

A non-parametric test for trend (Lehman, 1975, p. 129) was employed to test for evidence of a downward trend over time of both metrics toward zero for each subject. (Note that the value of either metric is always greater than or equal to zero.)

To perform each of the 14 tests, the statistic \( D^* \) was computed:

\[
D^* = \sum_{j=1}^{J} ((T^*(j) - j) ** 2)
\]

where

- \( J \) is the number of trials
**T*(j)** is the midrank of the metric \(Y(*, j)\) or \(Z(*, j)\)

The statistic

\[ (14) \quad (D* - E^0(D*)) / (VAR^0(D*))^{0.5} \]

has a limiting standard normal distribution. To calculate \(E^0(D*)\) and \(VAR^0(D*)\) we use

\[ (15) \quad E^0(D*) = 1/6*(N**3-N) - 1/12*SUM(d(k)**3-d(k)) \]
\[ c=1 \]

and

\[ (16) \quad VAR^0(D*) = 1/36*(N-1)*((N**2)*((N+1)**2)*\]
\[ C \]
\[ (1-\{SUM(d(k)**3-d(k))\}/(N**3-N)) \]
\[ c=1 \]

where

\[ E^0 \text{ and } VAR^0 \] are the expectation and variance operators under the null hypothesis

\[ C \text{ is the number of tied groups} \]

\[ d(k) \text{ is the size of the kth tied group} \]

### 3.3 Protocol Analysis

Concurrent verbal protocols are collected to provide a data base which will prove useful in answering the questions of subjects'  
(1) application of concepts from the normative model  
(2) information utilization and  
(3) use of the anchoring and adjustment heuristic.
3.3.1 **Statement of Research Questions**

The research questions in group 2 are related to information processes which are necessary to solve the problem using equation (7).

Q2a: Does the subject attempt to discover the decision maker's algorithm?

Q2b: Does the subject formally calculate expected payoffs?

Q2c: Does the subject calculate prior and posterior probabilities?

A large amount of information in the case is information which auditors typically utilize on an engagement which are not direct inputs to the normative model. This information consists of the flowchart of the internal control system, the unaudited financial statements, and miscellaneous information provided in the case. The third group of research questions is concerned with the subject's use of this information.

Q3a: What is the role of financial statements in the subjects' decision processes?

Q3b: What is the role of the flowcharts in the subjects' decision processes?
The research question which constitutes group 4 is concerned with the subject's pattern of sample size choices over time.

Q4: Are subject's sample size choices over time related to the anchoring and adjustment algorithm?

3.3.2 Procedure

Concurrent verbal protocols were collected from each subject. Each protocol was transcribed into a series of numbered lines; each line is a short phrase. These phrases were analyzed, or "scored", by assigning an operator to each line or a group of successive lines. Each operator may be thought of as a subprogram invoked by the subject to attempt to achieve a sub-goal of the problem to reach the ultimate goal of selecting a sample size. Information concerning subjects' sub-goals was obtained from the protocols by specifying program arguments in many cases.

The operators and arguments utilized in scoring the protocols are described below. The motivation for choosing this particular set of operators comes primarily from two sources: (1) extant accounting research described in the previous chapter which employed protocol analysis as a research methodology and (2) to address the research questions stated above. The arguments were chosen on an ex post basis. The protocols were examined for evidence of
various classes of information used by, or produced by, various programs. Those items which occurred frequently in such a context were used as arguments.

I. INPUT CODES
A. Read (R)—Information is read verbatim. Arguments consist of the classes of information which are readable: the case (C), a financial statement note (FSN), and sample size program instruction (SSP). Although other verbal messages are in the flowchart and accompanying memorandum on internal controls, these documents are considered to be primarily symbolic in nature.

B. Re-read (RR)—Information which has been previously read or paraphrased is re-read verbatim. Arguments are identical to IA, above.

C. Paraphrase (PARA)—Information is paraphrased. Arguments include the flowchart (FC) and the sample report (SR) as well as those in IA.

D. Re-paraphrase—Information is paraphrased which has been previously read or paraphrased. Arguments are identical to IC.

II. EXAMINATION CODES
A. Question (QUE)—The subject addresses a question to the experiment monitor. The task content (IC) operator is scored when the question pertains to interpretation of any information provided the subject in the case, such as the task requirements or characteristics of the company. The flowchart (FC) argument is used for any question pertaining to interpretation of the flowchart.

B. Compare (C)—Two or more items are evaluated in relation to each other. Arguments refer to the class of items compared: financial statement items (FSI), sample sizes (SS), and sample error rates (SAMPRE).!

C. Compare With Internal Norm (CI)—evaluate an item in relation to some internal criterion. As in IIB, arguments refer to the item being evaluated. Additional arguments include an internal control procedure found in the flowchart (FC), a decision made by the audit manager (DM), and the cost of sampling (SAMPJOST).
D. Comment (COM)—The subject makes a statement concerning some aspect of the experimental task. Arguments include DM, TC, SS, FSI, payoffs (PAY), expected payoffs (EXPPAY), information processing costs of cognitive effort (IPC), SAMPCOST, SAMPRATE, prior probabilities (PRIProb), and input parameters for the sample size selection program (FARM).

E. Meta Comment (MC)—A statement is made re the subject's problem solving process. Included in this classification are statements of the subject's intent to achieve a goal or a sub-goal. Possible arguments are the same as III.

F. Calculation (CALC)—A precise mathematical calculation or an estimate of a quantitative value from quantitative inputs. Arguments specified are the results of the mathematical operation or estimate. These include SAMPRATE, the range of sample error rates given in the case (RANGE), and PRIProb.

G. Explanation (EXP)—Fact 1 explains fact 2. Arguments refer to the subject of the explanation, e.g., DM, FSI.

H. Inference (INF)—The subject infers causes or consequences. Arguments refer to the subject of the inference.

I. Assumption (AS)—Arguments refer to the subject of the assumption.

III. MISCELLANEOUS CODES

A. Computation Interval (COMP)—Pauses in the tape recording of the protocol while the subject is waiting for computer output are transcribed as a blank, numbered line. No arguments are used.

B. Pick (PI)—The subject chooses a sample size or a parameter for input in the sample size selection program. The arguments FARM and SS are used as indicators.

C. Uncodable Phrase (UF)

D. Monitor Comment (MC)—The experiment monitor interjects a comment or an answer to a question. No arguments are utilized.
3.4 **COMPUTER SIMULATION**

The purpose of computer simulation in this research is to investigate the efficacy of a pure anchoring and adjustment heuristic search strategy over several types of decision makers.

3.4.1 **Statement of Research Questions**

Q5a: Since the anchoring and adjustment heuristic does not require any knowledge of the DM's decision algorithm, can the optimal sample size be approached over time without any knowledge of the decision making algorithm?

Q5b: Is the performance of the anchoring and adjustment heuristic for the information evaluator affected by different decision making styles of the DM?

3.4.2 **Procedure**

The simplified example of the task Uecker (1978, 1980) used in his experiments and described in Chapter I will help illustrate the fundamental problems of simulating the DM and IE. The problem for the subject (IE) is to choose the sample size. Since any sample size over three has no possibility of a positive return, we can limit consideration to sample sizes 0 through 3. The decision maker's choices can be completely represented by the matrix shown in Figure 5:
In each cell, \( P(*) \) represents the probability that the decision maker will choose an urn type, where that probability is conditional on \( y \), the number of black balls in the sample outcome, and \( N \), the sample size. In calculating the above chart for each type of decision maker, it will be assumed that each decision maker is consistent in the sense that he will always make the same inference given the same values for \( y \) and \( N \). Thus, one \( P(*) \) is equal to one in each cell, and the others are zero.
The anchoring and adjustment heuristic can be simulated using the following algorithm:
Figure 6
Flowchart of Anchoring and Adjustment Heuristic
The same non-parametric test for trend was employed to evaluate the performance of simulated subjects as was used to test for human subjects' ability to converge on the optimal sample size, \( N^* \), and the optimal expected value, \( E(S|N^*) \) (see section 3.2.2). Each simulation run consisted of 100 subjects paired with one of the two types of information evaluators used in the experimental task. For each run, one test for trend was performed on the average of simulated subjects' \( Y(i,j) \) responses (see equation 11). The metric \( Z(i,j) \) was not utilized in this phase of the research because the anchoring and adjustment algorithm does not consider the level of payoffs, and the tests for trend utilizing \( Z(i,j) \) would be redundant.

Eighteen (18) simulation runs were made. These consisted of all combinations of the two types of decision makers combined with three levels of the adjustment parameters and three levels for the anchor (3 x 3 x 2). Choices for the three levels of the anchor and adjustment parameters were based on results presented in chapter 4 and are therefore better discussed after these results are presented. The algorithms used to simulate each type of information evaluator are described in Appendix C.
Chapter IV

EXPERIMENTAL RESULTS

The results of the laboratory study and the computer simulations are presented in this chapter to provide answers to the research questions posed in Chapter III. Three types of results are presented: statistical tests on (1) "live" and (2) "simulated" subjects and (3) the analysis of the verbal protocols. The three types of data sources are (1) the sample size choices made by each subject, (2) the concurrent verbal protocols collected and (3) the sample size choices made by simulated subjects.

4.1 TEST FOR TREND

Q1: Will the subjects exhibit statistically significant convergence on the optimal sample size or the optimal expected payoff over time?

Recall from the previous chapter that Subjects C through F interacted with a "statistical" decision maker, and subjects G through I interacted with the Bayesian decision maker. The optimal sample sizes for the two DM's are 9 and 75 respectively, and the expected value of the optimal sample sizes are $24.34 and $53.41.
Table 2 lists the results of the test for downward trend on
the two metrics. Since the test is one-sided, the null
hypothesis is rejected only for large values of $z$, the
standardized value of $D^*$ (see equations (13-16)).

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>N</th>
<th>$Y(i,j)$</th>
<th>Z(i,j)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>z</td>
<td>SIGNIF.</td>
</tr>
<tr>
<td>C</td>
<td>28</td>
<td>-4.65</td>
<td>N.S.</td>
</tr>
<tr>
<td>D</td>
<td>50</td>
<td>1.79</td>
<td>.04</td>
</tr>
<tr>
<td>E</td>
<td>23</td>
<td>0.75</td>
<td>N.S.</td>
</tr>
<tr>
<td>F</td>
<td>50</td>
<td>2.28</td>
<td>.01</td>
</tr>
<tr>
<td>G</td>
<td>50</td>
<td>-5.04</td>
<td>N.S.</td>
</tr>
<tr>
<td>H</td>
<td>50</td>
<td>-1.97</td>
<td>N.S.</td>
</tr>
<tr>
<td>I</td>
<td>17</td>
<td>-2.04</td>
<td>N.S.</td>
</tr>
</tbody>
</table>

The results of the tests are clear, but disturbing. Only three of the fourteen tests indicated a downward trend at an acceptable level (.05) of significance. This finding is not surprising given the results of Uecker's work, but it was not anticipated that so many subjects would move in a systematic fashion away from optimal behavior. In fact, if the same test were conducted for upward trend (albeit ex post), nine of the fourteen tests would have resulted in rejection of the null hypothesis at a significance level of
less than or equal to .025. The evidence compels one to suspect that more experienced auditors may perform poorly as information evaluators.

Subject D is an example of the possible errors referred to in Chapter II that can be made if, for non-bayesian decision makers, performance is measured as a function of the sample size chosen. One metric, $Y(i,j)$, for this subject indicates convergence, and the other, $Z(i,j)$, indicates divergence.

It was mentioned in Section 3.2 that certain assumptions had to be made concerning auditors' criteria in decision making. The viewpoint was taken that auditors would subsume the goal of expected firm payoff maximization, and this was the basis for the calculation of the normative solution. No subject clearly indicated in his protocol that he took abnormal care to avoid the possibility of a large loss (consistent with the concept of risk aversion). The lack of evidence for risk aversion is supported by the results of the next section, which show little attention to any of the variables in the normative model by the subjects.

4.2 PROCESSING CONSISTENT WITH THE NORMATIVE MODEL

The normative solution to the experimental task is shown as Equation (6) in Chapter I (see section 1.2). It is virtually impossible to solve the problem by hand since the program written to solve the general equation consumes
approximately 35 minutes of CPU time on the university's Amdahl 470 computer to solve just one of the problems given in the experimental task. It is clear that the costs of time and effort for a human subject in such an endeavor would far exceed any benefit of obtaining the precise solution to the problem. Nevertheless, one purpose of the study was to see how, and to what extent, subjects will incorporate the concepts of probability and expected payoff maximization in their decision making processes.

Q2a: Does the subject attempt to discover the decision maker's algorithms?

Table 3 is a tabulation of the frequency that each operator and argument was scored for each subject.
### Table 3
**Operator Totals by Subject**

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Read</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Fin. Stat. Notes</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Re-read</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Sample Size Program</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td><strong>Paraphrase</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Financial Statements</td>
<td>3</td>
<td>17</td>
<td>8</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Fin. Stat. Notes</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Flowchart</td>
<td>3</td>
<td>2</td>
<td>23</td>
<td>8</td>
<td>6</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>Sample Report</td>
<td>29</td>
<td>36</td>
<td>10</td>
<td>27</td>
<td>2</td>
<td>33</td>
<td>13</td>
</tr>
<tr>
<td>Sample Size Program</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Re-paraphrase</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td>2</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>Financial Statements</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Flowchart</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Sample Size Report</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample Size Program</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td><strong>Question</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task Content</td>
<td>7</td>
<td>13</td>
<td>7</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Flowchart</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>Other</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td><strong>Compare</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial Statement Items</td>
<td>8</td>
<td>20</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Sample Sizes</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Sample Error Rate</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Compare with Internal Norm</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial Statement Items</td>
<td>1</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>Sample Size</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Sample Error Rate</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Flowchart</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>7</td>
<td>3</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Decision by Manager</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sampling Cost</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Comment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decision Maker</td>
<td>55</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Task Content</td>
<td>4</td>
<td>19</td>
<td>12</td>
<td>2</td>
<td>8</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Sample Size</td>
<td>2</td>
<td>7</td>
<td>3</td>
<td>12</td>
<td>0</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Financial Statement Items</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Table 3, continued</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Payoffs</strong></td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Expected Payoffs</strong></td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Costs of Effort</strong></td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Sampling Cost</strong></td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Sampling Rate</strong></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Prior Probabilities</strong></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Flowchart</strong></td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Sample Prog. Parameters</strong></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td>3</td>
<td>15</td>
<td>12</td>
<td>9</td>
<td>11</td>
<td>11</td>
<td>17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Meta Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample Size</strong></td>
</tr>
<tr>
<td><strong>Financial Statement Item</strong></td>
</tr>
<tr>
<td><strong>Payoffs</strong></td>
</tr>
<tr>
<td><strong>Prior Probabilities</strong></td>
</tr>
<tr>
<td><strong>Flowchart</strong></td>
</tr>
<tr>
<td><strong>Sample Prog. Parameters</strong></td>
</tr>
<tr>
<td><strong>Other</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample Error Rate</strong></td>
</tr>
<tr>
<td><strong>Range of Prior Probs.</strong></td>
</tr>
<tr>
<td><strong>Prior Probability</strong></td>
</tr>
<tr>
<td><strong>Other</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Financial Statement Item</strong></td>
</tr>
<tr>
<td><strong>Sample Size</strong></td>
</tr>
<tr>
<td><strong>Task Content</strong></td>
</tr>
<tr>
<td><strong>Other</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample Size</strong></td>
</tr>
<tr>
<td><strong>Decision Maker</strong></td>
</tr>
<tr>
<td><strong>Optimal Sample Size</strong></td>
</tr>
<tr>
<td><strong>Financial Statement Item</strong></td>
</tr>
<tr>
<td><strong>Other</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Assumption</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Decision Maker</strong></td>
</tr>
<tr>
<td><strong>Flowchart</strong></td>
</tr>
<tr>
<td><strong>Other</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Computation Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pick</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample Size</strong></td>
</tr>
<tr>
<td><strong>Sample Prog. Parameters</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Uncodable Phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Monitor Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
</tr>
</tbody>
</table>
Question 2a is addressed by searching each protocol for any operator which contains the "DM" argument. Table 3 shows that only 22 such operators were scored, and that subjects G and H made no reference to the decision maker. Subject I made only one comment (line I762), but no information concerning Question 2a was revealed. Recall that Subjects G, H and I constitute the group exposed to the Bayesian DM.

Subjects C, E and F observed that the decision maker was not using the representativeness heuristic (lines C147-C151, E566-E571, F534-540), but further on, Subject C also erroneously concludes that the DM's algorithm is not statistically based (lines C369-C390). All other remarks of Subject C concerning the decision maker are expressions of his dissatisfaction that the DM is unable to exactly predict the true state of internal control with greater frequency. Although these remarks have no direct relevance to the discussion of Question 2a, it should be noted that the subject never referred to sampling error as a possible cause of the decision maker's inaccuracy. Subjects D and E's only other material comments were an observation by each that the decision maker had inferred only once that the error rate was less than it actually was (lines D864-D868, E824-E826). Subject F had nothing further to remark about the DM.
Subject D assumed that the DM used some type of statistically based method by the seventh trial (lines D635-643), and as such, he is the only subject to have discovered the DM's algorithm for making inferences about the state of internal controls. There is no evidence, however, that he used this knowledge to make sample size choices.

Q2b: Does the subject formally calculate expected payoffs?

Q2c: Does the subject calculate prior probabilities?

Subjects made surprisingly few numerical calculations while performing the experimental task, as can be seen by inspection of the calculation operator in Table 3. No calculations involved the combination of probabilities with payoffs, and the answer to Question 2b must be an unequivocal "no" on the basis of this evidence. It is not true, however, that all subjects are unaware of the concept of each sample size having an expected value which can be used to compare one with the other. Subject C even suggested a method of combining the payoffs in the problem with statistical levels of significance to calculate an expected value for each sample size (lines C738-C763). He
did not apply his ad hoc suggestion to the problem at hand, he stated, because he did not have the "time and resources" to do so. There is, nonetheless, an indication that some auditors consider long-run average performance an important criteria in conducting an audit.

No subject explicitly stated an estimated prior probability of the occurrence of a state of internal controls. Subjects E, F and H estimated the range of prior probabilities by scanning the sample error rates, but it was not apparent from the protocols how this information was used in selecting a sample size. The only other calculation of prior probabilities involved the estimation of the mean error occurrence rate. In all cases, this was done to determine one of the input parameters for the sample size selection program which was used by subjects. Only one subject, Subject I, made the estimate of the mean error occurrence rate by referring to the sample error rates given in the case.

4.3 OTHER INFORMATION SOURCES

Q3a: What is the role of financial statements in the subjects' selections of sample sizes?

The subjects referred to the financial statements and accompanying notes for at least one of the following reasons: (1) to become acquainted with the general financial
characteristics of the company such as size, financial leverage, changes in financial characteristics over time, and (2) to determine the importance of the audit engagement as a whole, or the importance of the procedure they are asked to perform in the experimental task. The extent to which the financial statements were attended to varied greatly among subjects, but some commonalities did exist. It was significant that no subject computed a single financial ratio. This finding was somewhat unexpected since ratio analysis is probably the most heavily emphasized technique of financial statement analysis in college business courses. An accounting major would be exposed to it in his accounting principles courses, financial management course, and auditing course. Another commonality was that the financial statements were not referenced or referred to again after the selection of the first sample size.

Not all subjects explicitly stated the goals which they intended to achieve by analyzing the financial statements. However, we may gain some understanding of unstated goals by analyzing some other operators applied to the financial statements. The Compare operator indicates a desire to examine changes in financial data over time, the CI operator would indicate that the subject is comparing a financial statement item with some aspect of an internally stored, prototypical set of financial statements to assess the
degree to which the company is unusual. The Read and Paraphrase operators might be construed to represent the goal of becoming acquainted with the client, but they may only represent the process of information search.

The degree to which subjects inspected the financial statements delimits the subjects into two well defined groups. Subjects C, F and G engaged in a relatively limited analysis, as indicated by the number of lines of protocol, and the number of operators scored for them involving the financial statements. Subjects D, E, H and I made a more systematic and thorough investigation of this aspect.

Subject C made eight comparisons and concluded from the statements that the current year was "not a very good year" (lines C9-C42). Subject F's reasons for examining the financial statements were to become acquainted with the general nature of the company, and to find any significant fluctuations that would impact on the specific requirements of the case. He concluded: "nothing there that's going to too much impact on my compliance test, other than my dollar amounts I'm talking about" (lines F105-F140). Subject G (lines G15-G21) was only concerned with determining the size of the company.

Of the four subjects who spent large amounts of effort on financial statement analysis, three (C, F and H) decided to perform all or a large portion of the analysis before even reading what the task requirements were. These
The subjects evidently felt that a certain level of knowledge about a client must be obtained before any aspect of an audit engagement could be planned. The only other difference in their processing of the statements was their concern for specifying the significant types of other audit areas that should be emphasized on the engagement besides the one most pertinent to the task requirements. Subject I's protocol is different from the other three in the second group in the respect that he also thought it important to identify the various outside groups that will utilize the audited statements. He expressed concern for SEC reporting requirements (lines 156-157), estimates the number of shareholders (179-180), and mentions that the needs of the client's bank should be considered (181).

Table 3 shows that the Compare and Compare with Internal Norm operators were the most frequently used with a financial statement item argument. The compare operator was used to compare two statement line amounts across years to determine if there was a significant change from one year to the next. The internal norm comparison indicates that subjects were making an evaluation of an aspect of the company's financial condition based on their conceptions of what a prototypical firm would disclose in its financial statements.

Thus, the main differences between subjects were (1) the amount of effort expended in financial statement
analysis and (2) the individual goals of that analysis. These differences indicate the two basic points of view with which subjects considered the case problem. Subjects who read the task requirements first were able to adapt their analysis to the specifics of the case, and their analysis was presumably an abridged version of what they might do when first approaching an engagement where a broader range of audit steps must be accomplished. Three of the four longest analyses were made before the subject even knew the task requirements, and hence, they did not afford themselves the opportunity to adapt their "normal" processes to the specifics of the laboratory experiment.

Q3b: What is the role of the flowcharts in the subjects' selection of sample sizes?

Subjects displayed considerable commonalities in interpreting the flowchart. The basic procedure was to start at the beginning of the flowchart and to follow it to its termination. This process was interrupted to ask a question of clarification, or to state that a weakness or strength in the system had been identified. Some subjects made summary comments of the strength and weaknesses found after the analysis was completed, and then made an overall evaluation of the system.

Similar to the analysis of the financial statements, the most pertinent aspect of analysis of the flowcharts to
this study is the subjects' specifications of the goals they hoped to achieve from the analysis. The subjects can be classified into three fairly well-defined groups in this respect. Subjects C and H, the first group, gave no reason for examining the flowchart, but their behavior indicates that they were merely concerned with becoming familiar with the internal control procedures of the client. Subjects E, F, G and I, the second group, stated the specific goal of identifying strengths in the internal control system that can be tested (lines E180, F200-F210, G29-G43, I127-I131). However, subjects E and I were influenced by the internal control weaknesses they discovered in selecting their earlier sample sizes even though there was no explicit goals statement concerning weaknesses. Subject D, the third group, was the only subject to state the goal of examining the flowcharts for weaknesses in the internal control system.

The role of compliance testing in Generally Accepted Auditing Standards is to verify that stated internal controls are actually functioning. The auditor should use the flowcharts to establish the aspects of internal control that are purported by management to exist (Arens and Loebbecke). In this sense, the prescribed use of the flowcharts is to become familiar enough with the system of internal controls to identify the strengths to be tested. The role of the compliance test is to determine the degree
to which the controls are functioning. However, Subjects E and I, as well as D, were influenced in their sample size selection by the internal control weaknesses they found during their review of the system (lines E518, I542-I548, D475-D483). Thus, the auditor's apparent concern for weaknesses in the internal control system at the stage of the audit described in the case is puzzling.

Prior to this study, one explanation which could have accounted for subjects' attention to internal control weaknesses would have been that internal control weaknesses were inputs the subjects used to develop a subjective prior probability distribution of error rates. The plausibility of this explanation is now doubtful since the subjects paid scant attention to prior probabilities, and no sample size decisions were rationalized on any basis remotely reminiscent of prior probabilities. Another explanation is provided by the analysis of the protocols. Subjects F and G (lines F457, G29-G43) state that one of the decisions they must make before choosing a sample size for compliance testing is the degree of reliance they wish to place on the system. As the problem is stated, all aspects of this decision are made after observing the sample outcome. Hence, identification of weaknesses in a system of internal controls may help the auditor develop his prior beliefs as to how much reliance he will place on the system of internal controls.
This explanation has more appeal, but the protocols admittedly constitute weak support as the phrases cited were from subjects other than those who explicitly made use of internal control weaknesses in their decision processes.

Subjects E and I used the discovery of weaknesses in the internal control system to place a lower bound on the number of transactions they will select (E518, F542-F548). No explanation can be offered for this heuristic, except to note that at least some auditors used it in the context of this experiment to place a lower bound on the number of transactions they might choose to sample.

4.4 ANCHORING AND ADJUSTMENT

Q4: Are subjects' sample size choices over time related to the anchoring and adjustment algorithm?

The binomial test (Hollander and Wolfe, 1973), and analysis of the protocols were employed seeking evidence for the use of the anchoring and adjustment heuristic by the subjects. The flowchart of the anchoring and adjustment algorithm in Figure 6, shows that it is possible to count the number of times each subject changed his or her sample size consistent with the algorithm. This count can be compared with the total number of times the subject changed his sample size from one trial to the next. The binomial test can then be applied with the null hypothesis that a
change consistent with the heuristic will occur with a probability of .5 against the one-sided alternative of the probability greater than or equal to .5.

Note that the test alone is not an adequate response to Question 4. The null hypothesis to be accepted or rejected only states that changes in sample size choices occur on a random basis. Non-random behavior does not necessarily imply use of the anchoring and adjustment heuristic; further evidence is needed.

Cases where there is no change in the sample size from one trial to the next are ignored for purposes of this test. No change in sample size is considered to be ambiguous because it can either be interpreted as being (1) a special case of the heuristic where the adjustment parameter is equal to zero or (2) simply not supportive of the use of the heuristic. Either interpretation will virtually determine the results of each test since a large number of them exist. Table 4 presents the results of the binomial test for each subject.
TABLE 4

BINOMIAL TESTS FOR USE OF THE ANCHORING AND ADJUSTMENT HEURISTIC

<table>
<thead>
<tr>
<th>Subject</th>
<th>Total No. of Changes</th>
<th>No. of Changes in Support</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>16</td>
<td>14</td>
<td>.0001</td>
</tr>
<tr>
<td>D</td>
<td>8</td>
<td>7</td>
<td>.0352</td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>1</td>
<td>N.S.</td>
</tr>
<tr>
<td>F</td>
<td>10</td>
<td>4</td>
<td>N.S.</td>
</tr>
<tr>
<td>G</td>
<td>22</td>
<td>8</td>
<td>N.S.</td>
</tr>
<tr>
<td>H</td>
<td>12</td>
<td>11</td>
<td>.0032</td>
</tr>
<tr>
<td>I</td>
<td>3</td>
<td>3</td>
<td>.1250</td>
</tr>
</tbody>
</table>

Only three of the tests are significant at a level of .05 or less. The test on Subject F is not significant because there were many episodes in his behavior where his primary consideration was the collection of observations of outcomes at certain sample sizes for the purpose of finding the long run optimal solution. Subject G chose sample sizes extremely rapidly, and seemingly at random, after the sample size selection on the first trial. Beyond the use of the case materials, auditors have had no formal training in the task given and Subject G was at a particular loss for the remainder of the experiment. The null hypothesis cannot be rejected for Subjects E and I because of the paucity of data points—even though all of their observations support the
anchoring and adjustment heuristic. Excluding Subject G, a binomial test over all subjects' observations would reject the null at a significance level of less than .001. Thus, we can be very confident that changes in support of the anchoring and adjustment heuristic do not occur on a random basis across subjects. This indirect evidence in support of the use of the heuristic is enhanced by direct evidence from the protocol analysis.

The protocol analysis to discover evidence for the use of the anchoring and adjustment involved inspection of the "meta comment", "comment" and "inference" operators where the sample size argument was scored. The particular types of statements which were interpreted as evidence of the use of the heuristic depended on whether the sample size was increased or decreased by the subject. If the sample size was increased, then the subject should have made some comment to the effect that the sample size was not large enough to enable the DM to infer the actual state of internal controls. If the sample size was decreased, then some comment pertaining to the possibility of the DM doing just as well with a smaller size should have been made.

Table 5 presents the results of this investigation. The number of times the subject made a sample size choice consistent with the anchoring and adjustment heuristic is compared with the number of comments which support that
choice. The third column contains the starting line numbers in the protocol where the remarks are located.

**TABLE 5**

PROTOCOL LINES SUPPORTING THE ANCHORING AND ADJUSTMENT HEURISTIC

<table>
<thead>
<tr>
<th>Subject</th>
<th>Choices Supporting Heuristic</th>
<th>Remarks Supporting Heuristic</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>C*</td>
<td>14</td>
<td>10</td>
<td>133 157 223 250 277 283 332 400 408 426</td>
</tr>
<tr>
<td>D*</td>
<td>7</td>
<td>4</td>
<td>528 545 615 782</td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>1</td>
<td>736</td>
</tr>
<tr>
<td>F</td>
<td>4</td>
<td>1</td>
<td>400</td>
</tr>
<tr>
<td>G</td>
<td>8</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>H*</td>
<td>11</td>
<td>8</td>
<td>322 357 389 450 465 508 588 647</td>
</tr>
<tr>
<td>I*</td>
<td>3</td>
<td>2</td>
<td>820 993</td>
</tr>
</tbody>
</table>

*Significant binomial test results from Table 4

The results of the protocol analysis corroborate the results of the binomial test. The subjects with significant (insignificant) results in the binomial tests made a large (small) number of remarks supporting the use of the anchoring and adjustment heuristic as an indicator of their cognitive processes relative to their total choices made consistent with the algorithm. Re-inspection of Table 2
shows that the subjects who have strong evidence for use of the heuristic also show significant divergence over time from the optimal sample size (Subjects C, D, H and I).

4.5 SIMULATION OF ANCHORING AND ADJUSTMENT

Q5a: Since the anchoring and adjustment heuristic does not require any knowledge of the DM's decision algorithm, can the optimal sample size be approached over time without any knowledge of the decision making algorithm?

Q5b: Is the performance of the anchoring and adjustment heuristic for the information evaluator affected by different decision making styles of the DM, the anchor, or the size of the adjustment?

Tables 6 and 7 present the results of the 18 computer simulations of the anchoring and adjustment heuristic. Levels of the anchor and adjustment were made subjectively, using the protocols as a guide for reasonableness. The decision to restrict the setting of the adjustment parameters to the special case where the upward adjustment and downward adjustment are equal was based on the results of a statistical test. The sample means of the size of the downward adjustments and the size of the upward adjustments
were computed for subjects C, D, H and I (the significant subjects in Table 7), and the parametric test for the difference between two means was applied. We were unable to reject the null hypothesis of no difference at a reasonable significance level, and hence, we restricted ourselves to the special case just described.

TABLE 6

RESULTS OF TESTS FOR TREND ON SIMULATED SUBJECTS

<table>
<thead>
<tr>
<th>Run</th>
<th>Anchor</th>
<th>Adjustment</th>
<th>DM</th>
<th>Anchor</th>
<th>Minus</th>
<th>z</th>
<th>Two-Sided Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25</td>
<td>5</td>
<td>S</td>
<td>5.6</td>
<td>-2.68</td>
<td>.007</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>5</td>
<td>B</td>
<td>22.2</td>
<td>5.91</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>10</td>
<td>S</td>
<td>1.9</td>
<td>-2.36</td>
<td>.018</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>10</td>
<td>B</td>
<td>25.2</td>
<td>5.13</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>20</td>
<td>S</td>
<td>25.0</td>
<td>-2.55</td>
<td>.011</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>25</td>
<td>20</td>
<td>B</td>
<td>68.4</td>
<td>1.44</td>
<td>.150</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>50</td>
<td>5</td>
<td>S</td>
<td>-15.8</td>
<td>3.49</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>50</td>
<td>5</td>
<td>B</td>
<td>-14.9</td>
<td>-6.09</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>50</td>
<td>10</td>
<td>S</td>
<td>-7.5</td>
<td>.97</td>
<td>.332</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>50</td>
<td>10</td>
<td>B</td>
<td>20.1</td>
<td>5.04</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>50</td>
<td>20</td>
<td>S</td>
<td>29.2</td>
<td>-2.83</td>
<td>.005</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>50</td>
<td>20</td>
<td>B</td>
<td>25.0</td>
<td>.97</td>
<td>.332</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>100</td>
<td>5</td>
<td>S</td>
<td>-34.5</td>
<td>4.83</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>100</td>
<td>5</td>
<td>B</td>
<td>-21.0</td>
<td>6.67</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>100</td>
<td>10</td>
<td>S</td>
<td>-54.3</td>
<td>6.37</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>100</td>
<td>10</td>
<td>B</td>
<td>-19.6</td>
<td>1.47</td>
<td>.142</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>100</td>
<td>20</td>
<td>S</td>
<td>-56.4</td>
<td>.96</td>
<td>.337</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>100</td>
<td>20</td>
<td>B</td>
<td>19.0</td>
<td>-2.52</td>
<td>.012</td>
<td></td>
</tr>
<tr>
<td>Decision Maker</td>
<td>Number of Tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>----------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Up</td>
<td>Down</td>
<td>N.S.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistical</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bayesian</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results of each simulation run varied greatly, but unfortunately, we are unable to find any consistent pattern by grouping the results by main effects. The anchoring and adjustment heuristic seems to induce a trend in the sample sizes, but the direction of the trend cannot be predicted from the type of decision maker, anchor, or adjustment.
Chapter V

FINAL REMARKS

The purpose of this dissertation was to study the performance and cognitive processes of auditors in the task domain of information system choice. The specific task given them involved sample size selection for conducting tests of internal control procedures. This resulted in three types of research methodologies: statistical analysis, analysis of concurrent verbal protocols and computer simulation. The results from these three methodologies are summarized, and significant limitations of the research are considered.

The tests for downward trend on each auditor subject yielded little evidence to support the conclusion that subjects are able to approach the optimal sample size (as defined by the normative model used in this research) over time. In fact, evidence for divergence was stronger than evidence for convergence. Tests for upward or downward trend on the average performance of simulated subjects applying the anchoring and adjustment heuristic did not yield information with respect to the efficacy of the anchoring and adjustment heuristic when adjustment parameters, decision makers and anchors were varied.
The results of the tests on human subjects (as opposed to simulated subjects) must be tempered by the shortcomings of using the normative model as a benchmark to evaluate human performance, as was discussed in the literature review. On the other hand, Einhorn and Hogarth (1981) state that one purpose of training is to clarify goals. One implication of the negative results of the tests for convergence is that auditors might benefit from additional training as information evaluators. Yet, the largely negative results of this phase of the research suggest that additional research is needed to determine what criteria auditors utilize in planning audits in addition to research concerning the criteria of expected payoff maximization.

The cognitive processes of the subjects, as represented by the verbal protocols, bear almost no resemblance to the processing required to find the optimal solution using the normative model. There was some attention to payoffs, sampling costs and prior probabilities, but these were not combined to form any type of expected value.

Financial statements were utilized by the subjects to become acquainted with the company, and either to determine the importance of the audit engagement as a whole or the importance of the procedure they are asked to perform in the experimental task. The basic types of processing involved identification of significant changes in a financial statement item from one year to the next and the evaluation
of a financial statement item based on the subject's conception of what a prototypical firm would disclose on its financial statements. No financial ratios were computed. The main differences among subjects were in the amount of effort expended in financial statement analysis and the goal of that analysis.

The flowcharts were utilized to gain overall familiarity with internal control procedures and to identify strengths and/or weaknesses in the system. The attention by some subjects to weaknesses in the system was puzzling, since no justification for this type of process could be found in either the normative model or generally accepted auditing standards. A tentative explanation was supplied based on statements in two of the protocols: weaknesses may help the auditor develop his beliefs as to how much reliance he will place on the system of internal controls before the compliance test is conducted.

There is strong corroborative evidence that at least three subjects used the anchoring and adjustment heuristic for a number of trials. Statistical analysis of subjects' sample size choices demonstrated that sample size changes consistent with the anchoring and adjustment algorithm occurred non-randomly for at least three subjects. A high proportion of choices consistent with the application of the heuristic were supported by remarks in the protocols. Three subjects made only two remarks in total which relate to the anchoring and adjustment heuristic.
The ultimate goal of an empirical examination of auditor information system choice behavior should be to draw inferences to a larger number of auditors than is being sampled. This dissertation constitutes a preliminary step toward the attainment of that goal through the detailed examination of the behavior of a small sample of auditors. The demonstrated validity of verbal protocol analysis to provide answers to research questions has been heavily relied upon. It is an integral component of the research design.

Two characteristics of verbal protocol analysis limit the making of general statements about auditor behavior at this juncture. First, unlike statistically based methods of studying behavior, there exists no ability to quantitatively express confidence in any inference resulting from the data analysis. Second, and relatedly, the vast amounts of data to be closely scrutinized in a verbal protocol study constrains the practical sample size; this study utilized only seven subjects.

Inferential powers were gladly sacrificed in this study to achieve a higher level of descriptive detail of individual problem solving processes. These limitations should be lessened by considering the results of this study in the context of similar studies to be conducted in the future. Each future work would have limited applicability to non-sampled members of the population, but as evidence
accumulates, a general, empirical theory of information system choice by auditors may evolve.
Appendix A

NOTATION SYSTEM FOR EQUATIONS

A lineal presentation of symbols is used in combination with the Roman alphabet (upper and lower case) for all algebraic expressions and equations in this paper. The purpose of this system is twofold: to facilitate readability of equations, and to enable production of the equations by a conventional computer line printer where many special characters are unavailable, and subscripting and superscripting is impossible.

The hierarchy of operations is the same as most high level computer programming languages (e.g., FORTRAN, PL/I). Parentheses are used for two purposes:

1. Alter the sequence of operations
2. Denote the arguments of a function.

Commas are used for two purposes:

1. Delimit the arguments of a function
2. Denote joint random events in a probability function.

A "|" denotes conditional probability in a probability function, and "#" denotes the member of a set. The set name is denoted by an upper case letter, and the set member is denoted by a lower case letter. For example, "n#H" should be read as "n is a member of the set H."
Operators
+ addition
- subtraction
* multiplication
/ division
** exponentiation

The division operator may also be represented in "ratio" form, i.e.,

\[
\frac{\text{expression}_1}{\text{expression}_2}
\]

to facilitate readability. Also, "**" may be used to denote the optimal value of an expression when the expression is a function argument.

Functions

\begin{align*}
P(\cdot) & \quad \text{probability} \\
\text{SUM}(\cdot) & \quad \text{summation} \\
\text{MAX}(\cdot) & \quad \text{maximization} \\
U(\cdot) & \quad \text{utility} \\
E(\cdot) & \quad \text{expectation} \\
\text{VAR}(\cdot) & \quad \text{variance}
\end{align*}

For the iterative functions MAX and SUM, the basic iterative element is denoted below the function. The beginning and ending points of the iterative procedures are always the first and last elements in the set.
Appendix B
INSTRUCTIONS TO SUBJECTS AND CASE MATERIALS

The purpose of this research is to investigate methods by which auditors choose sample sizes, and to compare their sample size choice with a quantitative model of the sample size choice decision. Be assured that we do not expect you to be, or to become, familiar with this mathematical representation of the problem as we know of no examples of its direct application in auditing to date. We only ask that you do your best to be successful in the experimental task. The case study which follows is intended to be as plausible a situation which could occur in auditing as possible, and still allow the quantitative evaluation of certain hypotheses of interest.

Instructions

1. You are requested to perform the task given you by the monitor.

2. While solving the problem, please say out aloud what you are doing or thinking about. If this instruction confuses you, ask the monitor to elaborate.
3. You may use the calculator provided you to perform calculations.

4. Notify the monitor if you need information which is not contained in the problem statement. If possible, he will supply the information to you.

Thank you for your co-operation.
ROHR INDUSTRIES

You are the senior in charge of the audit of Rohr Industries, a manufacturer of cabinets for electronic equipment. This is the first year your firm is engaged to perform the year-end audit, and Rohr received an unqualified opinion the previous year from another Big Eight CPA firm. The company supplied comparative financial statements and accompanying notes are attached (Exhibit A).

You have just completed a review of the internal control procedures over purchases and payables and the results of your inquiries and observations are summarized in the flowcharts and the memorandum contained in Exhibit B. Your manager has examined all of the above information and he has instructed you to choose a random sample of cash disbursements and to trace these transactions back through the system of internal control and to classify each transaction as either complying or not complying with the internal control requirements for that transaction. Non-compliance is defined as failing to adhere to at least one procedure in the purchase/payment cycle. As the subject in this experiment, your only task is to choose the sample size for the compliance test. The actual test of compliance will be simulated by a computer program run by the monitor of the experiment.

The experiment will proceed by first having you choose a sample size. The monitor will then inform you of the following information:
1. **The sample outcome** (error rate in the sample).

2. **The inference made by your manager based on the sample size.** Your manager will make one of four inferences about the true state of compliance with internal controls:

   a. The population error rate is less than or equal to 3% and internal control procedures can be relied on to the maximum extent possible in performing substantive tests.

   b. The population error rate is between 3% and 7%. Internal control procedures are satisfactory, except for some minor departures, and substantive testing will be at a higher level.

   c. The population error rate is between 7% and 12%. There are significant departures from internal control procedures, and extensive substantive testing is necessary.

   d. The population error rate is between 12% and 20%. The maximum amount of substantive testing will be employed since internal control procedures are totally ineffective.

The manager's decision pertaining to his overall assessment of the degree to which the firm can rely on the existing internal controls for substantive testing is based on his general knowledge of the company and its environment, the narrative and flowchart of internal control procedures, the results of your sample, and the costs incurred to the firm resulting from any error in judgement he might make by either over-estimating or under-estimating the true amount that the internal control procedures can be relied upon.

3. **The true state of internal controls.** The true state of internal controls (a, b, c or d, above) becomes known sometime after completion of the audit.
4. **Your net payoff.** The following system of rewards and penalties will accrue to you.

a. In all cases you will be assessed $.20 per transaction sampled.

b. You will win $100 if the manager's decision corresponds with the true state of internal controls.

c. You will lose $4000 if the manager infers "a" or "b", and the true state of internal controls is "d".

d. If the manager errs by one level in either direction, no reward or penalty will accrue to you.

e. A two-level error other than the error discussed previously will cause your winnings to be decreased by $50.

Naturally, the larger the sample size you choose, the greater the likelihood your manager will have of making the correct inference. You will need to balance this intuitive statistical property against the cost of sampling.

You will have the opportunity to learn from experience, because the experiment will be repeated a large number of times. For each repeat of the experiment, assume you are in charge of an audit of another company, but each new company is identical in every respect to Bohr except for the number of cash disbursements in the population (7,000) which do not comply with the stated internal control procedures. This number is determined at random by computer in a manner which is intended to reflect the frequency of error rates observed over all clients audited by Big Eight CPA firms of similar characteristics. Because of your auditing experience, you
probably have a pretty fair conceptualization of error rate frequency, but to make the computer generation process more explicit to you, here is a sample of 100 error rates (in percentages) generated by the computer:

<table>
<thead>
<tr>
<th>8.27</th>
<th>1.39</th>
<th>9.72</th>
<th>5.31</th>
<th>2.48</th>
<th>2.00</th>
<th>4.09</th>
<th>5.32</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.23</td>
<td>0.51</td>
<td>8.47</td>
<td>14.63</td>
<td>1.59</td>
<td>19.79</td>
<td>3.35</td>
<td>1.39</td>
</tr>
<tr>
<td>0.61</td>
<td>0.52</td>
<td>5.40</td>
<td>11.04</td>
<td>8.07</td>
<td>0.69</td>
<td>2.52</td>
<td>5.87</td>
</tr>
<tr>
<td>3.39</td>
<td>12.05</td>
<td>6.11</td>
<td>0.55</td>
<td>6.48</td>
<td>4.89</td>
<td>4.72</td>
<td>2.06</td>
</tr>
<tr>
<td>0.89</td>
<td>1.33</td>
<td>3.78</td>
<td>4.47</td>
<td>3.11</td>
<td>2.53</td>
<td>2.41</td>
<td>2.83</td>
</tr>
<tr>
<td>6.16</td>
<td>1.78</td>
<td>5.37</td>
<td>4.29</td>
<td>3.72</td>
<td>0.28</td>
<td>1.16</td>
<td>9.83</td>
</tr>
<tr>
<td>1.29</td>
<td>1.35</td>
<td>2.08</td>
<td>1.94</td>
<td>1.64</td>
<td>1.20</td>
<td>0.96</td>
<td>2.26</td>
</tr>
<tr>
<td>2.99</td>
<td>2.03</td>
<td>4.63</td>
<td>0.53</td>
<td>2.37</td>
<td>5.12</td>
<td>1.08</td>
<td>7.02</td>
</tr>
<tr>
<td>2.72</td>
<td>2.68</td>
<td>3.90</td>
<td>11.00</td>
<td>1.35</td>
<td>2.02</td>
<td>2.79</td>
<td>5.71</td>
</tr>
<tr>
<td>1.30</td>
<td>2.96</td>
<td>1.84</td>
<td>0.78</td>
<td>6.09</td>
<td>6.47</td>
<td>5.61</td>
<td>4.74</td>
</tr>
<tr>
<td>0.52</td>
<td>3.33</td>
<td>2.24</td>
<td>15.37</td>
<td>2.30</td>
<td>7.94</td>
<td>8.08</td>
<td>1.64</td>
</tr>
<tr>
<td>0.42</td>
<td>1.95</td>
<td>0.29</td>
<td>2.38</td>
<td>2.52</td>
<td>4.13</td>
<td>6.41</td>
<td>15.34</td>
</tr>
<tr>
<td>2.64</td>
<td>3.75</td>
<td>4.68</td>
<td>7.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
SAMPLE SIZE PROGRAM

You need to specify three values to use the sample selection program:

1. Your estimate of the population error rate.
2. An upper confidence bound (UCB)
3. A level of risk.

Definitions of important terms follow to facilitate your use of the program.

1. **Sample selection basis** - Transactions will be drawn from the population randomly. That is, each cash disbursement has an equal chance of being selected, regardless of its magnitude.

2. **Confidence level** - Number of times in 100 that the sample will approximately represent the population (lie within a confidence interval), e.g. 95%, 90%, etc.

3. **Level of risk** - 1 minus the confidence level; for example, if the confidence level was 95%, the level of risk would be 5%.

4. **Population error rate** - The number of transactions in the population which do not comply with the internal control procedures divided by the total number of transactions in the population. This number may be thought of as the probability that a randomly selected transaction does not comply with internal control procedures.

5. **UCB** - The upper limit on the population error rate the auditor is willing to accept.
Flowcharting Symbols

Point of System Entry

Document

Manual Operation

File off line: alphabetically, numerically

File on line

Off-page connector

Destroy

Ledger, Register or Journal
When an order is taken, the engineering department is given a copy of the order to determine the specs. They then forward the materials requirements to the Inventory Control department head.

Based on EOQ and materials needs, Inventory Control determines the need to order materials. When the determination is made that an item should be ordered, the Inventory Control department forwards to the purchasing agent a requirements list from which he is to make purchase orders. This is done daily.
Appendix C

DESCRIPTION OF COMPUTER PROGRAMS

Four types of computer programs were written in PL/I for use in this research. One program was written to generate the results of the statistical test for trend; the details of this program need not be disclosed as its structure is apparent from the discussion of the test in Chapter III. The other three programs have the following functions:

Program 1: Calculation of the expected payoff of each allowable sample size to an information evaluator, given a specific decision maker.

Program 2: Simulation of an anchoring and adjustment information evaluator, given a specific decision maker.

Program 3: Interactive simulation of sample outcomes and decision making behavior of an audit manager during the experimental task.

The three programs have numerous common features, and for this reason, this appendix is organized by discussing important program features rather than by considering each program apart from the others. Also for clarity of exposition, the problem will be stated in terms of urns and
balls rather than the analogous terms of states of internal control and transactions used by auditors.

**PROBLEM STATEMENT**

There are $D \times N$ urns, each containing $N$ balls. $X$ balls in each urn are black, and $N-X$ balls are white. An urn is selected at random from the discrete approximation to a lognormal distribution. An information evaluator chooses a sample size, $ISAMP$, and the urn previously selected is sampled from with replacement. A decision maker is given the sample outcome, and then guesses which of four general types of urns the sample was drawn from. The four urn types are defined as follows:

<table>
<thead>
<tr>
<th>Urn Type</th>
<th>x ≤ A</th>
<th>B ≥ x &gt; A</th>
<th>C ≥ x &gt; B</th>
<th>D ≥ x &gt; C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

where

$x$ is the proportion of black balls in an urn

A = .03
B = .07
C = .12
D = .20
If the urn sampled from belongs to the class of urns guessed by the DM, the information evaluator is rewarded with \( \text{DCLWIN} \) dollars minus \( \text{SAMPRATE} \times \text{ISAMP} \) dollars of sampling cost. Errors of one level in either direction are penalized by \( \text{LEVEL1} \) dollars minus the sampling cost; the penalty for a two-level error is \( \text{LEVEL2} \) dollars minus the sampling cost unless the DM guesses "2" and the true state is "4". In this last circumstance, and for a three-level error, the penalty is \( \text{LEVEL3} \) dollars plus the sampling charge.

**PAYOFF MATRIX**

A payoff matrix is calculated for each sample size. The matrix has four rows, representing the four possible inferences made by the DM, and four columns, representing the four possible states which could obtain. The values of the variables used to calculate the payoff matrices were set for all runs as follows:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{SAMPRATE} )</td>
<td>$0.02</td>
</tr>
<tr>
<td>( \text{DOLWIN} )</td>
<td>100.00</td>
</tr>
<tr>
<td>( \text{LEVEL1} )</td>
<td>0.00</td>
</tr>
<tr>
<td>( \text{LEVEL2} )</td>
<td>50.00</td>
</tr>
<tr>
<td>( \text{LEVEL3} )</td>
<td>4000.00</td>
</tr>
</tbody>
</table>

**PRIOR PROBABILITIES AND RANDOM VARIATE GENERATION**

The lognormal distribution is used to generate states, and hence, it is the prior probability distribution.
The lognormal distribution has two parameters, a mean (MUX) and a variance (VARX). If Y is normally distributed with mean MUY, and variance VARY, then

\[(17) \quad MUX = e^{**(MUY+VARY/2)}\]

and

\[(18) \quad VARX=e^{**(2*MUY+VARY)*(e**VARY-1)}\]

Fishman (1973). After some subjective experimentation, it was decided that when MUY is set to and VARY is set to the resulting lognormal distribution yields a prior probability distribution which is judged to be a reasonable representation of what an auditor might face when performing compliance tests.

The prior probabilities of the four states were calculated by integrating the lognormal distribution over the four intervals \([0,A], [A,B], [B,C], [C,D]\), and dividing each result by the integral \([0,D]\). Note that the division operation is necessary because the actual probability distribution used is truncated at D.

It was also necessary to calculate the prior probability of each \(n=1,2,...,N\). This was accomplished by converting the continuous distribution to a discrete distribution by performing \(N\) integrations, \([0,1], [1,2], ... , [N-1,N]\). Simpson's rule (Swokowski, 1975) was used to approximate all integrals, as no integral exists for the lognormal probability density function.
For a description of the algorithm used to generate random variates from a lognormal distribution, see Fishman (1973).

**CALCULATION OF THE OPTIMAL SAMPLE SIZE**

The expected value of each sample size is calculated and the optimal sample size is the one corresponding to the maximum value. To calculate the expected value of a sample size, the following procedures are executed for each sample size, MINSAMP to MAXSAMP:

1. Calculate the probability of each sample outcome given the urn chosen and the sample size. I.e., calculate the "uncompressed" message/state matrix for the current sample size, SAMPSIZE. The resultant matrix has (SAMPSIZE+1) rows and (N+1) columns. The "compressed" message/state matrix is calculated by reducing the uncompressed matrix to four columns. Each column represents a class of urns.

2. Calculate the payoff matrix, inclusive of sampling costs.

3. Equation (7) is used to calculate the expected value of each sample size, and the optimal sample size. Inspection of the equation reveals that the
data inputs to the equation are the prior probabilities of the four states, the choice matrix, \( \text{CELL} \), and for each sample size, the payoff matrix and the compressed message/state matrix.

**SIMULATION OF DECISION MAKERS**

The DM's inferences are calculated for each sample size and each sample outcome and is stored in the choice matrix, \( \text{CELL} \). The matrix has \( \text{MINsamp} \) to \( \text{MAXsamp} \) rows, corresponding to the possible sample sizes, and \( (\text{MAXsamp}+1) \) columns corresponding to the number of black balls in the sample. We include descriptions of four classes decision making algorithms, even though only two types of DM's are used in this study.

**Representativeness DM**

The representativeness decision maker's rule is:

\[
\begin{align*}
\text{CELL}(I,J) &= 1 \text{ if } 0 < \frac{J}{I} < A \\
\text{CELL}(I,J) &= 2 \text{ if } A < \frac{J}{I} < B \\
\text{CELL}(I,J) &= 3 \text{ if } B < \frac{J}{I} < C \\
\text{CELL}(I,J) &= 4 \text{ if } C < \frac{J}{I} < D
\end{align*}
\]

**Hypothesis Testing DM**

The hypothesis testing DM uses a series of binomial tests (Hollander and Wolfe, 1973). For the first test, the null hypothesis is:

\[
H(0): p \leq C
\]
If the null is rejected, then $\text{CELL}(I,J)$ is set to "4". If the null cannot be rejected, then the hypothesis

$$H(0): p \leq B$$

is tested, and a similar procedure is followed until either a null hypothesis is rejected or it must be concluded that $\text{CELL}(I,J)$ should be set to "1".

The normal approximation to the binomial test was used when the sample size was greater than 40. The significance level was set at 10 percent for all tests.

**Bayesian and Conservative Decision Makers**

Recall that Equation (10) augments the Bayesian probability revision procedure with a "conservatism" parameter:

$$(10) \quad Z'(1) = (L * c) \cdot Z(0)$$

where:

- $L$ is the likelihood ratio
- $Z(0)$ is the prior odds
- $Z(1)$ is the posterior Bayesian odds
- $c$ is the conservatism parameter ($0 < c < 1$)
- $Z'(1)$ is the posterior conservative odds

If $c$ is equal to one, then the equation represents Bayesian updating instead of conservative updating. Therefore, the same algorithm can be used to simulate both Bayesian and conservative decision making. The algorithm works as follows:

1. Set $c$
2. Observe an outcome and a sample size, and calculate the posterior odds using Equation (10).

3. Convert the posterior odds to a posterior probability.

4. Set \( \text{CELL}(I,J) \) to the action (Choose urn type "1", "2", etc.) with the highest expected value based on the posterior probabilities. This operation is embodied in Equation (3).

**Simulation of Information Evaluator**

A flowchart for the simulation of the information evaluator using the anchoring and adjustment heuristic is contained in Chapter III.
Appendix D
SUBJECT PROTOCOLS

SUBJECT C

1 YOU ARE THE SENIOR IN CHARGE OF THE AUDIT
2 - - -
3 HERE IS A SAMPLE OF 100 ERROR RATES GENERATED BY THE
4 COMPUTER.
5 WHAT YOU WANT ME TO DO IS PICK THE SAMPLE SIZE FOR
6 CASH
7 DISBURSEMENTS TESTS, RIGHT?
8 (RIGHT)
9 I SEE IN 1979 SAYS NOTE 2,
10 SO I SHOULD LOOK AT NOTE 2 AND SEE WHAT'S GOING ON IN
11 1979
12 NOTE 2,
13 PRIOR PERIOD ADJUSTMENT,
14 FINANCIAL STATEMENTS FOR 1980 HAVE BEEN RETROACTIVELY
15 RESTATED TO CORRECT AN OVERSTATEMENT OF INVENTORIES
16 RESULTING
17 FROM THE INADVERTANT INCLUSION OF APPROXIMATELY
18 $45,000 IN
19 CUSTOMER PRODUCTS AND FINISHED GOODS INVENTORIES
20 THIS PRIOR PERIOD ADJUSTMENT HAD THE EFFECT OF
21 REDUCING INCOME
22 BEFORE EXTRAORDINARY ITEMS AND NET INCOME FOR 1979 BY
23 $19,779
24 OR 28 CENTS PER SHARE.
25 TOTAL BALANCE SHEET'S ABOUT THE SAME IN 1980 AS IT WAS
26 IN 1979.
27 SALES APPROXIMATELY THE SAME IN BOTH YEARS,
28 A LITTLE BIT HIGHER IN 1980.
29 COST OF SALES ARE A LOT HIGHER IN 1980.
30 GOT AN EXTRAORDINARY ITEM
31 BENEFIT OF TAX LOSS CARRYFORWARD,
32 GOES IN,
33 COMES BACK OUT,
34 NETS OUT TO ZERO.
35 EARNINGS PER SHARE IS DOWN SIGNIFICANTLY IN 1980,
36 FROM $5.63 IN '79 TO $1.68,
37 SO I WANT TO SEE IF THERE'S ANY CHANGE IN COMMON
38 STOCK,
39 THE SHARES.
LOOKS LIKE THERE HASN'T BEEN ANY CHANGE,
THERE'S A REFERENCE TO NOTE 5.
DURING '80, THE COMPANY REDEEMED 1,032 SHARES IN
NON-VOTING
COMMON STOCK,
SO IT LOOKS LIKE A DECREASE IN THAT.
EARNINGS PER SHARE ISN'T RELATED TO THE NUMBER OF
SHARES OUTSTANDING,
IT'S PRIMARILY NET INCOME'S DOWN QUITE A BIT,
MOST OF IT'S DUE TO COST OF SALES BEING HIGHER.
HAD A DECREASE IN WORKING CAPITAL OF $76,000 IN 1980
SO ALL THIS STUFF TELLS ME IT LOOKS LIKE 1980 WASN'T A
VERY GOOD YEAR.
FOR CASH DISBURSEMENTS.
I WANT TO SEE WHAT TYPE OF CONTROLS THEY HAVE.
OK, I'M LOOKING AT APPENDIX B NOW.
OK, PURCHASES, LOOKS LIKE THE ORDERS COME FROM THE
ENGINEERING
DEPARTMENT.
THEY HAVE APPROVALS ON PURCHASES,
ANYTHING OVER A THOUSAND THE PRESIDENT HAS TO APPROVE,
ANYTHING OVER 5000,
THE BOARD OF DIRECTORS HAS TO APPROVE,
FOR ASSET PURCHASES.
ANY PO OVER 10,000 FOR PRODUCTION MATERIAL,
THE PRESIDENT HAS TO APPROVE.
ALL OTHER PO'S HAS TO BE APPROVED BY THE PURCHASING
AGENT.
ENGINEERING DETERMINES THE MATERIAL REQUIREMENTS FOR
SPECIAL ORDERS.
LOOKS LIKE THEY HAVE 8 COPIES OF THE PO.
LOOKS LIKE COST ACCOUNTING TAKES CARE OF IT,
ANY VARIANCES MUST BE ON A STANDARD COST SYSTEM,
LOOKS LIKE.
LOOKS LIKE THEY MAY ACCEPT THE PO WITH THE FREIGHT
BILL
AND RECEIVING FILE,
REVIEW THE UNMATCHED RECEIVERS EVERY MONTH,
SEND THE INVOICE TO THE PURCHASING AGENT, WHO REVIEWS
IT,
I DON'T KNOW IF THEY REALLY NEED TO DO THAT,
I THINK ACCOUNTING COULD PROBABLY DO THAT.
PURCHASING AGENT REVIEWS THE INVOICE.
PAGE TWO STARTS HERE?
PURCHASING AGENT APPROVES THE PO,
SEND'S IT TO COST ACCOUNTING WHO ENTERS THE INVENTORY
CODING,
AND HE GIVES THE INVOICE TO AN ACCOUNTING CLERK FOR
PAYMENT.
ACCOUNTING CLERK TAKES ANY DISCOUNTS AVAILABLE,
CHECKS THE INVOICE FOR CLERICAL ACCURACY,
PAYMENTS ARE APPROVED BY THE CONTROLLER.
AUTHORIZED CHECK SIGNERS ARE THE PRESIDENT AND THE CONTROLLER.

LOOKS LIKE CHECKS ARE GENERATED BY AN NCR OPERATOR, WHO PUTS
THE CHECK AMOUNTS AND CHECK NUMBERS INTO THE CHECK REGISTER.

THE CHECK IS PRINTED OUT AND SENT TO THEM, THE VENDOR.

IN THE CHECK REGISTER, THE ENTRY IS MADE IN THE LEDGER.

MEMO,

INVENTORY CONTROL,

WHEN AN ORDER IS TAKEN THE ENGINEERING DEPARTMENT IS GIVEN
A COPY OF THE ORDER TO DETERMINE THE SPECS,
THEY THEN FORWARD THE MATERIALS REQUIREMENTS TO THE INVENTORY
CONTROL DEPARTMENT HEAD.

BASED ON EOQ, THE MATERIAL NEEDS ARE DETERMINED, THE NEED TO ORDER MATERIALS.

IF A DETERMINATION IS MADE THAT AN ITEM SHOULD BE ORDERED,
THE INVENTORY CONTROL DEPARTMENT FORWARDS TO THE PURCHASING AGENT, THE REQUIREMENTS LIST FROM WHICH HE IS TO MAKE PURCHASE ORDERS. THIS IS DONE DAILY.

WELL, FROM THIS,
ONE THING I DON'T KNOW,
I DON'T KNOW HOW MANY CHECKS ARE THEY WRITING.

IS THAT 7000?

IS THAT WHAT WE'RE SAYING?

7000?

OK

IT LOOKS LIKE THEY HAVE PRETTY GOOD CONTROLS OVER PURCHASE
APPROVALS.

7000 MIGHT BE ABOUT HOW MUCH A MONTH?

583 A MONTH.

OK. I'M JUST TRYING TO THINK HERE,

IF WE NEED TO TEST PROCEDURES A LOT
OR IF IT WOULD BE BETTER MAYBE AT THE END OF THE YEAR JUST
TO DO SUBSTANTIVE TESTING.

I THINK MAYBE,

20 AND 30 MIGHT BE A GOOD ENOUGH SAMPLE SIZE,

IT MAY JUST BE MORE EFFICIENT THAN TO SUBSTANTIALLY TEST

THES BALANCES AT THE END OF THE YEAR IF THEY TAKE A YEAR-END

PHYSICAL IN FIRM RECEIVABLES.

I'D SAY THAT 25 WOULD BE GOOD ENOUGH.
OK, HERE'S THE REPORT OF YOUR FIRST PICK. THAT'S HOW WE SCORE
AT THE TOP. YOU CHOSE A SAMPLE SIZE OF 25, YOUR TEST DETECTED
NO DEFECTIVE TRANSACTIONS, THE MANAGER CHOSE THAT THE STATE OF
INTERNAL CONTROLS WAS A, THE ACTUAL STATE WAS B, AND SO YOU
LOST $5.00)
I LOST $5.00?
(YES, IT COST 20 CENTS PER SAMPLE SELECTED, AND SINCE THERE
WAS ONE LEVEL ERROR THERE WAS NO PENALTY, SO YOUR LOSS WAS THE
SAMPLING CHARGE.)
HOW DO I KEEP GOING UNTIL I START WINNING MONEY, OR?
(WELL, WE'RE GOING TO KEEP GOING FOR AN UNSTATED NUMBER OF TRIALS)
SO WE SAID IF POPULATION ERROR WAS LESS THAN OR EQUAL TO 3%
INTERNAL CONTROLS SHOULD BE RELIED UPON TO THE MAXIMUM POSSIBLE.
THE TRUE ERROR RATE WAS BETWEEN 3 AND 7%.
SAID INTERNAL CONTROLS ARE SATISFACTORY,
EXCEPT FOR SOME MINOR DEPARTURES
AND SUBSTANTIVE TESTING WILL BE AT A HIGHER LEVEL, OK.
SO, EVIDENTLY I DIDN'T PICK ENOUGH TO BE STATISTICALLY VALID.
I THINK I'D TAKE 40,
JUST BASED ON EXPERIENCE.
---
IT STILL SAID IT WAS A, AND THE REAL STATE WAS B,
SO I LOST $8.00 THIS TIME,
SO THE REAL STATE WAS B,
IF WE TEST AT 40,
THE POPULATION ERROR RATE IS BETWEEN 3 AND 7%,
BUT WE STILL DIDN'T FIND ANY ERRORS,
IS THAT RIGHT?
OR IS THIS 2?
SAMPLE OUTCOME, IS THAT HOW MANY MISTAKES YOU HAD?
2?
I DON'T KNOW WHY IT WOULD SAY A WAS THE STATE THEN.
2 OUT OF 40 IS WHAT?
1/20TH, THAT'D BE ABOUT 5%.
WOULDN'T THAT PUT US INTO B?
POPULATION ERROR RATE?
(WELL, YOU HAVE TO REMEMBER THAT YOU'RE NOT DECIDING,
THE INFERENCE IS MADE BY SOMEBODY ELSE OTHER THAN YOU.)
OK. SO IT SEEMS LIKE IF YOU TEST AT 40 AND THERE'S 2 MISTAKES
THEN YOU WOULD BE AT THE B LEVEL, OK.
SO, WE'RE STILL OFF.
I DON'T KNOW WHAT THAT WOULD TELL US,
I GUESS, I DON'T KNOW,
IF I HAD 2 WRONG I DON'T KNOW IF I'D WANT TO DO MORE THAN 40,
IT MIGHT BE MORE EFFICIENT JUST TO SUBSTANTIVE TEST THIS AT
THE END OF THE YEAR AT A HIGHER LEVEL.
I GUESS I'M READY FOR ANOTHER ONE.
HOW ABOUT, 50?
...
4 ERRORS,
OUR DECISION WAS B AND THE TRUE STATE WAS B,
SO WE FINALLY ARE IN AGREEMENT,
PAYOFF THIS TIME.
WHAT'S THE PAYOFF?
(YOU WON $90.00)
I WON 90 THAT TIME?
(BECAUSE THE TWO AGREE, YOU WON A HUNDRED, LESS $10 SAMPLING CHARGE, $90)
IT SEEMS LIKE WHAT IT'S INDICATING TO ME SO FAR IS
THAT THERE IS A TENDENCY TO UNDERESTIMATE YOUR SAMPLE SIZE,
AT LEAST IN THIS SITUATION.
I MEAN I HAD TO GET UP TO 50 IN THE SAMPLE SIZE BEFORE I GOT TO THE TRUE STATE.
I DON'T KNOW IF I PICKED ANOTHER 50 IF I'D COME OUT WITH THE SAME DECISION AS WHAT THE TRUE STATE WAS,
BUT, SEEMS LIKE IT'S LEADING ONE TO BELIEVE THAT JUDGMENTAL
SAMPLE SIZES AREN'T LARGE ENOUGH.
THAT'S ABOUT ALL I CAN SAY ABOUT WHAT WE'VE GOT HERE SO FAR.
WITHOUT KNOWING A WHOLE HEIL OF A LOT ABOUT THE CLIENT,
I MEAN WE LOOKED AT THIS BUT I CAN'T IN A HALF HOUR GET A REAL GOOD FEEL FOR WHAT THE CLIENT'S LIKE,
WELL, 50 WORKED PRETTY GOOD,
I'D STAY WITH 50.
(OK)
...
(UNCLEAR).--THE STATE'S A.
SEEMS LIKE THIS COMPANY HERE'S DOING A BETTER JOB OF FOLLOWING
PROCEDURES THAN THE OTHER COMPANIES,
SAMPLE SIZE OF 50 I THINK WOULD BE PRETTY SUBSTANTIAL
AND IF YOU TESTED 50 ITEMS AND THERE WERE NO ERRORS,
THERE WOULD BE A GOOD DEGREE OF RELIANCE YOU COULD PLACE
ON INTERNAL CONTROLS.
I THINK WHAT YOU WOULD NEED TO DO THEN
IS JUST LOOK FOR UNUSUAL OR JUDGMENTAL TYPE OF AREAS LIKE SAY,
INVENTORY RESERVES OR ACCOUNTS RECEIVABLE RESERVES,
YOU'D HAVE TO FOCUS IN ON THOSE AT THE END OF THE YEAR AND,
THIS I THINK COULD GREATLY REDUCE YOUR SUBSTANTIVE TESTING AT
THE END OF THE YEAR,
IT LOOKS LIKE.
THAT'S ALL I'D SAY.
LOOKS LIKE I WON SOME MORE MONEY.
YOU WANT TO TRY ANOTHER ONE, OR?
(YEAH, I'D LIKE TO KEEP GOING UNTIL THE EXPERIMENT ENDS,
I'M NOT GOING TO DISCLOSE HOW MANY TRIALS,
IF YOU FEEL THAT FURTHER ONES WOULD BE USELESS,
OR IF YOU'RE GETTING FATIGUED, THEN...)
OH, I'M NOT GETTING FATIGUED,
NOT AT ALL,
I'M JUST SAYING THIS COMPANY LOOKS LIKE IT'S WELL RUN,
IS ALL.
I WANT TO TRY 40,
GO BACK TO 40.
...
ONE ERROR,
DECIDED IT WAS AN A STATE,
AND IT WAS AN A STATE,
ONE ERROR AGAIN OUT OF 40 IS NOT TOO BAD.
LOOKS LIKE OUR DECISION,
AS FAR AS THE STATE OF CONTROLS THAT EXISTED WAS ACCURATE.
THAT'S ALL I'D SAY ABOUT THAT.
TAKE ANOTHER ONE?
LET'S GO WITH 35 AND SEE WHAT HAPPENS.
...
THREE MISTAKES OUT OF 35,
WE DECIDED IT WAS A B STATE,
AND IT WAS A B STATE,
SO IT LOOKS LIKE THIS SAMPLE SIZE IS A GOOD INDICATION
OF WHAT
TYPE OF CONTROLS WERE THERE.
IT'S KIND OF INTERESTING,
THE SAME TYPE OF SITUATION,
WHEN YOU PICK 40 IN THE SAMPLE SIZE,
YOU GET 2 ERRORS,
SO YOU DECIDED THERE WAS AN A STATE BUT THERE WAS
REALLY A B STATE.
SOUNDS LIKE WHEN YOUR SAMPLE SIZE IS RELATIVELY LOW,
YOU DON'T HAVE AS GOOD A CHANCE OF PREDICTING IT AS
WHAT THE TRUE STATE IS,
AS IF THERE'S 50.
SEEMS LIKE EVERYTIME I CHOSE 50 WHAT WE DECIDED THE STATE WAS,
OR WHAT WE THOUGHT THE STATE OF THE CONTROLS WERE,
149

THAT'S WHAT THEY WERE.
I'D LIKE TO TRY 50 AGAIN.

6 ERRORS WE DECIDED WAS A C STATE
AND IT ACTUALLY WAS A C STATE,
LOOKS LIKE 50 SEEMS TO BE A PRETTY GOOD SAMPLE SIZE.
WE'VE TRIED IT THREE TIMES AND
WE'VE BEEN RIGHT 3 TIMES WITH IT,
LET'S GO DOWN AND TRY 45 AND SEE WHAT HAPPENS.

NO ERRORS,
DECIDED IT WAS AN A STATE,
AND IT WAS AN A STATE.
LOOKS LIKE SOME SAMPLE SIZE SOMEWHERE OVER 40 APPEARS
TO BE THE RIGHT SAMPLE SIZE,
LET'S TRY 45 AGAIN.

HAD ONE MISTAKE,
WE THOUGHT IT WOULD BE AN A STATE AND
IT REALLY WAS AN A STATE.
LET'S TRY 43

HAD 5 ERRORS,
DECIDED IT WAS A B STATE,
IT REALLY WAS A B STATE.
43 APPARENTLY SEEMS LIKE THERE WAS ENOUGH FOR THE
MANAGER TO MAKE A GOOD DECISION.
LET'S TRY 43 AGAIN AND SEE WHAT HAPPENS

SAMPLE OUTCOME IS ONE ERROR,
DECIDED IT WAS AN A STATE,
IT WAS REALLY A B STATE.
THAT INDICATES THAT THE SAMPLE SIZE PROBABLY WASN'T
ENOUGH.
LET'S TRY 50 AGAIN

HERE WE DECIDED IT WAS AN A STATE,
IT WAS REALLY A B STATE,
IT APPEARS THAT MAYBE 50 ISN'T A LARGE ENOUGH SAMPLE
SIZE TO GIVE YOU A TRUE INDICATION OF WHAT TYPE OF CONTROLS
EXISTED WITH THIS COMPANY.
LET'S TRY 55

9 ERRORS,
WE THOUGHT IT WOULD BE C,
STATE WAS C.
THAT SAMPLE SIZE WAS GOOD ENOUGH TO GIVE US AN ACCURATE
READING ON WHAT THE TRUE STATE OF THE CONTROLS WERE.
I'D TRY 55 AGAIN

...
295 2 ERRORS,
296 WE THOUGHT IT WAS AN A STATE,
297 IT REALLY WAS AN A STATE.
298 I THINK I'D TRY 55 AGAIN
299 ...
300 NO ERRORS,
301 THOUGHT IT WAS AN A STATE,
302 IT WAS AN A STATE,
303 THIS COMPANY LOOKS LIKE THEY'RE FOLLOWING THEIR
PROCEDURES
304 PRETTY WELL.
305 NO ERRORS OUT OF 55,
306 YOU COULD PLACE A HIGH DEGREE OF RELIANCE ON THEIR
CONTROLS.
307 I'D SAY IT'S EXCEPTIONALLY STRONG,
308 YOU WOULDN'T SEE TOO MANY COMPANIES,
309 I DON'T THINK,
310 IF YOU SELECTED 55 DISBURSEMENTS THAT WOULDN'T HAVE
ANY
311 ERRORS IN ONE OR ANOTHER PROCEDURE.
312 I'D TRY 55 AGAIN
313 ...
314 SAMPLE SIZE - 55
315 SAMPLE OUTCOME - 4 ERRORS,
316 DECISION - DECIDED IT WAS A B,
317 STATED IT WAS REALLY A C.
318 SIDE 2 OF TAPE C STARTS HERE
319 (UNCLEAR)... B AND A C STATE IS,
320 B STATE IS 3 TO 7% POPULATION ERROR RATE
321 CONTROLS ARE SATISFACTORY WITH SOME MINOR DEPARTURES.
322 C IS ERROR RATES BETWEEN 7 AND 12%,
323 CONSIDERED SIGNIFICANT DEPARTURES FROM INTERNAL
CONTROL PROCEDURES
324 AND EXTENSIVE SUBSTANTIVE TESTING IS NECESSARY.
325 I'D SAY BASED ON THIS IT LOOKS LIKE WE MADE A POOR
DECISION ON
326 WHAT THE STATE OF CONTROLS WERE.
327 WE THOUGHT WE COULD RELY ON THEM,
328 TO A PRETTY LARGE DEGREE AND REALLY,
329 THE CONTROLS WEREN'T THAT GOOD AT ALL.
330 COULD GIVE US PROBLEMS ON MAYBE ISSUING FINANCIAL
STATEMENTS
331 THAT WEREN'T PROPERLY STATED OR FAIRLY STATED.
332 LOOKS LIKE WE SHOULD HAVE DONE MORE SUBSTANTIVE
TESTING AT
333 THE END OF THE YEAR.
334 LET'S TRY 60.
335 ...
336 SAMPLE OUTCOME SHOWED 2 ERRORS,
337 WE THOUGHT IT WAS AN A STATE,
338 IT WAS AN A STATE.
339 IN THIS SITUATION, LOOKS LIKE WE MADE A GOOD DECISION
ON
340 EVALUATING WHAT TYPE OF INTERNAL CONTROLS EXISTED.
151

341 LET'S TRY 100 AND SEE WHAT HAPPENS.
342 ..
343 19 ERRORS,
344 WE DECIDED IT WAS A C STATE,
345 WAS REALLY A D STATE
346 DIFFERENCE BETWEEN C AND D IS POPULATION ERROR RATE IS
347 FOR C, AND D IS 12 TO 20%.
348 UNDER A C STATE WE DO EXTENSIVE SUBSTANTIVE TESTING,
349 UNDER D WE DO THE MAXIMUM AMOUNT OF SUBSTANTIVE
350 TESTING.
351 I THINK WE THOUGHT THE INTERNAL CONTROL PROCEDURES
352 WERE LACKING
353 IN MANY AREAS,
354 UNDER D WE THOUGHT THEY WERE TOTALLY INEFFECTIVE,
355 LOOKS LIKE WE MADE A WRONG DECISION AS FAR AS THE TRUE
356 STATE
357 OF THE CONTROLS,
358 A COMPANY THIS SIZE, WHEN I GET UP TO A SAMPLE SIZE OF
359 100, AND I STILL CAN'T RELY ON THE RESULTS OF THE
360 TESTING
361 IF YOU'RE LOOKING AT 100 DISBURSEMENTS,
362 THEN I BEGIN TO, I WOULD WONDER IF IT'S NECESSARY TO
363 DO THIS,
364 IT PROBABLY WOULD BE MORE EFFICIENT AND EFFECTIVE FROM
365 A TIME AND DOLLARWISE POINT OF VIEW,
366 JUST TO DO QUITE A BIT OF SUBSTANTIVE TESTING AT THE
367 END OF THE
368 YEAR.
369 AGAIN, THE ERRORS, I WOULD BE CONCERNED WHAT TYPE OF
370 ERRORS,
371 IT LOOKS LIKE THESE ARE PRETTY IMPORTANT ERRORS SINCE
372 IT SAYS
373 CONTROL PROCEDURES ARE TOTALLY INEFFECTIVE,
374 IF IT'S A D STATE, SO
375 IT LOOKS LIKE CONTROLS CAN'T BE RELIED UPON AND,
376 LITTLE SURPRISED THAT IF WE HAD A SAMPLE SIZE OF 100,
377 WE STILL CAN'T MAKE A PROPER DECISION ON WHAT THE TRUE
378 STATE
379 OF CONTROLS ARE.
380 (SO TO WHAT DO YOU ATTRIBUTE THE INABILITY TO MAKE A
381 PROPER
382 DECISION?)
383 WELL IT'S EITHER AN ERROR IN JUDGMENT ON THE MANAGER,
384 HE MAY BE MAKING A JUDGMENTAL DECISION,
385 WITHOUT HAVING ANY TRUE STATISTICS TO FALL BACK ON,
386 MAYBE WE'RE NOT RELYING ON STATISTICAL DATA AS MUCH AS
387 WE SHOULD.
388 PROBABLY I'D HAVE TO DO SOME TYPE OF STATISTICAL STUDY
389 AND SEE
390 WHAT THE CONFIDENCE LEVEL IS, AND THINGS LIKE THAT.
Again, I'm not that familiar with statistics but I think you can work through formulas and it would tell you if you have a population size of 7000, sample size of 100, you should be able to predict within a certain confidence level what the reliability of the tests are and, I would wonder if he really followed statistics, or whether he just made a judgmental decision as to the true state of controls are C.

I'd say he probably just made a judgmental decision that wasn't statistically valid.

Again, I would begin to wonder if we had to test 100 items and we still couldn't determine the true state of the internal controls.

I think it would probably be much more efficient just to do substantive testing at the end of the year, confirm receivables and observe the inventory, things like that.

Again, I don't know if it was, how the manager reached his decision that there was a C state but, I'm surprised that if we had such a large sample size we couldn't identify the true state of the controls.

Why don't we try 150 and see what happens.

The sample outcome of size of 150, we had 22 errors, we decided it was a C state, it really was a C state, correctly identified what the state of controls were, sample size told us what we wanted it to tell us, but it's, to me it seems like a large sample size.

Let's try 125.

---

We had one error out of 125 in the sample size, decided it was an A state, really was an A state, would hope 125 could give us a good indication of what was really going on.

And 1 error out of 125 would indicate that the controls are working effectively.

Why don't we try 100 again.

---

Ok, sample size of 100, we had 20 errors, decided it was a D state,
It really was a D state.
Sample size was able to tell us what the true state of the controls were.
Looks like maybe the sample size was too large,
I mean, maybe we could have sampled half as many and
still come up with the same decision.
Why don't we try 50?
...
Sample size 50 - we had three errors,
we decided it was an A state,
it was really a B state.
Difference between A and B is
A state is less than 3% error rate,
B state is between 3 and 7% error rate.
This'd probably cause us to do less substantive testing than what
we really should have done.
(Ok, why don't I say this, why don't we do 5 more and
then at the end of those 5, why don't you tell me what the
sample size, the best sample size, would have been)
Ok
Let's try 115.
...
Ok, we had one error out of 115,
we decided it was an A state,
it was an A state.
Looks like our sample size was large enough to tell us
what the true state of the controls were.
Let's try 125.
...
Sample size of 125,
sample outcome was 3 errors, we decided it was an A state,
it was really a B state, so
we didn't properly identify what the true state of the controls
were.
Let's try 150.
...
Sample size of 150,
we had a sample outcome - 8 errors,
we decided it was an A state,
it was really a B state.
So we were wrong.
Let's try a sample size of 175.
...
175,
we had 23 errors,
we decided it was a C state,
it really was a C state.
Let's try 175 again.

OK, sample size of 175,

had 18 errors,

decided it was a B state,

it really was a B state.

Looks like the sample size was large enough to tell us what

the true state of the controls was.

Let's try 175 again.

Had 8 errors,

we decided it was an A state,

it was really a B state,

so sample size didn't tell us what the true state of

the controls was,

that's interesting.

seems like an awful large sample size for a company of

this size.
SUBJECT D

1 I'M A SENIOR IN CHARGE OF AN AUDIT OF ROHR INDUSTRIES
2 YOU HAVE JUST COMPLETED A REVIEW OF THE INTERNAL CONTROL
3 PROCEDURES OVER PURCHASES AND PAYABLES, THE RESULTS OF YOUR
4 INQUIRIES AND OBSERVATIONS ARE SUMMARIZED IN THE FLOW CHARTS,
5 AND THE MEMORANDUM CONTAINED IN APPENDIX E.
6 SO I'M LOOKING AT APPENDIX B HERE
7 I SEE A FLOW CHART HERE,
8 SHOWING HOW FROM ENGINEERING, WORK ORDERS OR PURCHASE ORDERS ARE
9 DEVELOPED,
10 GOES INTO LOOKS LIKE, THE PURCHASING DEPARTMENT,
11 AND NUMEROUS COPIES OF THE PURCHASE ORDER ARE MADE,
12 A NUMBER OF WHICH ARE KEPT IN RECEIVING,
13 PURCHASE ORDERS AND PURCHASING DEPARTMENT ARE FILED AND
14 EVENTUALLY THEY ARE MATCHED WITH THE INVOICE RECEIVED FROM
15 THE VENDOR.
16 THE COPIES IN RECEIVING ARE COMPARED WITH THE BILL OF LADING
17 AND RECEIVING REPORTS,
18 AND SENT BACK TO PURCHASING.
19 THE PURCHASE ORDER, THE FREIGHT BILL AND THE RECEIVING REPORT,
20 AND VENDOR'S INVOICE,
21 ARE ALL MATCHED TOGETHER, AND
22 IT LOOKS LIKE IT'S SENT BACK TO THE PURCHASING AGENT.
23 TALKS ABOUT THAT THERE ARE SOME CERTAIN CONTROLS IN THE
24 PURCHASE DEPARTMENT AS FAR AS APPROVAL.
25 THE PURCHASE ORDER IS COMPARED TO LIKE, STOCK ITEMS ON HAND.
26 ENGINEERING DETERMINES MATERIAL REQUIREMENTS.
27 IN THE PURCHASING DEPARTMENT,
28 WE HAVE CONTROLS OVER
29 YOU HAVE A COST ACCOUNTANT TO DETERMINE THE COST DISTRIBUTION AND
30 THE CALCULATION OF ANY VARIANCE.
31 LOOKS LIKE THEY ALSO HAVE PROCEDURES FOR PARTIAL SHIPMENTS
32 WHAT'S PA APPROVAL?
33 (OR, PURCHASING AGENT)
34 PURCHASING AGENT APPROVAL, FOR STOCK PURCHASES SUPPLIES AND
35 OTHERWISE A DEPARTMENTAL (UNCLEAR) NORMAL DEPARTMENTAL PURCHASE.
Then I have the accounts payable and cash disbursements flow.

Chart in front of me.

And they receive from, where is this coming from? (That's the page connected to it)

Page 2, OK.

And after the purchasing departments match all the invoices, purchase orders, freight bills, receiver reports, and the

purchasing agent has reviewed and approved it, it goes to accounts payable,

whereby they forward it to cost accounting,

who stamps the invoice as paid,

and it enters the inventory account coding.

The invoice is then forwarded on,

where it is prepared for payment.

The accounting clerk will then take any discounts,

run a tape, do some clerical accuracy tests.

And then they get the payments approved by the controller.

The invoice is then passed on to the MCR operator,

where it is entered in by vendor by date,

and then at a different point in time,

a voucher is prepared with all the check and check copies

and it's filed for one year.

A check register is prepared,

you have some distribution tickets,

which I'm not sure these are,

are kept for each account and posted at the same time as

voucher register.

Looks like for each account,

we have a distribution ticket.

A tape is made of each of these

and used to post to the general ledger.

So these do not directly update the general ledger versus

per se, they update a distribution ticket which is used later.

Looks like a manual update.

The MCR operator has some controls,

as far as it looks like they batch them, or

I can't tell if they batch them,

and the controller approves all cash disbursements and there is

only 2 authorized signers,

the president and the controller.

I'm now looking at the financial statements,

Appendix A

where I have comparative between December 31, 80

and December 31, 79.

so it's between 79 and 80.
I see a note relating to 1979 numbers, we had a prior period adjustment. The financial statements for 1980 have been retroactively restated to correct an overstatement of inventories that resulted from the inadvertent inclusion of approximately $45,000 of customer owned products in finished goods inventory. This prior period adjustment had the effect of reducing income before extraordinary item and net income for 1979 by about $19,000, or 28 cents per share. So the inventory contains the products that have been shown as sold, but have not been shipped out, as what it looks like, or it's inventory that we're doing custom work on, we don't own the inventory, we're just processing on it. I'd be interested to see when their physical inventory was taken. In going down the balance sheet starting with assets, cash, it looks like their cash position has declined a little bit. Not too much, they've got the CD outstanding still. Accounts receivables has increased, notes receivable about the same, refundable income taxes for 1980, sure to see their income statement. They have a book income, but apparently, they must have a tax loss income. Inventories, approximately the same, there's a note on it. Lower cost or market, they're on a LIFO method beginning in 1980, so we have a change in accounting on this, we'll definitely have to test and make sure they're properly converted to LIFO, and that they're accounting for it properly, and that they got the right records. At December 31, costs are determined by the FIFO method, for approximately 81% of the inventories, and by the LIFO method for the remainder. So we got a lot of potential problems, here, in the pricing of this inventory. Again, I'm interested in seeing when their physical inventory is taken.
124 GOT PROPERTY, PLANT AND EQUIPMENT,
125 NO CHANGE AT ALL IN LAND.
126 BUILDING, HARDLY ANY,
127 MACHINERY AND EQUIPMENT, FURNITURE AND FIXTURES, MOTOR
VEHICLES,
128 NO PROBLEM THERE.
129 DEFERRED CHARGES;
130 WE HAVE DEFERRED FIT,
131 NOTE 6;
132 PRETTY STANDARD WORDING,
133 THEY HAVE TISING DIFFERENCES;
134 SCHEDULE IN DIFFERENCES ARE THE WAY THEY HANDLE
ADVERTISING
135 EXPENSES.
136 PENSION PLAN EXPENSES, RETIREMENT CONTRACTS, AND FRANCHISE
137 TAX FOR OHIO.
138 THE DIFFERENCES ARE IN DEFERRED TAX ACCOUNTS IN THE BALANCE SHEET.
139 BECAUSE OF INVESTMENT TAX CREDITS,
140 THEY REDUCED APPROXIMATELY 16,000 FOR THIS YEAR.
141 THEY HAD A NOL CARRYFORWARD FOR FRANCHISE TAXES,
142 AND HAS BEEN SHOWN AS AN EXTRAORDINARY ITEM IN THE STATEMENT OF INCOME.
143 WHICH I SEE.
144 DEPOSITS NO PROBLEM,
145 NOTES RECEIVABLE.
146 EXCESS OF COST OVER VALUE OF NET ASSETS OF ACQUIRED COMPANY,
147 LESS AMORTIZATION.
148 LOOKS LIKE WE HAVE SOME GOODWILL HERE.
149 AGAIN, I'D BE INTERESTED IN FINDING OUT THE HISTORY OF THIS.
150 LOOKS LIKE THEY'RE AMORTIZING IT OVER A PERIOD.
151 CASH SURRENDER VALUE OF LIFE INSURANCE;
152 I WOULD HAVE THOUGHT THAT THAT WOULD HAVE BEEN A PERMANENT
153 DIFFERENCE FOR TAX PURPOSES.
154 FLIPPING OVER TO LIABILITIES,
155 LOOKS LIKE WE HAVE A NEW NOTE PAYABLE.
156 ACCOUNTS PAYABLE ARE UP.
157 INCOME TAXES,
158 WE'RE NOT SHOWING A CURRENT LIABILITY FOR FIT
159 AND IT LOOKS LIKE WE ISSUED NEW LONG-TERM DEBT.
160 I'M SORRY, WE HAVE NOT,
161 THAT'S JUST THE CURRENT PORTION I'M LOOKING AT.
162 THERE'S A NOTE ON IT, WHICH I'M LOOKING AT NOW.
163 WE DO HAVE ONE NEW NOTE,
164 WHICH IS JUST FOR $150,000,
165 WE HAVE SOME OBLIGATIONS UNDER CAPITAL LEASES,
166 REFERS ME TO ANOTHER NOTE : 8
I see an analysis of the leased property under capital leases. Furniture and fixtures, the lease expired in 1980. We currently just have the machinery and equipment. Going down to stockholder's equity; I see we purchased some treasury shares, might want to test that and see if they've done the accounting for that.

That's not a normal recurring item. And because I can see that it's eaten into some of the additional paid in capital, in fact it has eliminated it.

Getting over to the income statement; sales are up slightly, probably due to inflation, more than volume. Interest income on their CD is up, that's to be anticipated because of the economic market. Cost of sales is up.

Selling and general administrative; again, looks like inflation, and interest is up for the same reason that interest revenue is up.

The interest rates have increased. We then have current income tax expense; $48,000 for this year, our income before income taxes and extraordinary item is down considerably. So we're showing net income of $130,000 down from $400,000.

So I would anticipate in this audit, that this is a new client, and I'd be interested in finding out why the other client, the other accounting firm has left, whether they were unhappy, was there conflicts over the accounting issues, I see there's been restatements here.

I think there's some potential problems here that should be ironed out early in the audit.

Net income, because it's down so low, I would think that you would find that the client would be hesitant and would argue certain accounting issues, especially of a subjective nature, as far as if they'd be making adjustments against income.

Retained earnings;
I CAN SEE THAT THE PRIOR PERIOD ADJUSTMENT IS THAT $19,000.

AS MENTIONED IN THE NOTE.

THEY'RE CONTINUING TO PAY THEIR DIVIDENDS.

THEY PURCHASED THAT TREASURY STOCK.

CHANGES IN FINANCIAL POSITION.

I'M LOOKING DOWN THE LIST NOW.

LOOKS LIKE IN 1979 WE HAD QUITE A BIT OF LONG-TERM DEBT,

I FAILED TO SEE THAT WHEN I ORIGINALY LOOKED AT THE FOOTNOTE.

I'M GETTING TO THE APPLICATION PORTION.

LOOKS LIKE WE HAD A SUBSTANTIAL AMOUNT OF PROPERTY, PLANT AND

EQUIPMENT ADDITIONS LAST YEAR.

THE ONLY OTHER ITEM WAS THE PURCHASE OF TREASURY STOCK WHICH I

HAD SEEN BEFORE.

THIS CONTINUES OVER, SHOWING THE COMPONENTS OF WORKING CAPITAL.

AGAIN, I SAW THAT CASH WAS DOWN.

CD WAS ABOUT THE SAME

ACCOUNTS RECEIVABLE HAS CONTINUED TO INCREASE.

THE REFUNDABLE INCOME TAXES AGAIN,

WE'D DEFINITELY BE LOOKING AT THIS AREA,

MAKING SURE THAT THEY'D PROPERLY ACCOUNTED FOR THIS.

INVENTORIES ARE DOWN.

IT'D BE INTERESTING, LOOKING AT THAT.

NOTES PAYABLE

NO OTHER UNUSUAL ITEMS THERE.

LOOKING AT THE FOOTNOTES,

I HAVE READ BEFORE,

THE INVENTORIES,

I'M NOW LOOKING AT DEPRECIATION -

NOTHING UNUSUAL THERE.

INVESTMENT TAX CREDITS,

AMORTIZATION, HERE,

IT TALKS ABOUT THE BUSINESS ACQUIRED IN 1964 -

STRAIGHT LINE METHOD FOR TEN YEARS.

OK.

PENSION PLAN,

I WOULD MAKE SURE THAT IF THEY REQUIRE AN AUDIT THAT WE GET IN ON

THAT,

AND MAKE SURE THAT WE'RE AWARE OF THE TIMING,

AS FAR AS GIVING BIDS,

AND THE TIMING OF THE AUDIT IF WE WOULD WIN THE AUDIT.

SO WE HAVE SOME RECLASSIFICATIONS,

ALONG WITH PRIOR PERIOD ADJUSTMENTS.

SO IF TYING IN LAST YEAR'S NUMBERS WILL BE A LITTLE MORE

DIFFICULT,

I WOULD BE SURE TO INCLUDE THAT IN TESTING THAT,

WHEN DEVELOPING MY AUDIT PROGRAM AND SCOPES.
NOW WE'RE TALKING ABOUT THE CHANGE IN INVENTORY VALUATION;
AGAIN, ABOUT SWITCHING TO LIFO,
WHICH PROBABLY HAS A LARGE IMPACT ON REDUCING THAT NET INCOME DOWN
AND REDUCING TAXABLE INCOME.
SO I WOULD THINK THAT IS A SIGNIFICANT PORTION,
WE CAN SEE COST OF SALES IS UP,
WELL OVER A MILLION,
WHEREAS SALES ONLY INCREASED $900,000.
SO I'D BE INTERESTED TO SEE WHAT PIECE OF THAT LIFO REDUCED INCOME
HERE I SEE THAT IT REDUCED ONLY BY $58,000, NET INCOME AFTER TAXES
I SEE WE DID NOT DO A PRO FORMA BACK,
I'D MAKE SURE THAT'S PROPER.
(LET ME MAKE A SUGGESTION, WHY DON'T YOU READ THIS ENTIRE DOCUMENT HERE, I THINK YOU MAY BE WORKING TOO HARD.)
FINE, I DIDN'T KNOW WHERE YOUR EMPHASIS WAS.
SO NOW I'M BACK TO THIS WHERE I'M NOW,
SUBJECT D CONTINUES READING DOCUMENT BEGINNING AND ENDING
AS FOLLOWS:
MY MANAGER HAS EXAMINED ALL THE INFORMATION AND HAS TOLD ME
TO CHOOSE A RANDOM SAMPLE OF CASH DISBURSEMENTS ....
SUBJECT D INJECTS THE FOLLOWING COMMENTS WHILE READING:
SO WE'RE MAKING A DECISION BASED UPON A GENERAL KNOWLEDGE OF
THE COMPANY,
A KIND OF A FEEL FOR IT,
A JUDGMENTAL.
A REVIEW OF SPECIFIC INTERNAL CONTROL PROCEDURES IN
THIS AREA,
THE RESULTS OF OUR WORK,
AND THE ECONOMICS OF OVER AUDITING OR UNDER-AUDITING.
THE OVER-AUDITING, OF COURSE WOULD BE OVERRUNNING THE JOB
INCURRING EXPENSES.
WHICH WE MAY NOT BE ABLE TO RECOUP.
UNDER-ESTIMATING COULD RESULT IN,
WE WOULD COME UP WITH THE WRONG AUDIT OPINION,
AND BE SUSCEPTIBLE TO FUTURE LITIGATION,
OR ADDITIONAL REWORK AFTER WE DETERMINE WE CAN'T RELY ON IT AND
HAVE TO GO BACK AND FIX IT.
NUMBER THREE,
THE TRUE STATE OF INTERNAL CONTROLS,
AGAIN, THIS IS WHAT THE MONITOR IS GOING TO BE SHOWING ME FROM
THE COMPUTER,
WHETHER IT'S A, B, C, OR D. OK.
I BELIEVE THIS IS AN ALPHA ERROR WHERE YOU
ACCEPT A FALSE
HYPOTHESIS.

SUBJECT D FINISHES READING:
HERE IS A SAMPLE OF 100 ERROR RATES GENERATED BY THE
COMPUTER.

I SEE SOME AS LOW AS 3 TENTHS OF A PERCENT,
AND AS HIGH AS 19%.
SO LOOKING AT WHAT WOULD FALL UNDER D,
THERE AIN'T THAT MANY THAT FALL IN THAT POPULATION D,
SO AS FAR AS LOSING THE $4000,
HOPEFULLY IT WON'T HAPPEN.
I DON'T THINK THAT COULD HAPPEN VERY EASILY.
OK, I'M DONE READING THIS PAGE.
I'M READY FOR THE NEXT STEP.
(WELL, YOU COULD CHOOSE A SAMPLE SIZE.)
(AS IT SAYS HERE, AS THE SUBJECT IN THE EXPERIMENT,
YOUR
ONLY TASK IS TO CHOOSE THE SAMPLE SIZE FOR THE
COMPLIANCE TEST).
SO IT'S THE FIRST TIME I'M GOING TO CHOOSE A SAMPLE
SIZE,
FOR RIGHT HERE,
AND AN ERROR RATE WOULD BE IF ANY OF THESE CONTROLS
THAT WE'VE
IDENTIFIED HERE, ARE NOT OPERATING FOR THE ONE THAT WE
SELECT.
(RIGHT)
IF FOR ONE ITEM, THERE'S 3 OR 4 EXCEPTIONS,
DOES THAT COUNT AS 4 EXCEPTIONS OR JUST ONE?
(JUST ONE)
AND THERE'S 7000...
I'M TRYING TO GET A LITTLE PERSPECTIVE HERE, THAT'S
7000 DIVIDED BY 12.
I SEE THERE'S ABOUT 600 A MONTH.
JUST TRYING TO GET A FEEL HERE NOW.
THIS IS A NEW CLIENT.
SO I WOULD EXPECT THIS YEAR, THAT I'D BE WILLING TO
INCREASE
MY SAMPLE SIZE,
JUST BECAUSE IT IS A FIRST YEAR CLIENT.
MOSTLY JUST TO DEVELOP A BETTER UNDERSTANDING OF THE
CLIENT'S
PROCEDURES, AND ALSO PROVIDE TRAINING FOR PEOPLE ON
THE JOB.
SO,
ALSO, NEXT YEAR THEN,
WE COULD PROBABLY REDUCE THE SAMPLE SIZE,
BECAUSE THIS LEARNING PROCESS WILL BE REDUCED.
BECAUSE WE'LL UNDERSTAND IT AND HAVE BETTER
DOCUMENTATION OF
THE SYSTEM.
AFTER WE'VE DONE THIS EXCERCISE ONCE.
BUT GETTING TO THE CASE ON HAND,
YOU CAN DO THAT?
I WOULD LIKE TO SEE THE FOLLOWING:
UPPER CONFIDENCE BOUND
WOULD BE 10% THE FIRST TIME,
AND I'D SEE MY LEVEL OF RISK,
I'D LIKE TO SEE IT AT 90%,
AND AT 95%.
I'D LIKE TO SEE WHAT TYPE OF SAMPLE SIZE.
(YOU'LL ALSO HAVE TO ESTIMATE WHAT YOU THINK THE
POPULATION ERROR
RATE IS)
LOOKING AT, GOING BACK TO THIS SAMPLE OVER HERE,
I DON'T SEE VERY MANY OVER TEN.
I DO SEE A LOT, LIKE 6, 7, OR 8.
SO WHY DON'T WE TRY 8%.
JUST TO SEE WHAT THIS GIVES US.
----
(HERE WE GO, SAMPLE SIZE FOR THE FIRST CASE IS 303).
(SECOND CASE, THIS IS WHAT THE SAMPLE SIZE IS,...)
IT'D BE EVEN HIGHER RIGHT?
(THAT'S CORRECT, LET'S SEE, ... IT'S 498).
WELL, BOTH OF THESE WOULD BE UNACCEPTABLE FOR ME,
THE ONLY TIMES I HAVE EVER USED MORE THAN 300 HAVE
BEEN ON CLIENTS THAT HAVE BEEN A LOT BIGGER THAN
THIS.
ONE EXAMPLE WOULD BE ONE THAT HAS 500,000 CUSTOMERS,
AND WE USED A SAMPLE SIZE OF ALMOST 700,
WE USED MONETARY UNIT SAMPLING.
SO IN THIS CASE,
HAVING ONLY 7000 TRANSACTIONS,
500, THAT'S 7% OF THE TRANSACTIONS,
I'D BE LOOKING AT THIS IF I DETERMINE THAT THEY HAD
GOOD INTERNAL
CONTROLS,
AND THROUGH OBSERVATION AND DISCUSSION WITH THEM,
DETERMINE THAT THESE ARE IN OPERATION.
WITHOUT REALLY ACTUALLY TESTING THEM.
IF I DETERMINED THAT THESE CONTROLS WERE GOOD,
AND ARE OPERATING,
I WOULD BE WILLING TO LOWER MY CONFIDENCE LEVEL DOWN
TO 85% TO SEE WHAT TYPE OF SAMPLE SIZE I WOULD GET.
(85%, OK)
I HAVE DONE THIS IN THE PAST ON CLIENTS THAT HAD
DEMONSTRATED
VERY GOOD INTERNAL CONTROLS,
AND THAT FOR A PERIOD OF THREE OR FIVE YEARS THAT
WE'VE
AUDITED THEM,
THAT THEY HAVE HAD VERY LOW ERROR RATES.
SINCE THIS IS A FIRST TIME THROUGH CLIENT,
I MAY BE RELUCTANT TO DO THIS.
OR I MAY MAKE A DECISION NOT TO USE STATISTICS.
RATHER JUST USE IT TO RANDOMLY SELECT.
(OK, WHAT'S YOUR ESTIMATE, STILL 8%?)
YEAH
(AND YOUR UPPER CONFIDENCE BOUND IS STILL 10%)
YEAH
(THE BRINGS IT DOWN 200)
(NO)
I WOULD MAKE MY FIRST,
LOOKING AT, BASED ON THESE STATISTICS,
I'M NOW GOING TO JUST COMPUTE HOW MUCH 200 TRANSACTIONS WOULD
BE WORTH.
NOW THIS ASSESSED 20 CENTS PER TRANSACTION,
IS THAT WHAT YOU'RE ASSUMING THE COST OF THE AUDIT?
(THAT'S RIGHT)
IS THAT A VALID NUMBER?
I'M TRYING TO...
LET'S SAY THE GUY'S BILLING RATES,
PLUS THE NEW STAFF MEMBER...
(WELL, I SHOULDN'T SAY THE COST OF THE AUDIT,
IT'S THE SYSTEM OF COSTS AND BENEFITS GIVEN IN THE PROBLEM)
OK.
BUT USING THIS ONE,
IF WE WOULD USE 200 TRANSACTIONS,
TIMES 20 CENTS,
THAT'S $40,
AND IF I'D PICK IT RIGHT ON THE NOSE,
I'D WIN A $100.
HOWEVER, IF I MISS IT BY ONE,
I GET NOTHING,
AND IF I MISS IT BY 2 I LOSE $50.
TO SPEED THINGS ALONG,
I'M GOING TO GO AHEAD AND PICK A SAMPLE SIZE OF 75.
...
(OK, WHAT THIS SAYS, IS THAT YOU CHOSE A SAMPLE SIZE OF 75,
YOUR RESULTS DETECTED 2 DEFECTIVE TRANSACTIONS, THE MANAGER'S DECISION WAS THAT THE STATE OF INTERNAL CONTROLS WAS A,
THE TRUE STATE OF INTERNAL CONTROLS WAS B. SO YOU LOST $15,
WHICH IS THE COST OF SAMPLING.)
SO IN THIS CASE, I HAVE UNDER-AUDITED,
I'VE DRAWN THE WRONG CONCLUSION.
AND MY LEVEL OF SUBSTANTIVE TESTING WAS NOT PROPER.
NOW, DO WE CONTINUE WITH ANOTHER ONE?
(YES, WE CAN GO ON FOR A LARGE NUMBER)
SAME CONDITIONS, RIGHT?
NOW WE'RE JUST KIND OF PLAYING AROUND PICKING NUMBERS HERE?
I'M NOW GOING TO LOOK BACK ON THESE INTERNAL CONTROLS.
I'M LOOKING NOW FOR CONTROLS,
TO FIND SIGNIFICANT WEAKNESSES WHERE PROBLEMS COULD OCCUR,
AND NOT GET DETECTED,
AND NOT CORRECTED,
SO I'M LOOKING FOR THE POSSIBILITY
OF AN ERROR OCCURRING,
AND THE POSSIBILITY OF THAT ERROR NOT BEING DETECTED.
AND THESE WOULD BE THE ERRORS THAT CAME OUT IN MY SAMPLE.
GOING BACK TO THE FIRST PAGE OF PURCHASES,
I SEE THERE'S A LOT OF COPIES OF THE PURCHASE ORDER.
The matching of the receiving and the freight bills,
AND THE INVOICES,
APPEAR GOOD, AS FAR AS THE MATCHING OF IT.
THE PURCHASING AGENT COMES BACK AND APPROVES IT.
THE ACTUAL PROCESSING ACCOUNTS PAYABLE,
THIS THING HERE WHERE THE PURCHASING AGENT APPROVES IT,
THE COST ACCOUNTANT STAMPS IT AS PAID.
THOUGH IT ACTUALLY HASN'T BEEN PAID UNTIL LATER ON.
WHEN IT'S ENTERED INTO THE SYSTEM.
THAT COULD BE A POTENTIAL PROBLEM.
MAYBE THE STAMP FOR PAYMENT SHOULD BE DONE WHEN THE VOUCHER IS PREPARED.
BASED UPON THIS REVIEW HERE,
WE FOUND SOME WEAKNESSES,
WE ARE NOW GOING TO MAKE OUR SELECTION OF ..
ONLY FINDING 2 ERRORS, THE FIRST YEAR,
THIS IS A DIFFERENT COMPANY NOW, ISN'T IT?
(THE NAME IS THE SAME COMPANY)
SO I DON'T HAVE THIS PRIOR HISTORY,
AND KNOWING THAT WE ONLY HAD 2 ERRORS LAST YEAR RIGHT?
(THE NAME IS THE SAME COMPANY)
SO IT'S A BRAND NEW CLIENT AGAIN.
(BRAND NEW CLIENT; IDENTICAL IN EVERY RESPECT)
SO NOT HAVING THIS PRIOR HISTORY OF KNOWING THAT THE CONTROLS...
(I JUST WANT TO MAKE SURE THAT YOU UNDERSTOOD EXACTLY WHAT I SAID, IDENTICAL IN EVERY RESPECT EXCEPT FOR THE NUMBER OF CASH DISBURSEMENTS IN THE POPULATION WHICH DO NOT COMPLY WITH INTERNAL CONTROLS.)
OK, SO WE STILL GOT 7000 OF THEM.
(YEAH, BUT THE NUMBER THAT ARE DEFECTIVE HAS CHANGED)
I WILL STILL PICK 75.
THE REASON THAT I PICKED 75 WAS THAT JUST BECAUSE OF THE LIMITED NUMBER OF TRANSACTIONS,
I think that 75 will give me a good representative sample...

...and now we lost another $15.

We're not doing too good here,

but at least I'm not making the mistake of

my...

I've been selecting A,

whereas, it's been B.

So in order for me to accurately pick B,

I would have to increase my population size.

And that I will do by picking 100.

Hopefully we made some money on this.

...

Well, we made some money on this one.

By increasing my sample size to 100,

we found seven errors,

which is a seven percent error rate for my sample size,

which equals,

the actual state which is between 3 and 7%.

So the actual state could have been anywhere from 3 to

7 percent,

and our sample outcome indicated 7%.

My payoff was $80.

which was a $100, less $20 for transaction sampled.

I'm going to reduce my sample size to 80.

Again, I feel that a sample size of 100

though in this case gave me a good pay-off,

I still feel that I would not pick a hundred

under these circumstances

...

Ok, we picked the 80.

My sample outcome was that we found 1

My decision was that it was a

population error rate is less that 3%

and the actual state was a,

less that 3%.

My payoff was $100, less 16 for my sample selection.

Again, I'd like to go back and try the 75,

my original selection.

...

Ok, by decreasing my sample size 5 here

we found a sample outcome of 2,

out of a sample size of 75.

we made the decision that it was a,

less than 3%.

the actual state was a.

payoff was $100, less $15 for sampling.

my cumulative payoff now is $219.

I'm curious, you know, how statistically valid my

sample size of

75 is.
570 SO WHAT I'D LIKE TO DO NOW IS GO BACK TO THAT CHART AND MOVE
571 MY ERROR RATE DOWN TO SOMEWHERE BETWEEN 3 AND 7%.
572 SO I WILL NOW MOVE MY POPULATION ERROR RATE THAT I EXPECT
573 TO BE 5%.
574 ---
575 AND MY UPPER CONFIDENCE TO BE, REMAIN AT 10%.
576 (AND WHAT LEVEL OF RISK?)
577 MY LEVEL OF RISK WILL BE 85%.
578 (OK)
579 WHAT I NORMALLY DO IS,
580 WHEN I HAD DECIDED ON A SAMPLE SIZE,
581 I WILL GO BACK AND MAKE SURE THAT IT IS
582 STATISTICALLY VALID.
583 AND WHAT I MEAN BY STATISTICALLY VALID IS THAT
584 IS THE POPULATION ERROR RATE AROUND 5%.
585 IS THAT REASONABLE?
586 WOULD I BE WILLING TO ACCEPT A 10%?
587 AND CAN I BE 85% CONFIDENT?
588 IN ORDER TO BE 85% CONFIDENT,
589 I WOULD HAVE TO EVALUATE THEIR CONTROLS AND
590 AFTER MAKING THAT EVALUATION,
591 DOCUMENT THAT THE 85% IS APPROPRIATE.
592 (OK, WITH THE UPPER CONFIDENCE BOUND AT 10%, LEVEL OF RISK,
593 85%, AND AN ESTIMATED OCCURRENCE RATE OF 5%, SAMPLE SIZE FROM
594 TABLE IS 21.)
595 21.
596 (YES)
597 21?
598 (21.)
599 21.
600 THAT'S JUST BY LOWERING MY POPULATION ERROR RATE DOWN FROM
601 8% TO 5%,
602 I'VE LOWERED IT FROM 200 TO 21?
603 (THAT'S CORRECT.)
604 THAT'S ABSOLUTELY AMAZING.
605 NOW I DON'T FEEL SO BAD ABOUT ONLY PICKING 75 OR 100.
606 BECAUSE I WOULD BE UNCOMFORTABLE PICKING ONLY 21.
607 I WILL NOW LOWER MY SAMPLE SIZE SELECTION FROM THE 75, TO 60.
608 ---
609 BY LOWERING IT TO 60 WE MADE THE DECISION OF A
610 THAT IT WAS LESS THAN 3%.
611 THE STATE WAS ACTUALLY 8, 3 TO 7%.
612 THE PAYOFF THIS TRIAL WAS A LOSS OF $12.
613 MY CUMULATIVE PAYOFF WAS $207.
614 WHAT I'M TRYING TO DO,
615 THE REASON I PICKED 60 IS THAT IT IS APPROXIMATELY 5 A MONTH.
OUT OF 5 A MONTH, AND THERE'S APPROXIMATELY 583 A MONTH
ON THE AVERAGE, THAT GIVES ME ALMOST 1%.
I WILL NOW INCREASE IT TO 72, MY SAMPLE SIZE SELECTION.
THAT GIVES ME OVER 1% OF THE ACTUAL TRANSACTIONS I WILL BE TESTING.
HOW MANY TIMES DO WE DO THIS?
(WELL, I'D RATHER NOT DISCLOSE THAT, YOU CAN PLACE A LIMIT
ON IT, IF YOU'RE REALLY TIRED, OR IF YOU THINK...)
NO, I JUST, I JUST WANTED TO KNOW, THIS IS LIKE (I DON'T WANT TO DISCLOSE IT BECAUSE I DON'T WANT TO ENCOURAGE ANY TYPE OF END OF GAME STRATEGY.)
YEAH, YEAH, I CAN SEE THAT...
...
OK, WELL IN THIS CASE MY SAMPLE SIZE WAS 72,
AND MY SAMPLE OUTCOME WAS 9.
JUST CURIOUS WHAT TYPE OF PERCENT THAT IS.
THAT'S 12.5%.
WHICH IS IN C.
REALLY IT'S UH, OH I'M SORRY I, NO, THAT'S NOT RIGHT.
YOU CAN'T DIVIDE LIKE THAT,
YOU HAVE TO PUT IT INTO THE FORMULA TO FIGURE OUT WHAT YOUR DECISION IS.
BUT MY POPULATION ERROR RATE WAS 12.5%.
WHICH PUT IT IN THE CATEGORY OF C.
THAT'D BE IN 7 TO 12%. THE ACTUAL STATE WAS BETWEEN 7 AND 12%,
SO I ACCEPTED A PROPER HYPOTHESIS.
MY PAYOFF WAS $85.60, WHICH IS THE NET OF $100, LESS MY TRANSACTION COSTS.
I AM COMFORTABLE AROUND THIS 72,
I WOULD LIKE TO TRY IT AGAIN.
...
OK, IN THIS CASE,
MY STAFF PERSON MUST HAVE BEEN SLEEPING,
BECAUSE HE FOUND NO ERRORS,
HE FOUND NO ERRORS IN A SAMPLE SIZE OF 72,
WE MADE THE DECISION OF A
THE ACTUAL STATE WAS A.
PAYOFF WAS $85.60.
(ASSUME YOUR STAFF PERSON IS COMPETENT.)
HA HA HA, HA.
I'M GETTING MORE AND MORE COMFORTABLE WITH THIS 72.
IF THIS WAS A RECURRING AUDIT,
I WOULD PROBABLY REDUCE IT THE NEXT YEAR DOWN TO PROBABLY 50.
ESPECIALLY AFTER FINDING NO ERRORS,
BUT AGAIN ASSUME THIS IS A BRAND NEW CLIENT,
WE ARE USING STATISTICS,
I WILL STAY WITH 72.
I KNOW ONE OF THESE TIMES I'M GONNA LOSE $4000.00
OK, 72.
THE LAST ONE AGAIN WE HAD A SAMPLE OUTCOME OF 2
ERRORS,
DECISION WAS A.
The outcome was A, the state was A.
AND AGAIN, I HAD A PAYOFF OF $85.60.

MY DECISION WAS B.
MY STATE WAS B.
PAYOFF AGAIN, $85.60.
I WOULD BE TEMPTED TO LOWER AGAIN, THE SAMPLE SIZE
SELECTION.
BUT, I WOULD RATHER OVER-AUDIT, THAN UNDER-AUDIT, AND
MAKE
THE WRONG INFERENCE.
SO WE'LL PICK 72 AGAIN.

IN THIS CASE,
WE FOUND 2 ERRORS.
I MADE THE DECISION OF A.
MY MANAGER DID.

THE ACTUAL STATE WAS B.
MY PAYOFF WAS NEGATIVE $14.40, THE COST OF
TRANSACTIONS.

I'M STILL COMFORTABLE WITH 72.

AGAIN, WE MADE THE DECISION OF A,
AND THE STATE WAS ACTUALLY B,
SO WE LOST THE UH, PAYOFF.
72 AGAIN.

HAD A LOT OF ERRORS.
WE HAD 11 ERRORS.
MY DECISION WAS C.
THE ACTUAL STATE WAS C.

THE PAYOFF THIS TIME WAS $85.60.
ONE THING I'D BE INTERESTED IN,
IS FIND OUT WHAT THE AVERAGE ERROR RATE OF THIS TABLE
HE'S GIVEN ME.

WOULD YOU KNOW THAT, HAVE THAT CALCULATED?
(I WOULDN'T HAVE THAT CALCULATED, NO.)
OK.

WELL, I'M NOT GONNA SIT HERE AND ADD THEM UP,
BUT, JUST SCANNING THROUGH THERE,
IT LOOKS LIKE THE ERROR RATE,
THE AVERAGE ERROR RATE NOTED WAS AROUND 5 OR 6%.
(I'LL GET THAT FOR NEXT TIME.)
(I THINK THAT'S ABOUT RIGHT.)
OK, AGAIN I'LL PICK 72.

WHETHER IT'S LUCK OR NOT I HAVE YET TO HAVE A 2 TIE
PENALTY ASSESSED ON ME.
MAYBE I SHOULD NOT HAVE SPOKEN.
OK, FOUND 2 ERRORS.
DECISION A.
STATE A.
PAYOFF $65.60.
I THINK I WILL CONTINUE WITH 72.

(ANOTHER INCOMPETENT STAFFER.)
HE FOUND NO ERRORS.
I GOTTA WAKE THESE GUYS UP.
BUT, THE DECISION WAS MADE PROPER, WAS A.
AND THE STATE WAS A.
LET'S CONTINUE WITH 72.

OH, WE GOT A 2 TIER PENALTY HERE.
SAMPLE OUTCOME WAS 4,
WHICH MADE A DECISION OF A.
THE ACTUAL STATE WAS C.
SO I LOST $50 ON THAT PLUS THE TRANSACTIONS.
WELL, THESE THINGS DO HAPPEN YOU KNOW?
WE'LL STAY AT 72.

IF I HAD THE TIME AND RESOURCES I WOULD BE ABLE TO CALCULATE
I WOULD BE ABLE TO CALCULATE.
I WOULD THINK GIVEN THIS, I COULD CALCULATE AN EXPECTED PAYOFF
FOR MY SAMPLE SIZE SELECTION OF 72.
BUT, I'M NOT GOING TO.
WHAT I WOULD DO IN THAT CASE ...
I WOULD CALCULATE,
GIVEN MY CONFIDENCE LEVELS, THAT I WOULD PICK,
THAT WOULD GIVE ME THE SAMPLE SIZE OF 72,
OK?
I WOULD CALCULATE...
I KNOW IT'S 72 TIMES 20 CENTS,
THAT'S GONNA BE MY TRANSACTION COSTS.
OK, MY ALPHA ERROR OF ACCEPTING A FALSE HYPOTHESIS,
I WOULD BE ABLE TO CALCULATE THAT FOR A 4 LEVEL ERROR,
CALCULATE WHAT % THAT IS,
WHAT MY ALPHA ERROR IS,
DO THAT FOR A 3 LEVEL AND 2 LEVEL ERROR,
AND ALSO FOR THE 1 LEVEL ERROR,
AND ALSO FOR GETTING IT RIGHT ON IT.
THEN I'D MULTIPLY THEM BY THE PENALTIES ON THE REWARDS.
I THINK IT CAN BE DONE.
AND IF I HAD ENOUGH TIME.
I COULD PROBABLY FIGURE OUT,
AGAIN THIS IS RANDOM,
BUT, PROBABLY FIGURE OUT A GOOD RANGE TO PICK.
BUT IN LIGHT OF EXPEDIENCY, I'LL CONTINUE TO PICK 72.
WE NORMALLY ON AN AUDIT, YOU KNOW,
DO NOT GO THROUGH THESE ECONOMICS OF CALCULATING IT OUT.

(BM HM)

BUT IF I WAS TAKING A STATISTICS COURSE, I MIGHT TRY SOMETHING LIKE THAT.

BUT, TO BE REALISTIC I DON'T THINK I'D EVER PICK MORE THAN 72.

IN THIS.

WELL, I HAD A BAD STATE HERE,
STATE OF D.
WHICH MEAN THEY HAVE A HELLOVA LOT OF ERRORS, RATE,
BETWEEN 12 AND 20%.
OUR DECISION WAS A C.
BECAUSE WE FOUND 11 ERRORS.
SO I LOST THE TRANSACTION COSTS, $14.40.
72 AGAIN.

WHAT IF THIS SAMPLE SIZE EXCEEDS MAXIMUM ALLOWED?
WELL, AGAIN WE MISSED IT BY ONE TIER.
WE SELECTED B.
AND THE ACTUAL STATE WAS C.
WE LOST THE TRANSACTIONS COSTS.
Uh, LET'S INCREASE IT TO 100 THIS TIME JUST FOR GRINS.

WELL WE HIT IT ON THE NOSE.
WE FOUND ONE ERROR,
OUR DECISION WAS A
THE STATE WAS A.
LOOKING BACK AT THIS,
I WOULD PROBABLY HAVE A, IF I WOULD ALWAYS PICK 500,
I WOULD, IT WOULD COST ME $100,
TO DO THE TRANSACTIONS,
BUT I WOULD PROBABLY BE VERY, VERY CONFIDENT
THAT I WOULD ALWAYS AT LEAST WIN THE $100.
FOR HITTING IT ON THE NOSE.
SO IF I WANT TO BE CONSERVATIVE IN THAT RESPECT,
I WOULD STICK WITH 100.
BY INCREASING MY SAMPLE SIZE HERE,
IN THIS PROBLEM,
I DON'T WANT TO OVER-AUDIT,
BECAUSE OF THE 20 CENT PENALTY PER TRANSACTION,
BUT YET,
THE $100 REWARD OUTWEIGHTS ANY OF THE 20 CENTS TRANSACTIONS.
SO LET'S TRY 100.

WE'RE HAVING SOME BAD LUCK WITH THE 72.
WELL,
WE FOUND 20 ERRORS,
WE DECIDED D,
The ACTUAL STATE WAS D.
MY PAYOFF WAS $80.
LETS CONTINUE WITH 100.
EVEN 100 IS NOT INHALLIBLE.
WE HAD AN OUTCOME OF 5.
DECISION A.
STATE WAS B.
I LOST $20.
BOY, WHEN I WAS ONLY PICKING 72,
I ONLY LOST $14.40.
LET'S PICK 100 AGAIN.

... (HOW ARE YOU DOING ON TIME?)
FINE, I GOT NOTHING TO DO TODAY.
IS THERE A LIMIT ON THIS?
(NO.)
(OH, ON THE NUMBER OF TRANSACTIONS?)
(YEAH)
(YES, THERE IS A LIMIT)
I WOULD THINK YOU'D HAVE TO MAKE IT EQUAL FOR EVERYBODY.
WELL,
DECISION A.
STATE A.
LET'S HUNDRED IT.

... MY GOAL'S TO BREAK 1000 HERE.
NO WE DIDN'T.
WE MADE THE WRONG DECISION.
WE MADE A, THE STATE WAS B.
LET'S TRY 100 AGAIN.

... AT THIS RATE I CAN MAKE FIVE ERRORS LIKE THIS,
AND UH, STILL BREAK EVEN.
OK,
AGAIN WE MADE THE WRONG DECISION,
WE MADE A.
THE STATE WAS ACTUALLY B.
LET'S TRY 100 AGAIN.

... WE FINALLY BROKE A THOUSAND HERE.
SAMPLE SIZE OF 12 OUTCOME, DECISION C, STATE C.
100 AGAIN, PLEASE.

... SO FAR NONE OF MY SAMPLE SIZES HAVE RANGED BETWEEN,
I HAVE GONE ALL THE WAY DOWN TO 60,
AND HAVE GONE ALL THE WAY UP TO 100.
WHEN I WAS AT 60,
I DID, WELL, I DID OK.
WHAT HAVE WE GOT HERE,?
WELL THIS IS INTERESTING HERE.
WE HAVEN'T SEEN TOO MANY OF THESE,
WHERE MY DECISION WAS A C,
AND THE ACTUAL STATE WAS A B.
IN OTHER WORDS,
MY SAMPLE SHOWED A WORSE CONDITION THAN THE ACTUAL POPULATION.
A HUNDRED.

DECISION A, STATE OF B.
JUST SO WE DON'T GET BORING HERE,

LET'S TRY 120.

IT <THE COMPUTER> DIDN'T LIKE IT, HUH?
(IT'S BURPING.)

HA HA HA.
HUH, WE STILL DIDN'T HIT IT.
AND IT COST US $24.00 JUST TO FIND THAT OUT.
I THINK WE'LL GO BACK DOWN TO 72.

WE'RE BACK ON THE RIGHT TRACK AGAIN.
72 AGAIN, DECISION A, STATE A.

STATE B,
DECISION B.
72 AGAIN.

BY LOOKING AT THIS IT'S INTERESTING THAT
IGNORING THE PAYOFFS,
THE BENEFITS OF MAKING A HIT RIGHT ON THE NOSE,
OR 1 TIEB, 2 TIEB, OR 3 TIEB, OR 4 TIEBS OFF,
IT'S INTERESTING TO KNOW THAT I'VE ONLY BEEN OFF BY
MORE THAN ONE TIEB ONCE.
AND THIS IS GIVING ME CONFIDENCE THAT POPULATION SAMPLE
SIZE BETWEEN 50 AND 100,
WILL SOMETIMES GIVE YOU PRETTY CLOSE RESULTS, AND UH,
THEY WON'T BE TOO MISLEADING.

SAMPLE SIZE, UH,
WE GOT DECISION A, STATE E.
72 AGAIN PLEASE.

WELL, AT LEAST WE DIDN'T MAKE THE DECISION A OR B.
BECause THE ACTUAL STATE WAS D.
72 AGAIN.

(I'M SURPRISED)
WAS IT REALLY GOOD, OR REALLY BAD?
(NO, A MARK THAT'S A NORMAL OUTPUT, I WAS THINKING TO
MYSELF
ABOUT SOMETHIN', I TRY TO ANTICIPATE THE OUTPUT AS IT COMES,
BEFORE IT COMES OUT.)
OH, WE HIT ON THE NOSE.
LET'S TRY 72 AGAIN.

HIT IT.
BY KEEPING MY SAMPLE SIZE AT 72,
I'M STILL GETTING CLOSE RESULTS,
VERSUS DOING IT AT 120,
WHERE I POSSIBLY GOT THE SAME.
PLAY IT AGAIN SAM.
---
OK.
(ONE MORE)
YEAH.
---
OF COURSE, IF I GET ONE OF THESE A-D TRUE STATE D,
I COULD LOSE ALL MY MONEY HERE.
OH, WOW, WE HIT THAT ONE.
IT WAS A D,
AND WE DECIDED D.
AGAIN.
---
SUBJECT E

1. You are the senior in charge of the audit of Rohr Industries, a manufacturer of cabinets for electronic equipment.

2. Subject E is reading case comparative financial statements and notes.

3. Well this sheet is a good place to start.

4. You want me to just stop whatever I'm looking at to whatever I'm thinking as I look at these.

5. (Sure)

6. I suppose the first thing I looked at is the total assets.

7. Just to get an idea of the size of the company.


10. Looking at significant items that make up the $8,000,000 accounts receivable, inventories, property and equipment.

11. Seems to have a lot of depreciation on it for what that's worth.

12. Deferred income taxes, and some cash surrender value, look to be some of the more significant types of audit areas to me on the balance sheet.


15. I don't see any significant fluctuations right off the bat.

16. I see a note 2 above 1979's, which raises my curiosity.

17. Prior period adjustment has been retroactively restated to correct for an overstatement of inventories that resulted in the inadvertent inclusion of approximately $45,000 of customer owned products in finished goods inventories at December 31.

18. This prior period adjustment had the effect of reducing income before extraordinary items in net income in 1979 by $19,000.

19. Sure seems like a moot point to me right off the bat.

20. $45,000, I wouldn't think would be a material amount to an $8,000,000 balance sheet.

21. For restatement anyway.

22. There's an error.
RETROACTIVE RESTATED CORRECTION OVERSTATEMENT IN INVENTORY
IS A RESULT OF FROM AN INADVERTENT INCLUSION, AN ERROR.
OK.
NET INCOME, IN 1979 WAS $406,000,
WITH A $45,000 RESTATEMENT.
OK.
I'LL GO ON AFTER JUST TAKING A QUICK LOOK AT THE INCOME STATEMENT HERE.
$15,000,000 IN SALES.
I'M JUST TRYING TO VISUALIZE MY - HOW I CAN RELATE THIS TO COMPANIES I'M FAMILIAR WITH, WITH ABOUT THE SAME NUMBER OF SALES.
COST OF SALES.
I THINK GENERALLY, IN A MANUFACTURING COMPANY,
I'M SURPRISED WE DON'T SEE A GROSS MARGIN AMOUNT ANYWAY, HERE, WE SEE SALES, COST OF SALES, A NET FIGURE.
BUT IT'S KIND OF A CONDENSED INCOME STATEMENT,
WHICH MAKES IT EASY TO READ.
LOSS CARRY FORWARD.
YOU'RE GETTING TOO TECHNICAL ON ME HERE.
UTILIZATION OF LOST..., OK.
OK., HAVING DONE THAT,
AND YOU'RE,
FIRST YEAR ENGAGEMENT,
PREVIOUSLY AUDITED.
WHAT IS THE TASK?
(THAT'LL BE IN THE CASE)
THAT'LL BE IN THE CASE AS I READ ALONG, OK.
YOU'VE JUST COMPLETED A REVIEW OF THE INTERNAL CONTROL PROCEDURES OVER PURCHASES AND PAYABLES...<READING>...
...RESULTING FROM ANY ERROR IN JUDGMENT YOU MIGHT MAKE BY OVERSTATING OR UNDERESTIMATING THE TRUE AMOUNT OF THE INTERNAL CONTROL PROCEDURES TO BE RELIED UPON.
OK.
AND AS YOU SAY BACK HERE,
MY TASK IS TO CHOOSE AN ADEQUATE SAMPLE SIZE,
BASED UPON THIS, BASED UPON THIS,
INFORMATION THAT WE KNOW,
ACCOUNTS PAYABLE PURCHASES,
OH MY,
THIS ISN'T THE KIND OF THING WE CAN SIT DOWN AND LOOK AT IN FIVE MINUTES.
WELL, AS IS INFERRED IN THE PROBLEM
I THINK THAT PAGE 2 OF THE PROBLEM HERE,
YOU SAY THAT THE MANAGER'S PERTAINING TO HIS OVERALL
ASSESSMENT OF THE DEGREE TO WHICH THE FIRM CAN RLY ON THE
EXISTING INTERNAL CONTROLS FOR SUBSTANTIVE TESTING,
OH WE'RE NOT DOING COMPLIANCE TESTING.
(WE ARE DOING COMPLIANCE TESTING.)
OH,
CONTROLS FOR SUBSTANTIVE TESTING,
IS BASED ON HIS GENERAL KNOWLEDGE OF THE COMPANY AND ITS
ENVIRONMENT.
THE NARRATIVE AND FLOWCHART OF INTERNAL CONTROL PROCEDURES,
AND THE RESULTS OF YOUR SAMPLE.
OK, WELL, LET ME JUST READ ON HERE,
THE TRUE STATE OF INTERNAL CONTROLS,
A B C OR D ABOVE BECOMES KNOWN SOMETIME AFTER COMPLETION
OF THE AUDIT.
YOUR NET PAYOFF.
THE FOLLOWING OF REWARDS AND PENALTIES WILL ACCRUE TO YOU,
IN ALL CASES, YOU WILL BE ASSESSED 20 CENTS PER TRANSACTION
YOU WILL WIN $100 IF THE MANAGER'S DECISION CORRESPONDS WITH THE TRUE STATE OF INTERNAL CONTROLS.
YOU WILL loose $4000 IF THE MANAGER INFERS A OR B AND THE TRUE STATE OF CONTROLS IS D.
OK.
IF THE MANAGER ERRS BY ONE LEVEL IN EITHER DIRECTION,
NO REWARD OR PENALTY WILL ACCRUE TO YOU.
A 2 LEVEL ERROR,
OTHER THAN THE ERROR DISCUSSED PREVIOUSLY,
WILL CAUSE YOUR WINNINGS TO BE DECREASED BY 50 DOLLARS,
A 3 LEVEL ERROR,
OTHER THAN THE ERROR DISCUSSED PREVIOUSLY,
WILL CAUSE YOUR WINNINGS TO BE DECREASED BY $100.
OH, MY.
I COULD GO BROKE IN A HURRY.
NATURALLY, THE LARGER SAMPLE SIZE YOU CHOOSE,
THE GREATER THE LIKELIHOOD YOUR MANAGER WILL HAVE OF MAKING THE CORRECT INFERANCE.
YOU WILL NEED TO BALANCE THIS INTUITIVE STATISTICAL PROPERTY AGAINST THE COST OF SAMPLING.
YOU'LL HAVE THE OPPORTUNITY TO LEARN FROM EXPERIENCE,
BECAUSE THE EXPERIMENT WILL BE REPEATED A LARGE NUMBER OF TIMES.
EACH NEW COMPANY IS IDENTICAL IN EVERY RESPECT,
EXCEPT...
FOR EACH REPEAT OF THE EXPERIMENT,
ASSUME YOU ARE IN CHARGE OF AN AUDIT OF ANOTHER COMPANY,
I SEE, WE'RE GOING TO GO THROUGH THIS SEVERAL TIMES.
(RIGHT)
DIFFERENT COMPANIES, I SEE.
BUT EACH NEW COMPANY IS IDENTICAL IN EACH RESPECT TO ROHR, EXCEPT THE NUMBER OF CASH DISBURSEMENTS IN THE POPULATION, 7000, WHICH DO NOT COMPLY WITH THE STATE OF INTERNAL CONTROL PROCEDURES.
I DON'T UNDERSTAND WHAT WE'RE SAYING HERE.
(OK, FOR EACH TRIAL, FOR EACH REPEAT OF THE EXPERIMENT, WE'RE GOING TO ASSUME THE EXACT SAME COMPANY, A CLIENT EXACTLY LIKE ROHR, EXCEPT THE TRUE STATE OF THE WORLD IS DIFFERENT. THE POPULATION ERROR RATE, THE CHANGE.)
I SEE.
OK, SO ONE COMPANY MAY BE AN A,
ANOTHER A D, ONE A C AND ONE A B.
(RIGHT).
FINE, OK.
THE NUMBER IS DETERMINED AT RANDOM BY COMPUTER,
AND THE MANNER IS INTENDED TO REFLECT FREQUENCY OF ERROR RATES OBSERVED OVER ALL CLIENTS AUDITED BY BIG 8 CPA FIRMS WITH SIMILAR CHARACTERISTICS.
WITH YOUR EXPERIENCE,
YOU PROBABLY HAVE A PRETTY FAIR CONCEPTUALIZATION OF ERROR RATE FREQUENCY,
BUT TO MAKE THE COMPUTER GENERATION PROCESS MORE EXPLICIT TO YOU, HERE ARE 100 ERROR RATES,
A SAMPLE OF 100 ERROR RATES IN %S GENERATED BY THE COMPUTER.
OK.
SO, THERE ARE VARIOUS ERROR RATES BETWEEN 14%, 19%,
SO THAT PRETTY MUCH COVERS YOUR GAMBIT OVER HERE ON A THROUGH D.
OK, SO THAT'S OUR SAMPLE OF COMPANY ERROR RATES.
THIS IS OUR INTERNAL CONTROLS.
FLOWCHART SYSTEM.
NOW,
YOU WOULD LIKE FOR ME TO GIVE YOU,
BASED UPON ALL OF THIS,
A SAMPLE SIZE FOR ROHR INDUSTRIES,
A TEST OF CASH DISBURSEMENTS, AND PURCHASING SYSTEM.
(RIGHT)
JUST AS I WOULD IN REAL LIFE HERE, HUH?
(RIGHT)
Well, I started off looking here at your order entry system,
Uh, number 1 where point of entry would happen,
in the engineering department.
Seems like a logical place to start.
And, focused in on the 40% of purchases special items,
request for purchases comes from engineering.
And is apparently, let's see what happens,
purchase approval limitations,
$1,000, the president, $5K, the board of directors,
over $10K, for production and material,
the president, all other po's, the purchasing agent.
I would say that it's good to know that they do have uh,
a system of purchase approvals,
which hopefully they would be following,
and that would be one item of course,
that we'd want to test for,
to see if in fact, all purchase orders are being approved.
As called for in their policy down here.
Are we saying here, tcm, that these bills of material,
at least some of these bills of material, are sent over
here to the purchasing department,
from which a po is prepared,
I'm a little confused as to where we actually do, either
authorize the purchase or get the information from engineering
to the purchasing department to make the purchase.
(um, well, uh, these are the purchase orders right here)
Ok, ok, so purchasing would,
I assume,
prepare a purchase order,
8 part purchase order,
based upon information received from engineering?
That's right
in the form of bill of materials.
(right)
Stock items, special items,
any item related to a contract job up to 40% of purchases,
Ok, let's see.
pa approval
(purchasing agent)
Purchasing agent approval needed for stock purchases and
supplies is a department expense,
it must have department approval.
Ok.
Manual operation,
cost accountant determines cost distribution.
AND DETERMINES ANY VARIANCE.
PO2, IT GOES ON FILE,
AND THEN GOES DESTROYED.
HA HA HA, YOU'RE KIDDING.
THEY WOULD, IT WOULD APPEAR FROM THIS FLOWCHART THAT
THEN THEY'RE DESTROYING ANY APPROVED PURCHASE ORDERS.
AS IT RELATES TO --
COST DISTRIBUTION DETERMINES ANY VARIANCE, OK.
PA APPROVAL.
PURCHASING AGENT APPROVAL NEEDED FOR STOCK PURCHASES
AND
SUPPLIES AT DEPARTMENT EXPENSE,
IT MUST HAVE DEPARTMENT APPROVAL.
OK, OK.
RECEIVE PURCHASE ORDERS OVER TO THE RECEIVING DEPT.
RECEIVE FREIGHT BILL AND RECEIVING REPORT,
FILE ON LINE,
COST DISTRIBUTION IN DETERMINING A VARIANCE,
I THINK I,
JUST TO EXPAND HERE,
I WOULDN'T RECOMMEND THAT THEY WOULD SEND THE PURCHASE
ORDERS
OVER TO THE RECEIVING DEPARTMENT.
IT WOULD NOT BE A GOOD INTERNAL CONTROL.
ALTHOUGH IT'S NOT UNUSUAL,
AND IT'S, WOULD BE BETTER TO KEEP THOSE OUT OF THE
RECEIVING DEPT.
FREIGHT BILL, RECEIVING REPORT,
RECEIVING REPORT FILED BY VENDOR
IN THE RECEIVING DEPARTMENT,
COMES BACK OVER HERE TO THE PURCHASING DEPARTMENT
WHERE IT IS MESHED WITH SOME MORE PURCHASE ORDERS
AND THAT'S QUANTITY.
FREIGHT SHIPMENT.
5, 6, OR 7 DEPENDING ON WHETHER THE SHIPMENT WAS
PARTIAL OR
COMPLETE.
OK, USUAL FREIGHT BILL, EVERY MONTH, THAT'S GOOD.
FROM THE VENDOR.
GOES TO THE PURCHASE AGENT WHO REVIEWS IT.
RIGHT NOW, I'M LOOKING FOR YOUR P2, PAGE CONNECTOR.
P2 OFF PAGE CONNECTOR, P2.
OH, P2, OR IS THAT P1?
(WELL, YEAH, GO TO P2 AND THAT'S FROM P1, THAT'S THE
OFF PAGE
CONNECTOR)
OH, I SEE, TO PAGE 2 FROM PAGE 1.
I SEE.
I SEE.
I'M JUST SITTING HERE TRYING TO THINK OF THE
ADVISABILITY OF
SENDING THE INVOICE BACK TO THE PURCHASING AGENT
AND OH, THEN HAVING IT GO THROUGH APPROVAL,
AND I WOULD PROBABLY DISCOURAGE AGAINST THAT KIND OF
PRACTICE.

In so far then the uh, purchasing agent would of course, be able to, to divert any sort of invoice that should happen to come in for an order which he may have purchased.

From a company he never knew about.

Upon approval for payment by the purchasing agent, the invoice and the po are forwarded to the cost accountant who stamps the invoice with a payment stamp, enters the inventory count coding, then gives the invoice to the accounting clerk for payment.

The accounting clerk takes any discount available, runs a tape to determine the accuracy of invoice summary, and has payments approved by controller.

Ok, so I think there we've got a couple of good controls from the standpoint that uh, that he has double checked the accuracy of the invoice summary, and that we have approval of the invoice from the controller.

Approves all the cash disbursements, authorized signers, president and controller.

Now we come over, this is what department are we in here.

Are we in the accounting dept. now?

I see up here we were engineering dept., purchasing dept., receiving, accounting? With the controller?

(Yeah)

Cost accountant, ok.

NCR operator.

They have an NCR computer.

(Well it's a posting machine)

Posting machine, ok, fine.

Not simultaneous, let me see, ok.

NCR operator.

Invoice flowing through the system, with the receiving report.

Where'd we pick up the receiving report?

Let's just see.

Receiving report, receiving report, freight bill, receiving report is just one of the copies of the purchase order.

Oh, I see.

So, one of these purchase orders is the receiving report.

(Right)
AND FOR THE APPROVAL OF THE INDIVIDUAL THAT LOOKED AT THE RECEIVING REPORT.

AS I SAID BEFORE, I THINK THAT UH, IT'S PROBABLY NOT A GOOD IDEA TO USE PURCHASE ORDER AS A RECEIVING REPORT.

UH, BUT, ON WE GO, BY DATE, THIS LOOKS LIKE SOME OF THE FLOWCHARTS I'VE MADE.

INVOICE, PO, CHECK COPY, UH OH, WE'VE PREPARED A CHECK SOMEWHERE, OH I SEE, COME DOWN HERE, AND WE PREPARE A VOUCHER FOR THE CHECK, SIMULTANEOUS POSTING TO THE VOUCHER AND DISTRIBUTION REGISTER.

OK, CAN RELATE TO THAT, FILED BY YEAR OVER THERE, DISTRIBUTION TICKETS.

UH, I THINK THAT'S UNUSUAL. THAT UH, LET'S SEE, I WAS JUST THINKING THAT THE CONTROLLER IS ACTUALLY UH.

REVIEWING AND APPROVING FOR PAYMENT OF THE INVOICE, OVER HERE, AND THEN A CHECK IS PREPARED AND THEN GOES BACK TO HIM FOR SIGNATURE.

I GUESS, I DON'T KNOW, SEEMS LIKE A DUPLICATE OF EFFORT, BUT UH, I DON'T SEE ANYTHING WRONG WITH HIM APPROVING IT BEFORE A CHECK IS PREPARED.

AND AFTER THE CHECK IS PREPARED, SIGNING IT IS AN APPROVAL ALSO, SC. MAY BE A DUPLICATE OF EFFORT, BUT, NO PROBLEM, UH, SIMULTANEOUS POSTING, DISTRIBUTION TICKETS, COMPARED TO THE CHECK REGISTER, WHICH IS GOOD.

DISTRIBUTION TICKETS ARE KEPT FOR EACH ACCOUNT AND POSTED AT THE SAME TIME AS THE VOUCHER REGISTER.

THE TAPE IS MADE AT THE END OF THE MONTH FOR EACH ACCOUNT WHICH IS USED FOR POSTING TO THE GENERAL LEDGER.

A CHECK FOR EACH ACCOUNT IS POSTED, UH SO IS A DISTRIBUTION TICKET, IF WE HAVE SOMEBODY OVER HERE, VOUCHERING ALL OF THESE INVOICES, WOULD A DISTRIBUTION TICKET JUST BE BASICALLY A SUMMARY OF THE ITEMS CHARGED TO A PARTICULAR ACCOUNT,
AND WE'D HAVE A DIFFERENT SHEET OR TICKET FOR EACH
ACCOUNT,

SHOWING EACH OF THE DOLLAR AMOUNTS CHANGED TO THAT
ACCOUNT?

(THAT'S RIGHT)

OK, DISTRIBUTION TICKETS ARE FOR EACH ACCOUNT.

AND POSTED AT SAME TIME AS VOUCHER REGISTER.

OK, PROBABLY PUT IT ALL IN THE MACHINE AT ONCE.

THE TAPE IS MADE AT THE END OF THE MONTH,

FOR EACH ACCOUNT,

WHICH IS USED IN POSTING TO THE GENERAL LEDGER.

SO THEY POST TO THE GENERAL LEDGER FROM THE
DISTRIBUTION TICKET.

AND THE TOTAL OF THE DISTRIBUTION TICKETS,

COMPALED TO THE CHECK REGISTER,

CHECK AND CHECK COPY COMES BACK OVER HERE,

THE ACCOUNTING CLERK RECEIVES THE CHECKS,

ATTACHES THE REMITTANCE ADVICES,

AND HAS THEM SIGNED, SEE BELOW,

BEFORE SENDING THEM OUT.

THE NCR OPERATOR POSTS THE GROSS AMOUNTS AND THE CHECK
NUMBERS INTO THE CHECK REGISTER.

I THINK THAT UH, THAT THERE'S SOME,

IS THAT RIGHT?

(YEAH)

CHECK COPY SIGNED BY...

(SEND IT TO THE VENDOR)

AND TO THE VENDOR.

SO IT'S SIGNED BY EITHER THE PRESIDENT OR THE
CONTROLLER,

BUT I WOULD SAY, IN LOOKING AT THIS,

THAT THE PRESIDENT OR THE CONTROLLER IS SIMPLY SIGNING
A CHECK,

AND NOT UH, NOT AT THE TIME HE SIGNS THE CHECK,

LOOKING AT THE SUPPORTING DOCUMENTATION.

SO I DON'T KNOW WHAT THE HELL GOOD IT DOES TO REALLY
APPROVE

IT UP HERE,

AND THEN SIGN A CHECK DOWN HERE,

ESSENTIALLY, YOU'RE SENDING OUT DISBURSEMENTS FOR
WHICH MAY,

OR MAY NOT BE AUTHORIZED.

OK.

THE GENERAL LEDGER GETS POSTED FROM THE CHECK
REGISTER.

I THINK TOO,

ONE OF THE THINGS I'M THINKING NOW, TOM, IS THAT UH,

IN ADDITION TO LOOKING AT A SYSTEM FROM THE POINT OF
ENTRY

NUMBER ONE HERE,

WHICH IS OBVIOUSLY THE UH,

ONE WAY TO LOOK AT IT,

THAT TOO, UH, I LIKE TO FOLLOW IT ALL THE WAY THROUGH,
THEN I WOULD LIKE TO COME BACK HERE AND LOOK AT THE GENERAL LEDGER AND SEE WHERE WE GOT TO FROM THE GENERAL LEDGER, BECAUSE OF COURSE, THAT WOULD BE THE BOOK OF ACCOUNTS FROM WHICH THE FINANCIAL STATEMENTS CAME FROM.

AND I SEE IT BEING POSTED FROM THE CHECK REGISTER, AND THE DISTRIBUTION TICKETS. UH, GENERAL LEDGER POSTED FROM THE CHECK COPY, FROM THE CHECK REGISTER WHICH IS...

I SEE ANOTHER PROBLEM IN HERE, IN SO FAR AS YOU UH, HAVE THE NCR, LET'S SEE IF I UNDERSTAND THIS; THE ACCOUNTING CLERK RECEIVES THE CHECKS, ATTACHES THE REMITTANCE ADVICES, AND HAS THEM SIGNED BEFORE SENDING THEM OUT. THE NCR OPERATOR POSTS THE GROSS AMOUNTS, AND CHECK NUMBERS TO THE CHECK REGISTER. SO YOU'VE GOT THE NCR OPERATOR DOING THE ENTIRE SHOT OVER HERE,

IN POSTING THESE DOCUMENTS TO THE SYSTEM AND PERHAPS A BETTER WAY MIGHT BE TO UH, SEGREGATE THOSE RESPONSIBILITIES. UH, OK NOW, I'M CALLED UPON TO SELECT A SAMPLE. (RIGHT)

TO TEST IT.

DO I HAVE TO SAY WHAT TYPES OF ATTRIBUTES I'M GOING TO BE TESTING FOR OR, JUST WHAT THE SAMPLE WILL BE?

(WELL, I WOULD SAY, AFTER ALL OF THIS) THAT UH, MY EVALUATION OF THE CONTROLS IN THIS AREA, ARE SUCH THAT UH, THERE ARE SOME CONTROLS IN PLACE AND I THINK THERE ARE SOME SERIOUS WEAKNESSES AND IN ADDITION, SOME WEAKNESSES THAT AREN'T SO SERIOUS THAT WOULD CERTAINLY NEED TO BE IMPROVED, OR SHOULD BE IMPROVED.

TO STRENGTHEN CONTROLS.

NOW, BEFORE I JUMP TO ANY RASH DECISIONS, I WANT TO GO BACK HERE AND TAKE A CLOSER LOOK AT THE FINANCIAL STATEMENTS.

AND OF COURSE, AS I SAID BEFORE, INVENTORIES ARE SIGNIFICANT ASSETS OF THE COMPANY AND PROBABLY ONE WHICH IS VERY SUSCEPTIBLE TO ERROR.

MORE SO THAN CASH, WHICH OF COURSE, IS NOT A SIGNIFICANT ASSET.

OR THESE CERTIFICATES OF DEPOSIT, OR EVEN RECEIVABLES.

GENERALLY INVENTORIES ARE MORE SUSCEPTIBLE TO ERROR.
READ FOOTNOTE 3 ON INVENTORIES.

And change in method of inventory valuation, I see.

Effective 1980, the company adopted the LIFO method of determining the cost for all inventories for which costs were previously determined.

Under the first in first out method, change to reflect earnings more realistically.

The effect of this change was to reduce inventories, $126,000 net income, approximately $58,000.

Effect on retained earnings at the beginning of 1980, and the pro forma amounts for 1970 have not been recorded because the different assumptions that they are required to apply.

LIFO method retroactively.

Inventory quantities at one of the company's plants were reduced, reduction resulted in liquidation LIFO.

The effect of this liquidation had decreased cost ratios.

Approximately $21,000 increased net income, by approximately $9000, 14 cents per share.

Long term debt, stock redemption, income taxes, prepaid advertising expense, pension plan, Ohio income franchise tax.

OK, pension plan, capital leases, operating leases, OK, Tom, having quickly gone through all of this, and at the risk of making a material error, for which my manager would, a few things I wouldn't want to say on tape, uh, let me just think here, were you supposed to give me any additional information in regards to uh, the uh, sample error rate other than what we've talked about here,

(no, if you feel you really need other information)

Feel free to ask right?

We've taken no presample.

Have no uh, pre uh, conceived notions as to what the error rate should be?

Other than the inference made by my manager based on the sample size?

(That's right)
YOU CHOOSE THE SAMPLE SIZE, THE MONITOR WILL THEN INFORM YOU OF THE FOLLOWING INFORMATION.

OK, SO I WILL SAY WHAT YOU SHOULD SELECT.

YOU'LL GO THROUGH AND TELL ME WHAT THE ACTUAL Uh.

THE INFERENCE MADE BY YOUR MANAGER BASED ON THE SAMPLE SIZE,

YOUR MANAGER WILL MAKE ONE OF FOUR INFERENCES.

OK, I WOULD, AFTER ALL OF THIS,

I WOULD ATTEMPT TO MAKE A DECISION THAT I FEEL CONTROLS ARE ARE VERY TIGHT AND IF THAT'S THE CASE,

I WOULD, OF COURSE, I THINK GO WITH A SMALLER SAMPLE.

THAT MAY NOT BE THE RIGHT ANSWER.

Uh, I WOULD BE INCLINED TO GO WITH A SAMPLE OF 60 ITEMS,

(OK)

AND I WOULD ATTEMPT TO Uh, EVALUATE THOSE 60 ITEMS STATISTICALLY.

OK.

(ALL RIGHT, NOW YOU SELECTED A SAMPLE SIZE OF 60)

RIGHT

(SAMPLE OUTCOME WAS 2, IN OTHER WORDS, YOU FOUND 2 TRANSACTIONS

THAT DID NOT COMPLETELY MEET INTERNAL CONTROL SPECIFICATIONS)

OK

(YOUR MANAGER'S DECISION WAS TO CHOOSE LEVEL A)

LEVEL A, POPULATION ERROR RATE IS EQUAL TO 3%, IS EQUAL OR LESS THAN,

IS LESS THAN OR EQUAL TO 3%.

(BUT THE TRUE STATE OF INTERNAL CONTROLS WAS B)

WAS B.

(SO YOUR PAYOFF WAS THE COST OF SAMPLING, 20 CENTS TIMES

60, OR YOU LOST $12.)

GREAT, HOW MUCH DID WE BILL THE CLIENT FOR ALL THAT WORK?

I SUSPECT WE DIDN'T LOSE $12.

BUT YEAH, I SEE,

WELL THAT WAS,

THE ACTUAL STATE WAS B.

(RIGHT)

OUR SAMPLE INDICATED AN A,

OF 3% OR LESS,

BECAUSE WE ONLY HAD 2 ERRORS,

OK.

NOW WHAT DOES ALL THAT MEAN?

I DID SEE SOME PROBLEMS HERE AND Uh, I WOULD BE RELUCTANT

TO GO WITH A SAMPLE MUCH SMALLER THAN 60.
ON THE SECOND CLIENT
OK.
(AND THE EXACT SAME FACTS ARE HELD,
AND THE ONLY ADDITIONAL INFORMATION YOU HAVE IS THIS FROM
THE FIRST CLIENT).
I AM GOING TO READ BACK HERE WITH UH, LET'S SEE WHAT THE
INSTRUCTIONS SAY.
FROM EXPERIENCE, BECAUSE THE EXPERIMENT WILL BE
REPEATED A
LARGE NUMBER OF TIMES.
FOR EACH REPEAT OF THE EXPERIMENT,
ASSUME YOU ARE IN CHARGE OF THE AUDIT OF ANOTHER COMPANY,
BUT EACH NEW COMPANY IS IDENTICAL IN EVERY RESPECT TO
ROHR.
EXCEPT FOR THE NUMBER OF CASH DISBURSEMENTS IN THE POPULATION,
7000,
WHICH DO NOT COMPLY WITH THE STATED INTERNAL CONTROL PROCEDURES.
THIS NUMBER IS DETERMINED AT RANDOM BY COMPUTER IN A MANNER
WHICH IS INTENDED TO REFLECT THE FREQUENCY OF ALL CLIENTS
AUDITED BY BIG 8 CPA FIRMS WITH SIMILAR CHARACTERISTICS.
BECAUSE OF YOUR OWN EXPERIENCE,
YOU PROBABLY HAVE A PRETTY FAIR CONCEPTUALIZATION OF THE ERROR
RATE, BUT TO MAKE THE COMPUTER GENERATION PROCESS MORE EXPLICIT
TO YOU, HERE IS A SAMPLE OF 100 ERROR RATES GENERATED BY THE
COMPUTER.
OK, SO, WE HAVE THE SAME SITUATION,
THAT IS MY ANALYSIS,
THE FIRST TIME AROUND,
WE HAD A VERY GOOD, UH,
THAT WAS A VERY GOOD OUTCOME,
OUTCOME OF A AS IT SAYS HERE.
ACTUAL STATE OF B.
COST US $12, WELL.
UH, I WOULD SAY I'M NOT PARTICULARLY UH,
CONCERNED, AND MY TRAINING HASN'T LED ME TO BE PARTICULARLY
CONCERNED ABOUT COST.
I THINK IT'S MORE IMPORTANT TO DO THE JOB RIGHT,
THAN TO WORRY ABOUT SPENDING THE ADDITIONAL HOUR OR DAY OR 2
ON, UH,
SO THAT,
I BELIEVE THAT THE SAME SITUATION OVER AGAIN,
I DON'T THINK THAT I WOULD BE UH, WELL,
IT'S DISTURBING TO NOTE THAT THE UH,
WE DIDN'T HIT RIGHT ON THE BUTTON,
THOUGH I DON'T THINK THAT WE WERE TOO FAR OFF.
I COULD GO WITH 60 AGAIN.

A AND B.
LOOK AT THAT,
SO WE CAN STILL HAVE AN OCCURRENCE RATE OF 3 ITEMS AND STILL
FALL INTO THE UH, INTO THE A CATEGORY,
WITH A 3% OR LESS.
I GUESS.
I DON'T KNOW, THAT DOESN'T MAKE A WHOLE LOT OF SENSE TO ME BUT, UH
(WELL, YOU HAVE TO RECALL THAT UH, THERE IS A MANAGER. THE
MANAGER SAW 3 OUTCOMES, AND HE, AND THE MANAGER SAYS, BASED ON
THAT, I'LL CHOOSE A. OK. YOU MAY THINK THE MANAGER'S A LITTLE
SCREWWY, BUT THAT'S OK.)
I MEAN NO, I'M NOT SAYING THAT I ABSOLUTELY FEEL THAT WAY,
I WOULD SAY THAT THE REOCCURRENCE IS OUT OF THE SAMPLE OF
60 ITEMS,
UH, WOULD NOT BE THAT SIGNIFICANT,
OF COURSE, DEPENDING ON WHAT THE NATURE OF THE OCCURRENCES WERE,
UH, ALTHOUGH IT'S STARTING TO LOOK RATHER SIGNIFICANT.
OK, SO HERE WE ARE,
WE'VE HAD AN A & A B,
AN A & A B.
60 AND 3 GOING INTO COMPANY NUMBER 3.
SAME SITUATION.
(RIGHT)
SAME MANAGER? HA HA HA HA
(OH, YEAH, SAME MANAGER.)
OK, WELL I'M PROBABLY GOING TO GET HUNG 
BUT I UH, WOULD STICK WITH 60,
WHICH IS KINDA JUST A STANDARD NUMBER.

AUDIT SAMPLING.
(AUDIT SAMPLING)
(ARE YOU LOOKING UP SAMPLE SIZE TABLE NUMBERS?, WELL, 
YOU CAN
IF YOU WANT.)
CAN I?
(WELL, THERE'S THE SAMPLE SIZE TABLES FROM HERE)
IS THAT RIGHT?
(YES, IF YOU WANTED TO USE IT)
OK, GOOD SAMPLE,
WE TESTED 60,
WE HAD 5 EXCEPTIONS OR OCCURRENCES.
THE MANAGER SAID THAT'S GETTING INTO THE B RANGE IN HIS
OPINION, HE'S GONNA GRADE THAT A B.
SO THAT, OH, CUMULATIVE PAYOFF,
WE'RE IN THE POSITIVE RANGE NOW,
WE HIT RIGHT ON.
ALL RIGHT, LET ME GO BACK AND JUST READ OFF WHAT UH,
I CAN SEE THE PAYOFF IS GOING TO BE THE BOTTOM LINE.
IN ALL CASES YOU WILL BE ASSESSED 20 CENTS PER TRANSACTION
SAMPLED.
YOU WILL WIN $100 IF THE MANAGER'S DECISION CORRESPONDS WITH THE
TRUE STATE OF INTERNAL CONTROLS.
YOU WILL loose $4000 IF THE MANAGER INFERS A OR B AND THE
TRUE STATE OF INTERNAL CONTROLS IS D.
I SEE.
OK, OK.
SO, WE'RE DOING PRETTY WELL.
NOW, BACK TO THE NEW CLIENT.
AND SAME SITUATION,
SAME INTERNAL CONTROLS,
I DON'T KNOW, I GUESS WITH THIS EXPERIENCE BEHIND ME,
IS THAT UH, YOU KNOW,
I GUESS I HAVE NO REASON TO SAY ANYTHING OTHER THAN 60.
IF I WERE JUST STARTING,
YOU KNOW, SAME COMPANY,
DIFFERENT COMPANY BUT SAME CONTROLS,
AND SO FORTH,
I GUESS I WOULD HAVE NO REASON TO UH, TO NOT SAY 60 AGAIN,
I GUESS.
(OK)
I'VE BEEN DOING PRETTY GOOD,
HAVEN'T GOTTEN BURNT YET.
GOT A FEELING IT'S COMING,
BLAME IT ON THE MANAGER THOUGH.
...
OK, OK.
NO EXCEPTION.
THE MANAGER SAID A,
I WOULD AGREE WITH THE MANAGER IN THIS CASE.
CHOOSE A, AND GET $100 PAYOFF.
LESS, UH, WELL,
AS I SAY,
PERHAPS THESE RESULTS MIGHT INDICATE I THINK, THAT UH,
PERHAPS WE'RE PLAYING IT PRETTY SAFE WITH 60,
AND HAVE HAD GOOD RESULTS WITH 60, UH,
I'M THE CONSERVATIVE TYPE AND I'M MORE INTERESTED IN HAVING
AN ACCURATE AND COMPLETE AUDIT.
I don't think I would be inclined under this system, with this company, or with one like it, to necessarily reduce my sample size even based upon the good results we had before. So you can punch in 60 one more time. I don't know if I'm ruining your experiment here, by being so conservative or consistent, based upon these, but I wouldn't do anything differently. (Don't worry about my experiment.)

Yes, I'm doing a conscientious job here, doing what I would do.

Okay.

Oh, ha ha, rated it an A and it was A.

I'll tell you, the 20 cents per transaction is awful cheap. (If you feel that 60 is the uh, the one that is going to max maximize your payoff, then you should choose 60.) Yeah, well I think that the 20 cents per transaction is not a significant cost compared to the ultimate $4000 cost if we really miss, and I think that while you can get a valid sample, with as few as 25 or 30 transactions, in certain cases with a sample this large, 7000, I would be reluctant to go any further, why don't we try 60 again, next company.

And I'm just gonna look up something, you say I can do this?

(Ok, sure)

I mean, I can do this?, I just want to see if maybe I'm uh, (as you're reading through that book, you'll, kinda describe what tables you're looking at, or..) Well, I haven't looked at this book in a while and I'm not, as I say, I'm not a statistician.

I've taken a course or 2, and so I'm trying to find the determination of sample size, tables, which I have right here.
(If it's that part you're looking up, I can do that right here on the computer). Ok, I would do that, however, uh, in that case I wouldn't be absolutely certain to tell you exactly what I think you should look up. In looking up this though, I would be inclined to settle for nothing less than a 90% reliability factor, and maybe inclined to go to a 95% reliability factor uh, the upper precision limit, percentage rate of occurrence uh, let's see, expected percentage rate of occurrence; in the last couple of examples, we've seen on companies similar to this between this is the last one, oh, not bad, not bad, come up to $228. Uh, I would suspect that the expected rate of occurrence would be somewhere between 2 and 5 occurrences out of 60 that we've tested which would uh, convert to uh, say a maximum of 5 divided by 60, or 8%. 8%, to a low of 2 divided by 60, actually we had one zero, but uh, 3%. That, an expected rate of occurrence would be 4%, in this population, and my upper precision limit at 90% reliability, I would assume to be 10%, that brings us back to uh, well, 4%. Let us uh, let us be a little more conservative on this next one, ok.

I'm going to assume on this one that we have a uh, manager who is real statistically oriented and, I can show him these charts and tables, he would understand them probably better than I, and I would go with a 70 sample on this, on this next one.

Primarily because I think that uh, this statistical sampling in many cases
Leads you to do more testing than really needs to be done,
you're assuming that you're not taking all of the judgment
away from the auditor,
however, I think that the auditor's judgment is awful important,
and therefore uh,
-----
c and c,
we tested 70,
so we still made some money,
but, uh,
it's probably a good thing that we tested 70.
because the actual occurrence rate was 9,
maybe those last 10 that we tested turned up 6 extras.
we would have selected a and ended up with a c or something.
ok,
and I'm onto the next job.
and I've got another uh,
another manager on this one,
and assuming that he's the same as all the best,
or some of the others,
I would go back to 60,
had good results with 60,
I feel confident with 60.
short staffing in the office,
don't have people out there testing ok.
-----
manager said a,
it's an a.
good. good.
where did we end up at the beginning?
just, we're off by one,
ok, tom, I have no reason not to say 60 again,
-----
a and a.
how do I make $4000?
I can't.
best I can do is about what I'm doing really.
except I were to cut it down,
and we still hit on.
for all I'm thinking here,
and I'm not trying to outsmart your machine,
but I think that after,
how many of those are we going to run?
(well, I'd rather not say,)
rather not say.
(let's just say there's a large number, um, if we get to the
point where you think it's pointless, and you know,
you get kinda...)
NO, I'M SURE THERE'S MORE GOING ON THAN WHAT I'M AWARE OF.

(OK, OK, WELL, WE CAN, ALL I'M SAYING IS WE CAN STOP ANY TIME
YOU WANTED, I WOULD LIKE TO DO A LARGE NUMBER OF THESE, OK)

UH HUH, UH HUH, I SEE.

I SEE,

(I ALSO RECOGNIZE THAT YOU'VE WORKED VERY HARD ON THE CASE,
AND THAT YOUR MENTAL GEARS STOP FUNCTIONING AT A CERTAIN POINT,
AND IF YOU FEEL THAT IT'S NOT HELPING ANY TO GO ON THEN WE'LL
STOP.)

NO, I DON'T HAVE A PROBLEM WITH GOING ON.

(I'M ALSO AWARE THAT YOU HAVE OTHER THINGS TO DO BESIDES
THIS EXPERIMENT).

UH, I GUESS WITH ALL OF THIS PAST EXPERIENCE,
AND Uh, THE FACT THAT IN EACH CASE,
I BELIEVE THAT WE'VE BEEN MATERIALLY CORRECT,
IN ALL THESE ITEMS,
WELL, I WOULD BE CURIOUS TO SEE WHAT HAPPENED IF WE TOOK A 50,
I'M GOING TO STICK WITH A 60,
CONTINUE MY CONSERVATIVE APPROACH,
AND SEE THAT CERTIFICATE THERE,
I WANT TO KEEP THAT ON MY WALL.
HA HA HA HA,
AND A COST OF 20 CENTS PER TRANSACTION IS NOT THAT SIGNIFICANT.
NOW, MAYBE, YOU KNOW,
A HIGHER COST FOR THE NUMBER OF TRANSACTIONS,
WOULD TEND TO DETER ME MORE FROM TESTING SO MUCH,
AND PARTICULARLY WITH THESE RESULTS,
I DON'T CONSIDER A 20 CENTS TO BE A VERY HIGH COST TO PAY.
SO LETS GO WITH 60 AGAIN,
I GOT A FEELING I'M GOING TO GET NAILED ONE OF THESE TIMES.

$4000 IS A VERY LARGE COST TO PAY.
(UH)
OBVIOUSLY.
(YEAH)
THAT TAKES MY CERTIFICATE OFF THE WALL.
UH OH,
YOU KNOW WE HAVEN'T COME OUT YET WITH A B AND AN A OR SOMETHING,
MY MANAGER IS,
HAS NEVER COME OUT TOO HIGH.
WELL I'M LEFT WITH $566.
OK, THIS TRIAL WAS A MINUS 12 DOLLARS.
829  UH, I WOULD LIKE TO STICK WITH 60,
830  I THINK I PROBABLY WILL FOR THE DURATION OF THE
     EXPERIMENT.
APPENDIX A FOOTNOTES,
JUST A COPY OF THEIR FINANCIALS.
APPENDIX B,
OVERALL FLOWCHARTS,
THIS SYSTEM IS CHANGING FOR US,
WE'VE DEVELOPED A NEW FLOWCHART SYSTEM.
(THIS BELONGS IN FOOTNOTE 2)
OK.
RANDOM SAMPLE;
(I DON'T MEAN TO HARASS YOU, BUT AS YOU ARE READING,
YOU MIGHT
WANT TO READ ALoud.)
DO YOU WANT ME TO READ THE WHOLE THING ALoud? IS THAT
WHAT...
(WELL, DO WHATEVER YOU WANT, BUT WHATEVER YOU ARE
READING,
READ OUTLOUD).
OK.
UH, CHOOSE A SAMPLE SIZE, THE MONITOR WILL INFORM YOU
OF
THE FOLLOWING INFORMATION,
SAMPLE OUTCOME,
INFEERENCE MADE BY YOUR MANAGER.
OK, SO WE'RE JUST GOING TO, IN EFFECT,
UH, DO A RANDOM SAMPLE THEN DRAW THE CONCLUSIONS
THROUGH THE
COMPUTER,
AS WHAT THE ERROR RATE IS
BASED UPON OUR SAMPLE SIZE,
AND IT LOOKS LIKE IT'S NOT 95 AND 5,
BECAUSE THE 3%, 7%, LOOKS A LITTLE
A LITTLE DIFFERENT.
UH, OK.
NOW, WE'RE GOING TO USE THAT AS AN OVERALL REVIEW.
IN WHAT, TRYING TO ESTIMATE WITH THE OVERALL,
WHAT WOULD HAPPEN IF THERE WAS AN ERROR,
AND HAVE TO RELY ON THAT.
OK,
THE NET PAYOFF,
AFTER WE GO THROUGH ALL OF THE NONCOMPLIANCE DECISIONS
IN OUR TEST,
AND WHAT'S, WHAT WE ESTIMATE THE TRUE SITUATION TO BE,
THEN WE'RE GOING TO TRY TO FIGURE OUT WHAT IT ACTUALLY
IS.
IS KNOWN,
THE PAYOFFS,
REWARDS AND PENALTIES,
A,
IN ALL CASES YOU WILL BE ASSESSED 20 CENTS PER
TRANSACTION
SAMPLED.
YOU WILL WIN $100 IF YOUR MANAGER'S DECISION CORRESPONDS WITH THE TRUE STATE OF INTERNAL CONTROLS.
AND THE MANAGER'S DECISION, I GUESS, RELIES UPON WHAT WE FIND IN OUR COMPLIANCE TEST, UH, EXTRAPOLATED, OR, OR, UH, SAMPLED, DECISION BASE IS THROUGH THE COMPUTER.
YOU WILL LOSE $4000 IF THE MANAGER INFERS A OR B, AND THE TRUE STATE OF INTERNAL CONTROLS IS D. THAT MEANS UH,
IN EFFECT, THAT THE MANAGER INFERS THAT THE POPULATION RATE IS LESS THAN OR EQUAL TO 7%, AND IT'S ACTUALLY IN EXCESS OF 12%, WE'RE GOING TO LOSE $4000.
IF THE MANAGER ERRS BY ONE LEVEL IN EITHER DIRECTION, NO REWARD OR PENALTY WILL ACCRUE TO YOU.
IN ONE LEVEL EITHER DIRECTION.
OK.
A 2 LEVEL ERROR OTHER THAN THE ERROR DISCUSSED PREVIOUSLY, WILL CAUSE YOUR WINNINGS TO DECREASE BY $50.
A 3 LEVEL ERROR OTHER THAN THE ERROR DISCUSSED PREVIOUSLY, WILL CAUSE YOUR WINNINGS TO DECREASE BY $100.
UH, OK.
NATURALLY, THE GREATER SAMPLE SIZE YOU CHOOSE, THE GREATER THE LIKLIHOOD YOUR MANAGER WILL HAVE OF MAKING THE CORRECT INFERENCES.
YOU WILL NEED TO BALANCE THIS INTUITIVE STATISTICAL PROPERTY AGAINST THE COST OF SAMPLING.
COST OF SAMPLING, 20 CENTS AN ITEM.
YOU'LL HAVE THE OPPORTUNITY TO LEARN FROM EXPERIENCE BECAUSE THE EXPERIMENT WILL BE REPEATED A LARGE NUMBER OF TIMES.
FOR EACH REPEAT OF THE EXPERIMENT,
ASSUME YOU WILL BE IN CHARGE OF AN AUDIT OF ANOTHER COMPANY,
BUT EACH NEW COMPANY IS IDENTICAL IN EVERY RESPECT TO BOTH, EXCEPT FOR THE NUMBER OF CASH DISBURSEMENTS IN THE POPULATION.
7000, WHICH DO NOT COMPLY WITH THE STATED INTERNAL CONTROL PROCEDURES.
EACH NEW COMPANY...
OH, OK, WE'RE JUST GOING TO INCREASE THE NUMBER OF ERRORS.
STEP IT UP, IT LOOKS LIKE,
EACH TIME WE HAVE A NEW SITUATION.

THIS NUMBER IS DETERMINED AT RANDOM BY COMPUTER,

IN A MANNER WHICH IS INTENDED TO REFLECT A FREQUENCY

OF ERROR

RATES OBSERVED OVER ALL CLIENTS AUDITED BY BIG 8 CPA

FIRMS.

WITH SIMILAR CHARACTERISTICS.

BECAUSE OF YOUR AUDITING EXPERIENCE,

YOU HAVE A PRETTY FAIR CONCEPTUALIZATION OF ERROR RATE

FREQUENCY,

BUT TO MAKE THE COMPUTER GENERATION PROCESS MORE

EXPLICIT TO YOU,

HERE IS A SAMPLE OF 100 ERROR RATES AND PERCENTAGES

GENERATED BY

THE COMPUTER.

THIS IS IN COMPLIANCE TESTING?

(THAT'S RIGHT)

ACROSS BIG 8 FIRMS.

HA, AS MUCH AS 11%, 15%, QUITE HIGH.

LOW IS .2%.

OK.

NOW I'M SUPPOSED TO CHOOSE A SAMPLE SIZE AT THIS

POINT?

LET'S LOOK AT THE FINANCIALS AND SEE WHAT KIND OF A

SIZE

INSTITUTION OR ASSOCIATION WE'RE TALKING ABOUT.

JUST REVIEWING THE VARIOUS BALANCES;

TOTAL ASSETS,

VARIOS CAPTIONS,

ANY SIGNIFICANT CHANGE OR FLUCTUATIONS THAT WOULD

CAUSE

ANY, ANY UNUSUAL ITEMS.

TOTAL ASSETS OF ABOUT $8,000,000.

'79, NOTE 2.

FINANCIAL STATEMENTS FOR 1980 HAVE BEEN RETROACTIVELY

RESTATED.

ADJUSTMENT IN INVENTORIES,

NO REAL EFFECT ON OUR COMPLIANCE TEST.

IN GENERAL,

IT LOOKS LIKE THEY'VE SWITCHED, ER, GONE FROM A HIGH

CASH

POSITION,

TO EXTENDED LIABILITIES.

AGAIN, THOUGH,

NO EFFECT ON OUR TEST.

OK,

REVENUES ARE ABOUT $15,000,000.

REVIEWING THE INCOME STATEMENT.

NO UNUSUAL ITEMS,

HAD AN EXTRAORDINARY ITEM IN '79,

AGAIN NO EFFECT ON OUR.

EARNINGS PER SHARE,

DECREASED SIGNIFICANTLY,

I GUESS THAT HAS TO DO WITH THE REORGANIZATION.
AGAIN, NO EFFECT ON A COMPLIANCE TEST.
JUST TRYING TO GET A FEEL FOR WHAT SIZE OF A COMPANY WE'RE TALKING ABOUT.
GETTING A GENERAL UNDERSTANDING.
OK.
REVIEW THE FOOTNOTES.
FOR ANY UNUSUAL ITEMS THAT MIGHT AFFECT ME.
NOTHING THERE THAT'S GONNA TOO MUCH IMPACT ON MY COMPLIANCE TEST.
OTHER THAN MY DOLLAR AMOUNTS I'M TALKING ABOUT.
REVIEWING APPENDIX B,
OF FLOWCHARTING SYSTEMS.
MEMO ON INVENTORY CONTROL,
WHEN AN ORDER IS TAKEN,
The engineering department is given a copy of the order to determine the specs.
Following the flow of materials requirements,
to the inventory control department head.
Inventory order quantities,
materials needs,
inventory control determines the need to order materials, uh,
looks like there's good controls there for inventory.
when a determination is made that an item should be ordered,
the inventory control department forwards to the purchasing agent,
the requirements list from which he is to make purchase orders;
this is done daily.
so we have edp help on the inventory control.
also the inventory didn't change significantly.
dollar amounts of inventories$ million 6 prior year,
$ million 4 this year so,
it appears that the controls over inventories are pretty good,
pretty stable.
reviewing flowchart,
comparing it to the flowchart symbols,
uh, just a flow of the paper basically,
i'm looking for strengths outlined,
uh, ok.
i'm gonna work from the general ledger back,
just to see if i can identify some of the strengths that we might want to do a compliance test.
uh, distribution,
review, uh,
uh, our requirements again,
purchases,
we're looking at the purchases cycle and payables,
We're going to take a random sample of cash disbursements.

Uh, so I'm looking for cash disbursements strengths. OK. One strength is going to be that they do have that inventory order quantity.

And, uh,

Purchasing agent,

Daily ordering, ok.

Purchase approval limitation;

Assets over $1000 have to have approval of the president,

Assets over $5000 Board of Directors,

Purchase orders over $10,000 for production material,

President,

All other PO's, purchasing agent.

Uh, ok it says assets over $1000.

So, uh, fixed assets over, ok.

Proper approval,

We might want to see that that's implemented.

Uh,

Purchase order to vendor,

5 copies or 8 copies filed.

Uh, cost accounting system;

One copy destroyed.

Purchase order copy 2 which went to accounting,

So the cost accountant determines the cost distribution,

And destroys the copy.

Not too good.

Uh, ok.

Just trying to get an overall feel for whether they're going
to a significant number of strengths that we're even going to
test,

And how far we want to take those.

If the system in general looks like a good system
Then we should determine what strengths that we might want
to rely on.

Uh, and certainly we have to consider materiality.

Uh, the balance sheet of what kind of items we're going to test.

Purchasing,

Unmatched receiver file reviewed every month,

That's, I guess.

Ok.

Looks like we have a pretty standard system;

The fact that purchase orders are matched with the freight bills

When they come in,

Uh, receiving reports filed,
RECEIVING REPORTS MATCHED TO THE INVOICE WHEN IT COMES IN.

UH, UH, ... ACTUALLY ACCOUNTS PAYABLE,
ACCOUNTING CLERK REVIEWS THE CHECKS,
ATTACHES,
ACCOUNTING CLERK, UH,
OK.

CONTROLLER APPROVES ALL CASH DISBURSEMENTS.
UH, SO THAT'S A GOOD CONTROL
THAT EVERYTHING'S GOING THROUGH THE CONTROLLER.
AND WE ALREADY SAID THAT INVOICES..
APPROVED, SO,
UH, LOOKS LIKE WE'RE GOING TO HAVE A RELATIVELY SMALL NUMBER
OF LARGE ITEMS,
A LOT OF SMALL DOLLAR TYPE ITEMS,
ESPECIALLY IF THEY'RE ORDERING EVERY DAY.
SO WE'RE GOING TO HAVE A POPULATION OF, OF,
UH, OUR 7000 ITEMS,
LOOKS LIKE THEY WILL BE PRIMARILY SMALLER TYPE ITEMS,
UH, ALL HANDLED MUCH THE SAME WAY.
DISTRIBUTION TICKETS, UH,
GET A VOUCHER FOR PAYMENT,
UH, NOT SIMULTANEOUS, OK,
SO WE,
PROCESSING NCB OPERATOR,
POSTS THE GROSS AMOUNTS AND THE CHECK NUMBERS INTO THE CHECK
REGISTER.
THAT'S AFTER THE CONTROLLER HAS APPROVED THEM.
The invoice, the purchase orders are filed,
ALPHA FILE, OK.
SO ACCOUNTING CLERK IS THE FINAL PERSON WHO HANDLES
THE CHECKS,
HAS THEM SIGNED,
SEE BELOW.
MAKE SURE THAT THE PROPER APPROVAL NCB POSTS THE GROSS AMOUNTS
AND CHECK NUMBERS.
OK.
UH AT THIS POINT WE WANT TO TRY AND DETERMINE WHAT
SAMPLE SIZE UH,
WE'RE GONNA HAVE.
JUST GONNA REVIEW THE REQUIREMENTS,
UH, YOUR MANAGER HAS INSTRUCTED YOU TO CHOOSE A RANDOM
SAMPLE OF CASH DISBURSEMENTS AND TRACE THESE TRANSACTIONS BACK
THROUGH THE SYSTEM,
CLASSIFY EACH TRANSACTION AS EITHER COMPLYING OR NONCOMPLYING
WITH THE INTERNAL CONTROL REQUIREMENTS FOR THAT TRANSACTION.
AND AGAIN, NONCOMPLIANCE IS IF IT FAILS TO MEET ALL OF THEM.

ONLY ONE KNOCKS IT OUT.

THE EXPERIMENT WILL FIRST PROCEED BY HAVING YOU CHOOSE A SAMPLE SIZE AND THEN THE COMPUTER WILL TELL US THE ERROR RATE.

WE'RE GOING TO MAKE A CONCLUSION,

AND THEN WE'RE GOING TO FIND OUT ACTUALLY WHAT THE TRUE STATE OF AFFAIRS IS.

WE'RE ALSO GOING TO DETERMINE OUR NET PAYOFF.

UM, AND AGAIN,

NET PAYOFF, 20 CENTS PER TRANSACTION SAMPLED.

SO WE HAVE A LOW COST PER TRANSACTION, UH, COMPARED TO UH,

OK.

$100 IF WE, OK.

7000 UH, POPULATION.

MATERIALITY, I GUESS, DOESN'T ENTER IN AT THIS POINT.

WE'RE DOING A COMPLIANCE TEST.

UM, WE WANT TO DETERMINE HOW MANY WE WANT TO DO.

MINIMUM OF 20.

UM, LET ME SEE,

MAXIMUM OF 60.

THAT'S 95 AND 5.

POPULATION IS LESS THAN OR EQUAL TO 3%.

SO UH, OK,

TYPICALLY,

OK, NATURALLY, THE LARGER SAMPLE SIZE YOU CHOOSE,

THE GREATER THE LIKELIHOOD YOUR MANAGER WILL HAVE OF MAKING THE

CORRECT INFERENCE.

INTUITIVE STATISTICAL PROPERTY AGAINST THE COST OF SAMPLING.

UM, ACCOUNTS PAYABLE, CASH DISBURSEMENTS TYPICALLY, WITH ONLY 7000,

LET'S SEE 7000 CASH DISBURSEMENTS,

THEY'RE ALL RELATIVELY SMALL DOLLAR AMOUNTS.

THEN IT'S OUR FIRST YEAR AUDIT.

WE'D TYPICALLY MAYBE EXPAND THAT SOMEWHAT.

UM, I WOULD CHOOSE A SAMPLE SIZE I THINK FOR THE FIRST YEAR

OF 60 ITEMS.

ALL THAT TO GET TO THERE.

(OK, THAT'S FINE)

... (OK, HERE'S THE RESULTS OF THE FIRST TRIAL OF THE EXPERIMENT.)

(UM, YOU CHOSE A SAMPLE SIZE OF 60, UH, THE COMPLIANCE TEST

FOUND 2 TRANSACTIONS NOT COMPLYING, BASED ON THAT, YOUR
Manager inferred the true state of internal controls was A.

When in fact it was B. Therefore, you are assessed a $12 penalty which consists of B here, which is zero, plus the sampling costs, 20 cents, for $12.)

B which is zero.

OK, wait a minute.

(IT was off by one level)

Oh, OK, if the manager ended by one level in either direction,

OK, no reward or penalty will accrue to you.

OK, so we missed the boat by one level.

OK, we're saying that the true error rate is between 3 and 7%.

We said it was less than 3%.

We placed maximum reliance when there were some minor departures.

Uh, and substantive testing will be at a higher level.

Uh, OK.

$12, am I supposed to react to that, as far as...?

(Oh well, you're now working for another client)

(for another client which is identical in every respect to Rohr)

Why don't you read exactly what it says there.)

OK.

You will have the opportunity to learn from experience,

Because the experiment will be repeated a large number of times.

For each repeat of the experiment,

Assume you are in charge of the audit of another company

But that each new company is identical in every respect to Rohr

Except for the number of cash disbursements in the population

Which do not comply with the stated internal control procedures.

Again, the only thing that is going to change is the number of errors.

This number is determined at random by computer

In a manner which is intended to reflect the frequency of error rates observed over all clients audited by Big 8 CPA firms

With similar characteristics.

Because of your auditing experience you probably have a pretty fair conceptualization of of error rate frequency.

OK. Um, so I think we got off pretty good on that one.
I think typically we overaudit on cash disbursements, or I'm sorry, on the vouching of items anyway, so uh, I think that would balance out. We typically don't reduce it as much as we can, so I think that the first situation turned out pretty good. The only thing that's gonna change is the number of errors, uh, I want to go with 60 again.

(and virtually the same result, you're only difference being that the sample outcome this time was 3 instead of 2.) Ok, so we're getting 2 or 3 errors, and again, I...

in the case that they're fairly minor, in amount, our substantive tests are going to compensate for it. Uh, after some time we might reduce it as low as 40, and again, balancing that with;

"Hey look guys, you're going to do a fair amount of substantive auditing anyway", just to get a good gut feeling for the thing, we tend to be happier with substantive testing so let's try a sample of 40.

hopefully for a moderate reliance, sample size 40, we had 4 errors that time, decision was B, in other words, I was shooting for moderate reliance, we got moderate reliance.

payoff for this trial was $92. in effect we, so we made the right decision.

$100, we got a bonus, but then we net that with the cost of 20 cents per sample.

uh, I'm going to try another one at 40, just for the moderate reliance.

ok, good job.

no errors on that one.

so, we would have guessed A, in other words, we would have placed maximum reliance on the system, truly was maximum reliance, and we had another 92 dollar payoff.
SAME AS BEFORE WE GUESSED IT RIGHT.

ALTHOUGH, I GUESS,

I WAS REALLY SHOOTING FOR MODERATE RELIANCE,

OH SO, I WAS PLANNING ON DOING SOME SUBSTANTIVE TESTING.

I COULD MAYBE EVEN REDUCE THAT.

UH, IN SOME SITUATIONS I COULD SEE WHERE YOU MIGHT WANT TO DO VER MINIMAL AMOUNT OF TRANSACTION TESTING,

OR WALK THROUGH THE SYSTEM AS YOU WILL,

UH, AND AGAIN,

IN HOPES THAT WE'RE GOING TO DO ANALYTICAL REVIEW PROCEDURES,

WE'RE GOING TO VOUCH SIGNIFICANT CAPITAL ITEMS, ETC.

SO WE MIGHT DO AS LOW AS 20 I THINK,

WE'LL TRY 20.

WE HAD ONE ERROR IN A SAMPLE OF 20.

BUT THAT STILL ALLOWED US TO INFERENCE THAT THE POPULATION ERROR RATE WAS LESS THAN 3%.

THE TRUE SITUATION WAS LESS THAN 3%.

PAYOFF THIS TRIAL, $96.

SO, THE SAME THING.

UH, SO WE'VE BEEN BETTER OFF WITH THE SMALLER SIZE SAMPLES.

I'M GOING TO REVIEW THE SAMPLE OF 100 ERROR RATES JUST FOR CURIOSITY.

AGAIN, SITUATION C - POPULATION ERROR RATE BETWEEN 7 AND 12%

THERE ARE SIGNIFICANT DEPARTURES FROM INTERNAL CONTROL PROCEDURES AND EXTENSIVE SUBSTANTIVE TESTING TESTING IS NECESSARY, THAT'S C.

SO B YOU MIGHT CALL MODERATE RELIANCE,

C IS NO RELIANCE,

WE'RE GONNA DO,

OUR TEST TELLS THAT WE CAN'T REALLY RELY ON THEM, OK,

SO IT'S 7 TO 12%.

THERE ARE SEVERAL OVER,

MAYBE TEN TO TWELVE,

12% OF THE COMPANIES,

IT APPEARS,

ARE IN THAT CATEGORY.

AGAIN THE MAXIMUM LOSS;

YOU WILL LOSE $4000 IF THE MANAGER INFERS A OR B,

WHICH IS MAXIMUM OR MODERATE,

AND THE TRUE STATE IS D.

D IS THE WORST,

I GUESS YOU STILL LOSE IF IT'S A OR B AND THE TRUE SITUATION

IS C,
UH, BUT IT'S JUST BY A DEGREE OF THE SUBSTANTIVE TESTING.

D IS GREATER THAN 12%,

AND OUT OF THIS SAMPLE OF 100 RATES,

THERE'S 1, 2, 3, 4, 5, 6 OF THEM OUT OF 100.

THE WORST POSSIBLE SITUATION.

AND THOSE ARE AVERAGE UH,

FREQUENCY OF ERROR RATES OVER ALL CLIENTS AUDITED BY BIG 8 FIRMS

WITH SIMILAR CHARACTERISTICS.

I THINK I TEND TO LIKE THE MODERATE RELIANCE, UH,

I KNOW THERE'S POTENTIAL FOR PROBLEMS THERE.

UH, I THINK I'D TEND TO GO WITH 20 ITEMS.

(OK)

20 ITEMS;

WE HAD AN ERROR OF 1 WHICH TELLS US THAT OUR,

WE GOT MAXIMUM RELIANCE,

UH, TRUE STATE OF AFFAIRS,

UH WAS B.

OR AN ERROR RATE, MODERATE RELIANCE.

PAYOFF THIS TRIAL WAS WE LOST $4.

UH, THAT'S BECAUSE WE MISSED IT BY ONE LEVEL EITHER DIRECTION;

WE'VE GOT NO PENALTY,

20 ITEMS COST US $4 TO PROCESS.

UH, SO WHAT WE HAVE LOOKS LIKE

IF WE GO WITH 20 ITEMS I THINK WE GO IN WITH THE IDEA THAT

WE'RE REALLY GOING TO HAVE MODERATE RELIANCE BECAUSE

HERE IF WE HAVE,

I DON'T THINK WE CAN EXPECT TO HAVE MAXIMUM RELIANCE,

WITH 20 ITEMS.

SO IF I CHOOSE 20 ITEMS,

I GO IN WITH THE IDEA THAT UH I'M GOING TO DO MODERATE, OR UH,

MODERATE RELIANCE I'M GOING TO DO SOME DETAILED TESTS,

OR

I'M SORRY, SOME SUBSTANTIVE TESTS.

I'M GOING TO DO SOME VOUCHING,

I'M NEVER GOING TO PLACE MAXIMUM RELIANCE ON 20 ITEMS,

UH,

THE OTHER SITUATION;

60 ITEMS,

UH, THAT I HAD,

I'M DOING MORE HOPING TO PLACE MAXIMUM RELIANCE,

I'M CLOSE,

IT LETS ME GET THE SITUATION I WANT.

BUT IT COSTS ME TO DO THAT.

AND WHAT AM I SAVING?

I'M SAVING SUBSTANTIVE TESTING,

WHICH WITH 7000 TRANSACTIONS,

OVER 12 MONTHS,

UH, SO WE GOT 500 TRANSACTIONS A MONTH,
THAT'S OVER DAYS,
AND ALL THE LARGER AMOUNTS REVIEWED,
I THINK, BASED ON THIS,
I WOULD GO WITH ANOTHER SAMPLE OF 20 AND TRY TO GET
MODERATE
RELIANCE.

HOW MANY OF THESE DO YOU WANT TO DO?

(WELL, I'D RATHER NOT TELL YOU HOW MUCH, HOW MANY
WE'LL DO,
UM, LET'S JUST SAY THERE'S A LARGE NUMBER OF THEM)
OK,
(UH, WHEN YOU GET TO THE POINT WHERE YOU FEEL THAT
IT'S
POINTLESS, THEN, OF COURSE, WE'LL STOP.)

SAMPLE SIZE FACTOR OF 20.
OH OH OH OK.
SAMPLE SIZE,
OK.

SO WE HAD SAMPLE SIZE FACTOR OF 20,
WE HAD 2 ERRORS,
THAT MADE US PICK B,
WHICH IS MODERATE.
RELIANCE.
THE SITUATION WAS ACTUALLY C.
SO WE LOST $4.
IN OTHER WORDS,
WE WOULD HAVE DONE MODERATE TESTING AND WE SHOULD HAVE
BEEN
DOING ExtENSIVE.
SO, UH, WE BROKE EVEN ON THE ERROR,
BUT WE PAID FOR OUR PROCESSING.
WE HAD ANOTHER SITUATION OF 20 WHERE WE HAD ONE ERROR,
AND WHERE WE'LL SAY MAXIMUM.
OK, SO WITH 20 ITEMS,
ONE ERROR MAXIMUM,
2 ERRORS MODERATE,
UH, OK.
DO A SAMPLE SIZE FACTOR OF 40,
TO LOOK AT SOME TREND HERE.

AGAIN, 40.
ZERO ERRORS,
WE SAID IT WAS MAXIMUM RELIANCE,
WE PLACED MAXIMUM RELIANCE,
IT REALLY WAS THE SITUATION,
WE GAINED $92.
UH, OK.
(WHAT, ARE YOU CALCULATING?)
OK, I'M SORRY,
I'M CALCULATING OVERALL UH, PERCENTAGES,
I'M SORT OF DOING IT IN MY HEAD.
2 OUT OF 20,
532  SO 10%,
533  4 OUT OF 40,
534  SO THE ONLY SITUATIONS WE'VE HAD SO FAR,
535  HAVE BEEN WHERE WE'VE HAD MAXIMUM RELIANCE,
536  OR WHERE WE'VE HAD 10% ERRORS,
537  IN OTHER WORDS;
538  WE'RE ALLOWING 10% ERRORS IN THE POPULATION.
539  AND YET WE CAN STILL PLACE MODERATE RELIANCE ON THE
      SYSTEM.
540  10% ERRORS,
541  I'M TRYING TO DETERMINE WHAT KIND OF AN IMPACT THAT
      WOULD
542  HAVE ON MY SUBSTANTIVE TESTS,
543  WHETHER I WOULD BE UH,
544  10% ERRORS,
545  THAT'S UH,
546  IF THEY WERE ALL LARGE DOLLAR AMOUNTS WE'D HAVE A
      COLUMN.
547  UM, I'M JUST TRYING TO GET A HANDLE ON WHAT I'M
      ALLOWED ON
548  ERRORS.
549  CERTAINLY IF I ANTICIPATE ERRORS I MIGHT,
550  THAT'S GOING TO ENTER INTO MY UH, MY UH,
      SAMPLE SIZE.
551  FACTOR.
552  I'M GOING TO SELECT ANOTHER SAMPLE OF 40,
553  JUST SO I CAN DETERMINE WHERE,
      WHERE THE BREAK EVEN IS.
554  ...
555  OK,
556  HAD A SAMPLE SIZE OF ONE,
      WHICH ALLOWS MAXIMUM RELIANCE.
557  NOTHING UNUSUAL THERE.
558  ANOTHER 40.
559  ...
560  OK, THIS,
      SAMPLE SIZE FACTOR OF 40.
561  HAD AN ERROR OF 4,
      AGAIN, THAT 10%.
562  ALLOWS MODERATE RELIANCE.
563  AND THE TRUE STATE OF RELIANCE WAS MODERATE.
564  WE GOT A PAYOFF OF $92.
565  UM, OK.
566  JUST FOR, UH,
567  FOR CURIOSITY,
568  WE'LL TRY ONE UNUSUAL,
      I'VE NEVER DONE A SAMPLE SIZE LESS THAN 20,
569  BUT THERE'S SITUATIONS WHERE WE MIGHT JUST WANT TO DO
      A
570  WALK THROUGH, UH,
571  ONE PER MONTH,
572  TO SEE THAT THE SYSTEM'S THERE.
573  BE CURIOS TO KNOW WHAT WE GET WITH 12.
574  ...
SAMPLE SIZE FACTOR OF 12,
WE HAD NO ERRORS,
GAVE US AN A.
MAXIMUM RELIANCE,
WHICH UH,
WE DIDN'T REALLY INTEND TO PLACE MAXIMUM RELIANCE UH,
EVEN THOUGH DECISION TELLS WE MIGHT DO THAT.
TRUE STATE OF AFFAIRS WAS MODERATE.
AND THAT'S REALLY WHAT WE'VE INTENDED IT.
IT WOULD COST US $2.40 TO TELL US THAT.
SO, UH,
LET'S TRY ANOTHER ONE OF 12 JUST TO COMPARE THAT TO 20.

AGAIN NO ERRORS,
UH SAME SITUATION AS BEFORE.
ONE MORE AND I'D LIKE TO GET SOMETHING DIFFERENT THAN 12.

SAMPLE SIZE FACTOR, OK.
WELL, SO WE GET A PAYOFF ON THIS UH,
OK, SAMPLE SIZE FACTOR OF 12,
WE HAD 2 ERRORS,
WHICH SAYS NO RELIANCE IN EFFECT,
OR EXTENSIVE TESTING.
BUT YET THAT'S WHAT WE SAID WE WERE GOING TO DO,
AND THAT'S WHAT WE ENDUED UF.
ALRIGHTEE, UH,
OK, SO, SO FAR, WE'VE HAD SAMPLE SIZE OF 12, 20, 40
AND 60.
UH, ONE MORE OF 12,
JUST TO GET A BROAD SAMPLE.

AGAIN, NO ERRORS,
The 12's ARE COMING OUT LOOKING PRETTY GOOD,
BUT I THINK THAT WOULD BE HARD TO SELL,
THAT WE ONLY NEED TO DO 12 ITEMS.
TO TEST THEM,
I THINK THAT WE'VE GOTTEN PRETTY LUCKY THERE BUT,
I THINK THAT WE'RE IN A PRETTY GRAY AREA AS OPPOSED TO
THE 20.
1 OR 2 ERRORS OUT OF 20 ITEMS, UH,
LETS TRY ANOTHER SAMPLE SIZE OF 20.

OK, NO ERRORS, SO WE CAME OUT LOOKING PRETTY GOOD.
UH,
LET ME ADD A COMMENT HERE,
THAT THE REASON THAT WE TEND TO GO SO,
ER, THE REASON I'M PICKING SMALLER SAMPLE SIZES
IS THAT I REALLY DON'T FORSEE A SITUATION WHERE I'D HAVE
MAXIMUM RELIANCE IN MOST CLIENTS THAT I'M FAMILIAR WITH.

AND MAXIMUM,

IN OTHER WORDS,

WE WOULD REDUCE SUBSTANTIVE TESTING TO THE BONE,

I THINK WE STILL TEND TO DO SOME MINIMAL AMOUNT OF

SUBSTANTIVE TESTING, THEREFORE,

IF I DON'T PERCEIVE THE MAXIMUM SITUATION OCCURING,

THAT I COULD SELL THAT,

"HEY, LOOK WE GOT, WE TESTED 120 ITEMS AND WE ONLY FOUND

5 ERRORS, WE CAN RELY ON THAT SYSTEM,"

I CAN'T SELL THAT AS MUCH AS I CAN SELL 20 OR 40 OR 60 ITEMS,

AND GET HAPPY WITH THE SYSTEM,

OR FAMILIAR WITH THE SYSTEM,

THAT IT WORKS,

AND REDUCE OUR SUBSTANTIVE TESTING.

I'M, IT APPEARS THAT I'M PREJUDICED TOWARD THE SMALLER

SAMPLE SIZES.

THAT'S WHERE WE'VE BEEN GOING,

BUT I THINK IT WOULD BE CURIOUS TO KNOW HOW IT WILL COME OUT.

BUT LET'S SAY 100 ITEMS,

100 ITEMS,

WE HAD 8 ERRORS,

8, OK.

THERE'S MODERATE RELIANCE,

THAT WAS ACTUALLY,

WE MISSED IT.

IT WAS ACTUALLY MINIMAL RELIANCE,

WE GOT $20.

SAMPLE SIZE FACTOR OF 20, JUST TO--.

HAD ONE ERROR.

OK, OK.

THIS IS INTERESTING.

20 ERRORS.

OK, SO.

WE SAID WE'D PLACE MAXIMUM RELIANCE,

AND THAT WAS THE TRUE STATE OF AFFAIRS.

IN ANOTHER SITUATION WHERE WE SAID WE'D HAVE MAXIMUM RELIANCE,

WE HAD,

ER, I'M SORRY.

WE SAID WE HAD MAXIMUM RELIANCE;

IT CAME OUT MODERATE,

SO WE LOST MONEY.

THIS TIME WE CAME OUT LOOKING GOOD.

BUT AGAIN,

LET ME SAY THAT WITH 20 ITEMS,

I THINK THAT WE STILL INTEND TO DO SOME,
WE WOULDN'T TEND TO PLACE MAXIMUM RELIANCE, SO,
I THINK WE'RE JUST FOOLING OURSELVES IF IT COMES UP A,
THAT WE'D ACTUALLY DO THAT.
I THINK THAT WE'RE SAYING THAT A OR B IS MODERATE
RELIANCE.
UH, 20 ITEMS.

HA HA HA.

OK, 20 ITEMS TOLD US WE FOUND 4 ERRORS,
YOU SAID, HEY GUYS, WE GOT PROBLEMS HERE,
UH, POPULATION ERROR RATE BETWEEN 12 AND 20%.
THIS IS THE WORST POSSIBLE SITUATION.
THAT'S WHAT WE GUESSED IT WOULD BE,
FINDING 4 ERRORS,
WE'VE REALLY GOT TO DO SOME, SOME,
BEAT THE HECK OUT OF SOME NUMBERS OUT THERE.
AND UH, SOME PROBLEMS YET,
THE PAYOFF LOOKS GOOD TO US BECAUSE THAT'S WHAT WE
WANTED TO FIND.
IN OTHER WORDS,
THAT WAS THE TRUE SITUATION.
UH, SO WE SPENT A FEW DOLLARS,
TRYING TO FIND,
WE SPENT, WHAT, $4,
TRYING TO FIND WHAT, WHAT WE WANTED TO KNOW.
OK, ANOTHER ONE OF 40.
I'M GOING TO BREAK A THOUSAND DOLLARS THIS TIME,
CUMULATIVE PAYOFF.

$1,074 CUMULATIVE PAYOFF, OK.
ON A SAMPLE SIZE FACTOR OF 40,
WE GOT 6 ERRORS,
WHICH TOLD US WE HAD EXTENSIVE SUBSTANTIVE TESTING IS
NECESSARY.
THAT'S WHAT WE GUESSED THE SITUATION WAS.
AND THAT'S REALLY WHAT IT WAS.
SO, WE SPEND A FEW EXTRA DOLLARS,
TO GET THIS THING PLAYED.
UH, AND ONE FOR 60, SO WE HAVE A FAIR REPRESENTATIVE
SAMPLE.

60 ITEMS,
1 ERROR,
WE PLACED MAXIMUM RELIANCE.
UH, THAT WAS OUR TRUE SITUATION,
WE GUESSED IT RIGHT.
WE HAD A PAYOFF OF $88.
PAYOFF DOESN'T LOOKS AS GOOD BECAUSE WE SPENT A LOT OF
MONEY
I'M TRYING TO FIND THE TEST.
UH, 20.

20 ITEMS,
WE HAD THREE ERRORS,
UM, IT SAID THAT WE WOULD EXTENSIVE,
EXTENSIVE SUBSTANTIVE TESTING IS NECESSARY WHEN
ACTUALLY
WE SHOULD HAVE DONE THE MAXIMUM AMOUNT.
SO WE HAD A BAD SITUATION.
WE DIDN'T GUESS RIGHT.
OK.
TRY AND GET A HANDLE ON WHERE WE'VE BEEN,
WHERE WE'RE HEADING.
LET'S TRY AND IDENTIFY WHAT UH,
TYPE OF PROBLEM,
WHAT IT ACTUALLY MEANS IF WE MISS HIGH,
IN OTHER WORDS,
IF THE WORST SITUATION WOULD OCCUR,
IF THE C OR D WOULD OCCUR AND WE GUESSED LOWER;
FROM 20 SAMPLE ITEMS.
IN OTHER WORDS,
WE SHOULD HAVE DONE A LOT OF TESTING AND DIDN'T.
WHICH IS WHAT THE WORST THAT COULD HAPPEN.
UH, WE ONLY HAD ON THE 20 ITEMS THAT WE SELECTED,
1, 2, 3, 4, 5, 6, 7;
OUT OF 7 WE HAD ONE,
LETS DO ANOTHER 20.
...
20 ITEMS,
WE HAD 2 ERRORS,
UH, OK.
WE HAD A PAYOFF OF $96.
OK.
I'M PRETTY PLEASED WITH THE RESULTS THAT WE'VE GOTTEN OUT OF
20 ITEMS.
BUT, BUT UH,
TRYING TO SELL THAT 20 ITEMS IS ENOUGH,
OR TRYING TO TAKE THAT APPROACH, UH,
LETS DO 20 ITEMS AND SEE IF THE TREND CONTINUES.
...
20 ITEMS,
NO ERRORS.
ONE MORE.
...
UH, OK.
WE GUESSED MAXIMUM RELIANCE,
UH, IT CAME OUT MODERATE RELIANCE;
SO WE LOST $4.
BUT AGAIN,
THAT'S MY SITUATION WHERE I SAY WE REALLY DIDN'T INTEND
MAXIMUM RELIANCE,
EVEN THOUGH THE THING WOULD HAVE TOLD US THAT.
SO 20 ITEMS GIVES YOU AN IDEA OF WHERE YOU'RE HEADING AND
WHETHER YOU NEED MAXIMUM OR MODERATE RELIANCE.
UH,
WE'VE HAD PRETTY GOOD PAYOFFS WITH THAT.
60 ITEMS TEND TO BE PRETTY COSTLY.
Uh, we didn't really seem to do that much better.
We had two,
out of three situations,
we had two situations where um, we missed it by one,
out of 60 items.
40 items;
uh, out of 10 items on 40 we've only had uh,
1 situation,
2 situations where we've missed it.
Uh, ok, just compare 20 to 40.
Out of 10 items on the 20 sample size,
we had 4 that we didn't guess right, our of 40, em,
10 on the 40 sample size we had,
what'd I say,
1, 2 that we didn't guess right.
2 out of 40.
Um, ok.
So I guess,
I don't know how far you want to go but uh,
I can tell you where I'm heading;
If I want,
If I see a situation where I could easily do 7000 items,
I can't really determine how big that is in materiality,
and enter in,
that's a fairly unlimited sample size.
But yet,
we have good controls,
somebody reviews it,
so, with that sort of review,
it's going to be relatively easy for me to do a substantive test.
I'm going to do 20 items to determine whether I need moderate or maximum.
If I want to go a step further,
maybe there's 100,000 items out there;
20 items, I can,
I don't think that's going to give me a very good sample.
Maybe I want to go with 40, uh,
or even if we had 7000 items,
40 gives me much more-
and I might start to rely on the maximum reliance.
I might tend to do a detail test and an analytical review.
With 40 items.
But again,
I think I'm in a gray area,
I think that if I want to go that maximum reliance,
I'd want to get a much more larger sample.
WITH A SAMPLE UH, UH,
A CLIENT THIS SIZE,
WITH ABOUT $8,000,000 ASSETS,
$15,000,000 IN REVENUE;
IT'S NOT REALLY THAT LARGE TO TRY AND SPEND A LARGE
DOLLAR AMT.
IN COMPLIANCE TESTING.
SO I DON'T THINK WE NEED TO GET 100 OR 200 ITEMS OR
WHATEVER,
I THINK WHERE WE'RE AT, 20, 40, 60 GIVES US AN
ADEQUATE FEELING OF
THE COMPLIANCE TESTING,
BUT YET,
WE HAVE A TRADEOFF, BECAUSE OUR SUBSTANTIVE TESTING IS
RELATIVELY
EASY TO DO,
THE DOLLAR AMOUNTS AREN'T THAT LARGE,
I THINK 50% OF THEM WERE INVENTORY.
INVENTORY, WE WANT TO LOOK AT THAT AREA ANYWAY.
TO SUM UP,
I GUESS I'M SAYING IF I SEE A SITUATION WHERE I WANT
MODERATE
RELIANCE,
I WANT TO HAVE A BASIS FOR MODERATE RELIANCE,
AND DO LIMITED AMOUNT OF TESTING,
AND IT'S NOT A CRUCIAL AREA;
I THINK I'D BE HAPPY,
BASED UPON THESE RESULTS,
WITH 20.
YOU ARE THE SENIOR IN CHARGE OF THE AUDIT OF ROHR INDUSTRIES,
A MANUFACTURER OF CABINETS FOR ELECTRONIC EQUIPMENT.
SUBJECT G CONTINUES READING
HERE IS A SAMPLE OF 100 ERROR RATES GENERATED BY THE
COMPUTER.
(HOW WHAT ARE YOU DOING AS YOU ARE LOOKING AT THEM?)
I'M READING THEM.
ARE YOU SIMPLY READING THE NUMBERS OR WHAT?)
RIGHT NOW I'M READING THE NUMBERS,
THEN I INTEND TO FIND OUT IF ANY OF THEM ARE
RECURRING...
THEY'RE PRETTY MUCH EVERYWHERE.
OK, NOW WHAT?
(WELL YOUR TASK IN THE EXPERIMENT IS TO CHOOSE THE
SAMPLE SIZE.)
ARE YOU AWARE THAT YOU HAVE 2 APPENDICES)
UH HUH.
NOW I'M LOOKING AT THE FINANCIAL STATEMENTS.
(REMEMBER TO VERBALIZE)
SORRY, RIGHT NOW I'M JUST GLANCING.
JUST TO GET AN OVERALL FEEL FOR HOW,
WHAT THE SIZE OF THE COMPANY IS.
$8,000,000 IN ASSETS,
$15,000,000 IN REVENUE.
NOW I'M LOOKING AT THE FLOWCHARTS.
THIS SHOULD TAKE A WHILE BECAUSE THEY'RE NOT THE SAME
AS
I'M USED TO SEEING.
BACK TO THE INSTRUCTIONS.
(TRY TO MAKE AN EFFORT TO VERBALIZE.)
I DON'T KNOW WHAT YOU WANT ME TO SAY.
(OK, WELL, LET ME TURN THAT OFF FOR A SECOND)
WELL, MY THOUGHT PROCESS IS THAT I NEED TO FIND OUT,
BEFORE I CAN MAKE A SAMPLE SIZE DECISION,
HOW MUCH RELIANCE I WANT TO PLACE ON THE SYSTEM,
AND WHAT KIND OF ERRORS I EXPECT TO START WITH.
BASED ON THAT,
THEN I CAN DETERMINE HOW MANY ITEMS I NEED TO SELECT.
GIVEN THAT,
OBVIOUSLY IF I WANT TO PLACE A HIGH RELIANCE ON THE
SYSTEM
AND I EXPECT A LOT OF ERRORS,
THEN I GONNA CHOOSE A LOT HIGHER SAMPLE SIZE.
IT WORKS THE OTHER WAY TOG.
IF I DON'T EXPECT VERY MANY ERRORS,
THEN THE SAMPLE SIZE IS GOING TO BE LOW.
SO HOPEFULLY BY LOOKING AT THE FLOWCHARTS,
I'LL BE ABLE TO FIND OUT WHAT KIND OF CONTROLS THERE
ARE.
That's where the problem is because I'm not used to these flowcharts.

Are these departments engineering, purchasing, receiving?

(Um hm.)

Do you want me to read these too?

(Well, whatever comes to mind. Feel like saying what you're doing, just think aloud.)

OK, I'm just following right now.

The order system.

So I'm just following the process of the order through the system.

Now I'm reading the memo in inventory control.

When an order is taken the engineering department is given a copy of the order to determine the specs.

They then forward the materials requirements to the inventory control department head.

Based on EOQ, materials needed, inventory control determines the need to order materials.

When the determination is made that an item should be ordered, the inventory control department forwards to the purchasing agent a requirements list from which he is to make purchase orders, this is done daily.

Back to the flowchart.

Now I'm in the purchasing department.

I'm following the purchase order through.

Cost accountant determines the cost distribution and the determination of any variance.

I still have a little bit of trouble following this.

The invoice is sent to the purchasing agent, he reviews it.

OK. Receiving matches the receiving report and the purchase order.

Before that's sent back to purchasing, they get the invoices.

That's when the purchase agent reviews it.

What is PA?

(Purchasing agent.)

OK. Purchasing agent approval.

Stock purchases and supplies, purchasing agent approval.

OK, accounts payable and cash disbursements flowchart.

Right now I'm just going to read the processes because I can't follow the flowcharts real well.

...this will tell me all I need to know.

Upon approval for payment by the purchasing agent,
THE INVOICE AND PURCHASE ORDER ARE FORWARDED TO THE COST 
ACCOUNTANT WHO STAMPS THE INVOICE WITH A PAYMENT STAMP.
AND ENTERS THE INVENTORY ACCOUNT CODE.
HE THEN GIVES THE INVOICE TO THE ACCOUNTING CLERK FOR PAYMENT.
THE ACCOUNTING CLERK TAKES ANY DISCOUNT AVAILABLE,
RUNS A TAPE TO DETERMINE ACCURACY OF INVOICE SUMMARY
AND HAS PAYMENTS APPROVED BY THE CONTROLLER.
THE ACCOUNTING CLERK RECEIVES THE CHECKS,
ATTACHES THE REMITTANCE ADVICES,
AND HAS THEM SIGNED BEFORE SENDING THEM OUT.
NCR OPERATOR POSTS THE GROSS AMOUNT AND CHECK NUMBERS INTO
THE CHECK REGISTERS.
CONTROLLER APPROVES ALL CASH DISBURSEMENTS.
AUTHORIZED SIGNERS ARE THE PRESIDENT AND THE CONTROLLER.
DOES THAT MEAN DUAL, OR JUST THE CONTROLLER?
(UM, EITHER ONE)
OK.
DISTRIBUTION TICKETS ARE KEPT FOR EACH ACCOUNT AND POSTED
AT THE SAME TIME AS THE VOUCHER REGISTER.
A TAPE IS MADE AT THE END OF THE MONTH FOR EACH ACCOUNT WHICH
IS USED FOR POSTING TO THE GENERAL LEDGER.
OK,
THAT'S THAT.
SO I KNOW THAT THERE'S 7000 TRANSACTIONS.
THAT'S PROPER, OK.
AND NOWHERE IN HERE DOES IT SAY THE EXTENT OF RELIANCE THAT
YOU PLAN ON PLACING.
(READ THE TOP OF PAGE 2).
THAT'S AFTER SAMPLE SIZE,
OK,
BUT BEFORE SAMPLE SIZES,
WE DON'T KNOW HOW MUCH RELIANCE WE PLAN ON PUTTING ON THE
SYSTEM.
SO THAT'S A DETERMINATION THAT NEEDS TO BE MADE.
WE ALSO DON'T KNOW THE SIZE,
DOLLAR VALUE OF THE SAMPLES, ITEMS.
(NO)
(I'M NOT SAYING THAT YOU HAVE TO GUESS OR YOU HAVE TO LOOK)
RIGHT.
RANDOM SAMPLE, SO IT'S NOT STATISTICALLY BASED,
IS THAT A FAIR ASSUMPTION?
(I DON'T KNOW WHAT YOU MEAN BY STATISTICALLY BASED)
WELL, IF YOU HAVE A STATISTICAL SAMPLE,
VERSUS A NON-STATISTICAL SAMPLE,
There's going to be a difference, if you can go in there and take a statistical sample.

(you mean a stratified sample?)

Stratified and statistical.

You can stratify a non-statistical sample too.

(give me an example of what you mean by a non-statistical sample)

OK.

Um, non-statistical;

If I just go in there and,

they have a file of invoices,

and I go in and just start pulling them out,

and say I want 50 of them.

That's a random sample, you call that non-statistical.

That's that's, right.

OK, I call that non-statistical, right.

If I say I want to go in there and pull out every 6th item,

starting with the 3rd.

You know,

something that's going to do that,

or every, every thousandth dollar, then it's statistical.

(Ok)

But we're saying it's a non-statistical, random sample,

just go in there and pull them out.

OK.

So I would determine that from looking at the flowcharts,

Um, I would probably go in at first and say that we're going to place maximum reliance on the system.

There's a lot of segregation of duties,

and a lot of double checking,

so that would be my first assumption before I started picking a sample size.

Then, uh,

Costs and expenses are $15,000,000.

Obviously that doesn't tell me anything about purchases that are still in inventory, but,

so there's $15,000,000 worth,

and there's 7000 items.

If I was compliance testing,

first thing I would do is go to our standard form that tells me,

that's why this is difficult.

(Oh, do you want to use a sample size table)

I mean we have a sample size table that tells us.

(well I can)

And I know that it would probably tell me to pick 20 items.

Yeah.
(OK, SO THIS IS YOUR FIRST SAMPLE REPORT. WHAT IT SAYS;

IT SAYS THE SAMPLE SIZE YOU PICKED WAS 20, YOU DID NOT DISCOVER ANY BAD TRANSACTIONS, AS A CONSEQUENCE YOU INFERRED THAT THE TRUE STATE OF INTERNAL CONTROLS WAS A, BUT IN FACT IT WAS B. SO ACCORDING TO THE PAYOFF SYSTEM YOU LOST $4.)

OK.

(OK, NOW IF YOU'RE READING PAGE 3 OF THE EXPERIMENT, SEE THE LITTLE PARAGRAPH THAT SAYS THAT, WELL YOU CAN READ IT.)

YOU HAVE THE OPPORTUNITY TO LEARN FROM EXPERIENCE BECAUSE THE EXPERIMENT WILL BE REPEATED A LARGE NUMBER OF TIMES.

FOR EACH REPEAT OF THE EXPERIMENT, ASSUME YOU ARE IN CHARGE OF THE AUDIT OF ANOTHER COMPANY,

BUT EACH NEW COMPANY IS IDENTICAL IN EVERY RESPECT TO BOHR

EXCEPT FOR THE NUMBER OF CASH DISBURSEMENTS IN THE POPULATION.

7000,

WHICH DO NOT COMPLY WITH THE STATED INTERNAL CONTROL PROCEDURES.

THIS NUMBER IS DETERMINED AT RANDOM BY COMPUTER IN A MANNER WHICH IS INTENDED TO REFLECT THE FREQUENCY OF ERROR RATES OBSERVED OVER ALL CLIENTS AUDITED BY BIG 8 CPA FIRMS WITH SIMILAR CHARACTERISTICS.

BECAUSE OF YOUR AUDITING EXPERIENCE,

YOU PROBABLY HAVE A PRETTY FAIR CONCEPTUALIZATION OF ERROR RATE FREQUENCY.

BUT TO MAKE THE COMPUTER GENERATION PROCESS MORE EXPLICIT TO YOU,

HERE IS A SAMPLE OF 100 ERROR RATES AND PERCENTAGES GENERATED BY THE COMPUTER.

(OK, WELL THE BASIC POINT IS THAT WE ARE NOW ON ANOTHER CLIENT).

OK.

(WHICH IS IDENTICAL IN EVERY RESPECT TO BOHR, EXCEPT FOR THE NUMBER OF TRANSACTIONS IN THE POPULATION WHICH ARE DEFECTIVE,

SO PICK A SAMPLE)

(SO YOUR TASK IS TO PICK ANOTHER SAMPLE SIZE)

AND YOU HAVE NO INFORMATION ON IT?
(WELL, EVERYTHING THAT YOU'VE HAD BEFORE, AND IN ADDITION YOU
HAVE THAT PARTICULAR PIECE OF INFORMATION)
30.
(OK NOW, ON THIS CASE YOU HAD A SAMPLE SIZE OF 30, THE SAMPLE
OUTCOME WAS 2, YOUR MANAGER INFERRED STATE C, WHEN IN FACT IT
WAS STATE B.)
OK.
GLAD I'M NOT A GAMBLER.
OK, NOW, SO YOU'RE GONNA JUST SAY THAT'S THE SITUATION AND
WE'RE JUST GONNA KEEP ON MOVING SAMPLE SIZES.
(WELL, YEAH)
I GUESS I JUST DON'T UNDERSTAND THIS WHOLE,
I MEAN I COULD JUST SIT HERE AND BATTLE OFF 25, 40, 60, 90,
70, 50, I MEAN,
BUT WHERE DOES IT GET YOU IN THE END.
(WELL, THE OBJECT OF THE EXPERIMENT IS FOR YOU TO DO AS WELL
AS YOU CAN, MAXIMIZING YOUR PAYOFF.)
OH.
OK.
25.
SUBJECT H

1. YOU ARE THE SENIOR IN CHARGE OF THE AUDIT OF ROHR INDUSTRIES, A MANUFACTURER OF CABINETS FOR ELECTRONIC EQUIPMENT.

2. SUBJECT H IS READING EXPERIMENT THE RESULTS OF YOUR INQUIRIES AND OBSERVATIONS ARE SUMMARIZED IN THE FLOWCHARTS AND MEMORANDUM CONTAINED IN APPENDIX B.

3. AT THIS POINT I BETTER TAKE A LOOK AT THE STATEMENTS. APPENDIX A AND UH, COMPARATIVE FINANCIAL STATEMENTS. AN $8,000,000 ASSETS, THAT'S THE TOTAL.

4. LOT OF RECEIVABLES AND INVENTORIES, NOT A GREAT DEAL OF PLANT AND EQUIPMENT. WE KNOW CABINETS FOR ELECTRONIC EQUIPMENT.

5. QUITE A BIT OF LONG TERM DEBT. SOME SUBSTANTIAL RETAINED EARNINGS.

6. CURRENT RATIO LOOKS REAL GOOD.

7. CURRENT ASSETS WHICH INCLUDE INVENTORIES AND RECEIVABLES WOULD COVER ALL OF THEIR DEBT.

8. GETS BACK TO THAT KIND OF SMALL PLANT & EQUIPMENT. NET INCOME WAY DOWN IN THE PAST YEAR. RATHER SMALL COMPARED TO THEIR RETAINED EARNINGS.

9. LOOKS LIKE THEY REALLY HAD A HARD TIME MEETING THEIR MARGINS THIS PAST YEAR.

10. MINOR EXTRAORDINARY ITEM, VERY MINOR.

11. EPS, WAY DOWN.

12. DIVIDENDS RATHER LEVEL, NOT VERY LARGE.

13. IN '80 CHANGE TO LIFO.

14. WHEN WE GET TO NOTE 3 IT'LL PROBABLY TELL ME WHETHER THEY RESTATE THE PRIOR CHANGE IN LIFO.

15. COULD HAVE MADE A DIFFERENCE IN THE INCOME.

16. DEPRECIATION LOOKS RATHER STANDARD.

17. THERE'S BEEN SOME RESTATEMENT OF INVENTORY. IT DOESN'T LOOK LIKE IT WAS THAT SIGNIFICANT, $45,000.

18. EFFECTIVE CHANGE TO LIFO IS DISCLOSED HERE.

19. THAT HAD SOME OF THE EFFECT ON THE DECREASE IN NET INCOME.

20. BUT APPARENTLY NOT THAT SIGNIFICANTLY.

21. LONG TERM DEBT HAS A LONG WAY TO GO.

22. INCOME TAX ...(UNCLEAR)

23. I DON'T SEE THE KIND OF NUMBERS I'D BE LOOKING FOR.

24. WE HAD APPENDIX B,

25. WHICH WAS SOME EVALUATIONS.

26. OBSERVATIONS, EVALUATIONS,
RIGHT INTO PURCHASES AND PAYABLES.
INVENTORY CONTROL MEMO.
WHEN AN ORDER IS TAKEN THE ENGINEERING DEPARTMENT IS
GIVEN A
COPY OF THE ORDER TO DETERMINE THE SPECS.
THEY THEN FORWARD THE MATERIALS REQUIREMENTS TO THE
INVENTORY
CONTROL DEPARTMENT HEAD.
BASED ON THE ECONOMIC ORDER QUANTITY AND MATERIAL
NEEDS,
INVENTORY CONTROL DETERMINES THE NEED TO ORDER
MATERIALS.
WHEN THE DETERMINATION IS MADE THAT AN ITEM SHOULD BE
ORDERED,
INVENTORY CONTROL DEPARTMENT FORWARDS TO THE
PURCHASING AGENT,
THE REQUIREMENTS LIST FROM WHICH HE IS TO MAKE
PURCHASE ORDERS.
THIS IS DONE DAILY.
IT DOESN'T SAY WHETHER THE TIME THEY NEED IT TO COME
IN IS
LISTED THERE,
BUT I WOULD ASSUME THAT.
SO IT SOUNDS LIKE THEY GOT A PRETTY GOOD HANDLE ON
THAT.
FLOWCHARTS, PURCHASING.
GOT SOME RATHER LOW LIMITS ON PURCHASE ORDER, DOLLAR
LIMITS
ON PURCHASE ORDER APPROVALS.
I'M TRYING TO FIGURE OUT WHAT THE SYMBOLS ARE HERE.
GOT SOMETHING HERE SOMEWHERE.
MUST BE THE OTHER ONE.
FILE ON LINE.
DOES THAT MEAN COMPUTER FILE? OR ..
(WHICH ONE)
THERE'S AN OFF-LINE FILE AND AN ON-LINE FILE.
OR DOES IT JUST MEAN THEY FILE SOMETHING AT THAT
POINT.
(THE IS A COMPUTER FILE.)
OK.
PURCHASE ORDER.
IN THE PURCHASING DEPARTMENT, THEY KEEP A COPY OF IT.
SEND A BUNCH OF THEM ON TO RECEIVING.
THEY'VE ALSO GOT ONE THAT'S FILED HERE.
LET ME SEE.
THE RECEIVING DEPARTMENT CHECKS TO MAKE SURE THEY WERE
SUPPOSED
TO BE RECEIVING SOMETHING.
CHECKING THE PURCHASE ORDER.
CHECK A COPY OF THE PO AND THE RECEIVING REPORT,
AND SEND IT BACK TO PURCHASING.
QUANTITIES ARE CHECKED THERE.
THEY ALSO CHECK IT TO THEIR OWN COPY OF THE PO.
REVIEW ON MASTER RECEIVER FILE EVERY MONTH.
223

86 WHICH IS A GOOD THING,
87 BECAUSE OTHERWISE, YOU'D END UP WITH STUFF IN THERE THAT
88 YOU'VE GOT ON HAND,
89 THAT MIGHT NOT BE IN YOUR INVENTORY.
90 AND IT SHOULD BE IN PAYABLES.
91 CAN'T TELL HERE WHETHER IT'S INVENTORY OR NOT.
92 UH,
93 PURCHASING AGENT REVIEWS THE INVOICE
94 AFTER WE'VE MATCHED UP THE RECEIVER AND THE PURCHASING ORDER,
95 SAME GUY THAT'S ORDERING THE STUFF IS CHECKING TO MAKE SURE
96 THAT THEY GOT WHAT THEY ORDERED.
97 WHICH ISN'T QUITE WHAT IT SHOULD BE.
98 ACCOUNTS PAYABLE AND CASH DISBURSEMENTS;
99 THIS COMES ON PAGE ONE.
100 PURCHASING AGENT MATCHED ALL THAT UP.
101 ENGINEERING APPROVED.
102 I'M NOT SURE BECAUSE OVER HERE...
103 I GUESS THEY'RE RUNNING IT BACK THROUGH THE PURCHASING AGENT
104 TO MAKE SURE THAT THE SPECS ARE RIGHT
105 SOMETHING THAT HE WOULD RECOGNIZE THAT MAYBE ACCOUNTING WOULDN'T.
106 FIRST TO ACCOUNTING.
107 APPARENTLY THE RECEIVER IS NOT SENT ON TO ACCOUNTING.
108 SO THEY DON'T KNOW WHETHER THEY GOT THE STUFF OR NOT.
109 THEY JUST GET THE PO AND THE INVOICE.
110 OK.
111 CONTROLLER REVIEWS THE DOCUMENTS AND AUTHORIZES THE CASH
112 DISBURSEMENTS.
113 NCR OPERATOR WOULD ...
114 (UNCLEAR)
115 OH, IS THAT ALL IT IS.
116 HERE WE GOT AN INDICATION THAT THE INVOICE, THE PURCHASE ORDER,
117 AND THE BB,
118 APPARENTLY RECEIVING REPORT ARE ALL TOGETHER THERE.
119 IT ONLY SHOWS 2 DOCUMENTS,
120 BUT APPARENTLY THE RECEIVING REPORT IS THERE WITH IT.
121 OK, SO ACCOUNTING DOES KNOW THAT THEY ACTUALLY RECEIVED SOMETHING.
122 FILED UNDER BY DATE
123 PARTICULAR FILES SO THEY KNOW WHEN TO PAY THE STUFF,
124 I ASSUME.
125 UM,
126 LET'S SEE,
127 (UNCLEAR)...CHECKS, THEY RUN THEM THROUGH THE NCR MACHINE.
128 WHICH TAKES THEM BACK OFF ACCOUNTS PAYABLE.
APPARENTLY, THE POSTING THERE DISTRIBUTES IT TO THE DIFFERENT REPORTS.
THE ARROWS ARE COMING FROM A COUPLE DIFFERENT WAYS INTO HERE.
OK,
WHEN THEY POST THE PAYROLL IN THE FIRST PLACE,
IT GOES TO THE DIFFERENT EXPENSES,
INVENTORY OR WHATEVER.
WHEN THEY RUN THE CHECK THROUGH,
THEY TAKE IT BACK OUT OF PAYABLES,
WHICH COULD BE A COUPLE WEEKS LATER.
(REMEMBER TO VERBALIZE)
YEAH, I'M TRYING TO FIGURE OUT EXACTLY WHAT'S HAPPENING HERE.
THE UH,
THEY'RE DOING SOME KIND OF A MONTH-END RECONCILIATION OF THE CHECKS THAT HAVE BEEN WRITTEN,
I THINK.
THE DISTRIBUTION TICKET, I'M NOT QUITE SURE WHAT IT IS.
UH,
IS THAT WHAT THEY USE TO POST TO THE DIFFERENT EXPENSE OR ASSET ACCOUNTS.
(THAT'S RIGHT)
OK,
AND THEN, THOSE ARE ACTUALLY KEPT,
POSTED,
THEY RUN A TAPE AT THE END OF THE MONTH FOR EACH ACCOUNT TO MAKE SURE THAT WHAT GOT POSTED IS WHAT THEY HAVE ON HAND.
CHECKING THE TOTALS.
OK.
THEY CHECK THE GRAND TOTAL OF DISBURSEMENTS INTO THE CHECK REGISTER.
THROUGH THE GENERAL LEDGER.
OK.
UM.
YOUR MANAGER HAS EXAMINED ALL OF THE ABOVE INFORMATION AND HAS INSTRUCTED YOU TO CHOOSE A RANDOM SAMPLE OF CASH DISBURSEMENTS AND TO TRACE THESE TRANSACTIONS BACK THROUGH THE SYSTEM AND CLASSIFY EACH TRANSACTION AS EITHER COMPLYING OR NOT COMPLYING WITH THE INTERNAL CONTROL REQUIREMENTS FOR THAT TRANSACTION.
Non-compliance is defined as failing to adhere to at least one procedure in the cycle.

So apparently, we're testing for all of the compliance with each procedure, all the way along, rather than one or two particular controls.

As the subject in this experiment, your only task is to choose the sample size for the compliance test.

The actual test of compliance will be simulated by the computer program run by the monitor of the experiment.

The experiment will proceed by first having you choose the sample size.

The monitor will then inform you of the following information:

- The sample outcome,
- Error rate,
- Inference made by your manager based on the sample size,
- One of 4 inferences.

So I need to choose a sample size,

I'm going to go ahead and read this.

Inference by the manager based on the sample size.

Population error rate is less than or = to 3%.

That can be relied on,

is between 3 and 7.

Satisfactory, maybe you'll do more substantive testing.

Population error rate is between 7 and 12.

Significant errors,

More extensive substantive,

and population error rate between 12 and 20.

And maximum substantive testing,

which means the compliance testing wasn't relied upon at all.

So I need to pick a sample size.

Apparently I've read far enough to know everything that is.

Gonna go into the sample size decision.

Uh,

Manager's decision pertaining to the overall assessment of the degree to which the firm can rely on the existing internal controls for substantive testing is based on his general knowledge of the company and its environment.
THE NARRATIVE AND FLOWCHART OF INTERNAL CONTROL PROCEDURES
AND THE RESULTS OF YOUR SAMPLE
THE COSTS INCURRED TO THE FIRM RESULTING FROM ANY ERROR IN
JUDGMENT.
YOU MIGHT MAKE BY OVERESTIMATING OR UNDERESTIMATING THE TRUE
AMOUNT OF INTERNAL CONTROL PROCEDURES WHICH CAN BE RELIED UPON.
THE TRUE STATE OF INTERNAL CONTROLS BECOMES KNOWN SOME TIME
AFTER COMPLETION OF THE AUDIT.
YOUR NET PAYOFF.
THE FOLLOWING SYSTEM OF REWARDS AND PENALTIES WILL ACCRUE TO YOU.
IN ALL CASES YOU WILL BE ASSESSED 20 CENTS PER TRANSACTION
SAMPLED.
YOU WILL WIN $100 IF THE MANAGER'S DECISION CORRESPONDS WITH THE
TRUE STATE OF INTERNAL CONTROLS.
LOSE $4000 IF THE MANAGER ERRS A OR B AND THE TRUE STATE OF
INTERNAL CONTROLS IS D.
WHICH WOULD BE NO RELIANCE.
IF THE MANAGER ERRS BY ONE LEVEL IN EITHER DIRECTION,
NO REWARD OR PENALTY.
A 2 LEVEL ERROR OTHER THAN THE ERROR DISCUSSED PREVIOUSLY WILL
DECREASE THE EARNINGS BY $50.
A 3 LEVEL ERROR OTHER THAN THE ERROR DISCUSSED PREVIOUSLY WILL
CAUSE YOUR WINNINGS TO BE DECREASED BY $100.
NATURALLY, THE LARGER THE SAMPLE SIZE YOU CHOOSE THE GREATER
THE LIKELIHOOD THAT YOUR MANAGER WILL HAVE OF MAKING THE CORRECT
INFERIENCE.
YOU WILL NEED TO BALANCE THIS INTUITIVE STATISTICAL PROPERTY
AGAINST THE COST OF SAMPLING.
YOU HAVE THE OPPORTUNITY TO LEARN FROM EXPERIENCE BECAUSE THE
EXPERIMENT WILL BE REPEATED A LARGE NUMBER OF TIMES.
FOR EACH REPEAT OF THE EXPERIMENT ASSUME YOU ARE IN CHARGE OF
AN AUDIT OF ANOTHER COMPANY,
BUT EACH NEW COMPANY IS IDENTICAL IN EVERY RESPECT TO ROHR,
EXCEPT FOR THE NUMBER OF CASH DISBURSEMENTS IN THE POPULATION,
7000.
242 WHICH DO NOT COMPLY WITH THE STATED INTERNAL CONTROL PROCEDURES.
243 THIS NUMBER IS DETERMINED AT RANDOM BY COMPUTER IN A MANNER WHICH
244 IS INTENDED TO REFLECT THE FREQUENCY OF ERROR RATES OBSERVED
245 IN ALL CLIENTS OBSERVED BY "BIG 8 CPA FIRMS WITH SIMILAR
246 CHARACTERISTICS.
247 BECAUSE OF YOUR AUDITING EXPERIENCE,
248 YOU PROBABLY HAVE A PRETTY FAIR CONCEPTUALIZATION OF ERROR RATE
249 FREQUENCIES, BUT TO MAKE THE COMPUTER GENERALIZATION PROCESS
250 MORE EXPLICIT TO YOU, HERE IS A SAMPLE OF 100 ERROR RATES AND
251 PERCENTAGES GENERATED BY THE COMPUTER.
252 SO THAT'S APPARENTLY THE TYPICAL RANGE OF ERRORS AND PERCENTS.
253 WE GOT A RANGE FROM ALMOST NOTHING UP TO A FEW ABOVE 10.
254 THERE'S A COUPLE 11'S,
255 A 19, ALMOST 20.
256 MUST OF BEEN 8 OR LESS.
257 INITIAL SAMPLE SIZE,
258 SAY 40 ITEMS.
259 ... AT THIS POINT YOU'RE GOING TO TELL ME SOMETHING.
260 (YES)
261 (OK, THESE ARE THE RESULTS, YOU PICKED A SAMPLE SIZE OF 40,
262 THE PROCESS DETECTED ONE ERROR, YOUR MANAGER INFERRED THAT THE
263 TRUE STATE OF INTERNAL CONTROLS WAS A, IN FACT IT WAS 8,
264 SO ACCORDING TO THE PAYOFF SCHEME, YOU LOST $8.)
265 SAMPLE OUTCOME OF A.)
266 OK.
267 I'M TRYING TO FIGURE OUT EXACTLY WHAT HAPPENED HERE.
268 THERE IS APPARENTLY,
269 IF APPARENTLY MY SAMPLE SIZE,
270 IF IT HAD BEEN LARGER,
271 WOULD HAVE DISCOVERED THERE WAS A LITTLE BIT HIGHER ERROR RATE.
272 OK.
273 NOW WHAT HAPPENS AT THIS POINT?
274 (WELL, YOU HAVE A NEW CLIENT, IDENTICAL IN EVERY RESPECT TO ROHR,
275 WITH THE ONLY EXCEPTION THAT THE POPULATION ERROR RATE HAS
276 CHANGED)
277 OK, SAME POPULATION, IT'S JUST THAT THE NUMBER OF ERRORS
278 IN THERE IS GOING TO CHANGE.
AND WE WANT TO KNOW, LET ME READ THROUGH THIS NEXT PAYOFF
THINK A LITTLE BIT AGAIN,
SO I CAN SEE HOW THAT HAPPENED.
PAYOFF ON THIS TRIAL WAS DUE TO THE TRANSACTION CHARGE AND
BEING WITHIN ONE UH, OK.
GO WITH 40 AGAIN.
---
SAMPLE SIZE 40,
2 ERRORS
DECISION C,
BETWEEN 7 AND 12.
ACTUAL IS B.
OK, AND,
PAYOFF AT THIS POINT IS SAME AS THE OTHER ONE,
AND THAT'S BECAUSE I'M WITHIN 1 ON THE DECISION, AND,
AND THEY CHARGE FOR THE NUMBER OF ITEMS TESTED.
AND THE POINT HERE IS TO BE AS EFFICIENT AS YOU CAN.
AND COME OUT REASONABLY CLOSE TO ACTUAL.
LET'S GO WITH 40 AGAIN.
---
THIS COULD GET REAL INTERESTING IF I MISS IT BY A WHOLE BUNCH.
(WELL)
DID THAT JUST HAPPEN?
(NO)
SAMPLE OUTCOME 4 ERRORS, WOW.
DECISION C.
ACTUAL WAS B.
I END UP ABOUT THE SAME AS WE DID BEFORE.
LET'S TRY 60 ITEMS.
(THE PAPER GOT KINDA STUCK HERE)
CAN YOU READ IT?
(YEAH, I CAN READ IT BUT THAT'S ALL THAT CAME OUT)
THE PRINT ON TOP OF IT.
(I WILL READ IT TO YOU)
THAT'S PRETTY POOR PAPER ISN'T IT.
(WELL, YEAH. IT'S A POOR TERMINAL.)
(YOUR SAMPLE OUTCOME WAS 60. I MEAN YOUR SAMPLE SIZE WAS 60.
SAMPLE OUTCOME WAS 0, THE MANAGER INFERRED A, THE TRUE STATE
WAS A. AND SO YOU WON ON THAT TRIAL, $88, FOR A CUMULATIVE
TOTAL OF $64.)
OK.
NOW WE'RE GONNA HIT IT AGAIN.
(UH, 60 ITEMS IS GIVING ME A MORE ACURATE,
AT LEAST IN THIS CASE,
OF COURSE WITH ZERO ERRORS,
YOU'RE GOING TO COME OUT WITH THAT INFERRENCE IN ANY CASE.
UH, LET'S GO WITH 60 AGAIN.
...  
2 ERRORS,  
A, ACTUAL A.  
WE'RE IN GOOD SHAPE HERE.  
WE SEEM TO BE COMING MUCH CLOSER WITH A LITTLE BIT  
SAMPLE  
SIZE, UH,  
BUT THAT'S WITH SMALL ERRORS TOO.  
LET'S GO WITH THAT AGAIN.  
IT'D BE INTERESTING TO SEE WHAT HAPPENS WHEN IT COMES  
OUT TO  
BE 4 OR 5 ERRORS.  
...  
4 ERRORS,  
DECISION 7 TO 12,  
AND THE ACTUAL IS BETTER THAN THAT.  
WHICH IS WITHIN ONE.  
LOOKS GOOD.  
GO WITH 60 AGAIN.  
...  
6 ERRORS, WOW.  
STILL PICKED UP  
NO WORSE THAN ACTUAL.  
PICKED A 7 TO 12,  
AND IT WAS REALLY 3 TO 7.  
SO WE'VE GOT NO RISK THERE.  
OK.  
GO WITH 60 AGAIN.  
...  
NO ERRORS.  
A.  
ACTUAL HAPPENED TO BE A.  
WHICH WOULD PROBABLY COULD HAVE TOLD WITH 40 ITEMS  
BUT 60 GIVES YA A LOT BETTER SHOT AT IT THERE.  
3%,  
LET'S GO BACK TO 40 AND SEE WHAT HAPPENS.  
...  
I BET YOU'RE TIRED OF TEARING AFTER DOING THIS FOR,  
WHAT, YOU  
SAID TEN PEOPLE.  
(HA HA)  
40,  
ONE ERROR.  
ACCEPTED THE MAX AND THAT'S WHAT IT ACTUALLY WAS.  
OK.  
GO WITH 40 AGAIN.  
I REALLY DON'T WANT TO GO ANY LOWER THAN THAT  
BECAUSE IF YOU GET MUCH LOWER,  
THERE'S A CHANCE OF GETTING BLOWN TOTALLY OUT OF THE  
WATER,  
IF YOU DON'T HAVE ANY ERRORS,  
JUST IF YOU HAPPENED TO PICK 20 GOOD ONES,  
The actual could be pretty bad.
There's a lot more risk there of accepting a population that's no good.

---

At that point we came up with a 7 to 12, and the actual was a little better than that, so there's no big problem there.

Okay.

Go with 40 again.

---

Okay, no errors, except for the A, which is natural.

Actual was B.

At this point we don't seem to be getting anywhere again.

Uh, try 50.

The reason being, with 40 we're usually missing it by 1.

There's not much of a payoff that way.

No penalty, but no payoff either.

---

2 errors.

Came up with a 3 to 7, and that was the actual.

Got a payoff.

Go with that again, 50.

---

8 errors, wow.

Uh, our decision is 12 to 20.

Don't accept anything at all.

Actual was a little bit better than that.

Try 50 again.

---

(What are you calculating?)

16%, I don't need the calculator.

That's a lot of errors.

That's almost as high as these 100 error rates.

One of them was 19.

Okay.

Got one.

Hit it right on.

With the 50 we got 8 errors and that's 16%.

And that's about as high as we're ever going to get.

And yet we were only one off.

In fact we accepted a little worse than what the case was when actually we ended up not relying.

When we could have.

Let's go back to 60.

---

No errors and we hit it right on.

Okay, go with 60 again.
426 ...  
427 OK,  
428 4 OUT OF 60,  
429 OUR DECISION IS ON THE CONSERVATIVE SIDE  
430 FROM THE ACTUAL.  
431 LET'S GO TO 50.  
432 ...  
433 2 ERRORS.  
434 ACCEPT ONE MORE CONSERVATIVE THAN ACTUAL.  
435 WHICH, AFTER DOING THE AUDIT WOULD BE FINE,  
436 THAT'S THE KIND OF THING WE'RE LOOKING FOR,  
437 BUT THAT'S NOT DOING US ANY GOOD IN THE CUMULATION OF  
438 OUT OF 60,  
439 OUR DECISION IS ON THE CONSERVATIVE SIDE  
440 FROM THE ACTUAL.  
441 LET'S GO TO 50.  
442 ...  
443 2 ERRORS.  
444 ACCEPT ONE MORE CONSERVATIVE THAN ACTUAL.  
445 WHICH, AFTER DOING THE AUDIT WOULD BE FINE,  
446 THAT'S THE KIND OF THING WE'RE LOOKING FOR,  
447 BUT THAT'S NOT DOING US ANY GOOD IN THE CUMULATION OF  
448 POINTS  
449 50 AGAIN.  
450 7 ERRORS,  
451 AND WE HIT IT RIGHT ON,  
452 14%, WHICH IS KINDA HIGH.  
453 YET WE HIT THAT RIGHT ON.  
454 I WONDER IF WE JUST GOT LUCKY.  
455 THAT'S A LOT OF ERRORS TO BE PLAYING WITH AND  
456 COME OUT WITH A, THAT CLOSE.  
457 TRY 50 AGAIN.  
458 ...  
459 6 ERRORS.  
460 WHICH KIND OF INDICATES THAT LAST TIME WE DID GET  
461 LUCKY  
462 BECAUSE WE HAVE THE SAME  
463 EXCEPT FOR THE ACTUAL IS LOWER THAN THAT, BETTER THAN  
464 THAT.  
465 WE HIT OUR SAMPLE SIZE CHART...  
466 50 AGAIN.  
467 ...  
468 OK, AND WE HIT IT.  
469 WITH JUST ONE ERROR.  
470 60 AGAIN.  
471 ...  
472 10 ERRORS,  
473 WHICH IS WAY UP THERE.  
474 WE END UP A LITTLE MORE CONSERVATIVE,  
475 BUT THE OBJECT HERE IS TO HIT IT ON THE NOSE AND AT  
476 20 CENTS A HEAD,  
477 LET'S TRY 80 ITEMS.  
478 ...  
479 WITH 5 ERRORS,  
480 WE STILL MISSED IT.  
481 TRY 80 AGAIN.  
482 ...  
483 NO ERRORS, OK.  
484 TRY 80 AGAIN.  
485 I'M INTERESTED TO SEE HOW...  
486 I WOULD THINK THAT WE WOULD HIT IT ON THE NOSE A LOT  
487 MORE OFTEN.  
488 THAN WITH 50 OR 60.
Which is apparently the name of the game,
because the charge per transaction isn't that great.

3 errors, and it hit it.
Go with 80 again.

2 errors,
and we missed it.
Is this based on...
Does this thing actually set a population that has
 certain
items keyed as being errors
and you just pick a random sample and it hits 'em or
it doesn't.
(That's a close way of putting it)
So these are very realistic.
(Yes, they are)
Go with 80 again.

(IT's not a fixed game)
No, I mean these are very realistic...
It simply sets a population,
of 7000,
and says these ones are bad,
whichever percentage it happens to be this time.
(Yes)
And then you randomly select 80 or 40 or whatever,
and you get some out.
8 items, which is 10%,
and we hit it.
We're getting,
go with 80.

We're getting a lot better matching of actual
side 2 of tape ii begins here.
Keep on missing the $100 payoff.
go to 140.

Still missed it, wow.
Try that again.

24 errors.
That's about 20%, 18 to 20%.
Which is awful high.
That's not going to happen much more.
We're still getting off our...
We don't seem to be doing a hell of a lot better,
go with 140 again.

It's kinda hard to believe
that with that many items,
you'd be that far off very often.

(Um hm.)
I think that somewhere in the 80 to 100 range
would pretty much get you barely there.
(I'd like to reserve my comments for now.)
4 errors, and we hit it.

GO WITH 140 AGAIN.

1 error, got it.

At this point, the $28 isn't costing us much from what we're getting.

Then again, if I wanted to be real certain of getting that...

It's only gonna cost me $30 to check 200 items.

Go with 140 again.

OK.

Try that again.

Unintelligible.

First 2 times I did that we missed it.

Keep going with 140.

... 29 errors,

which is 20%,

which, although it's happened a couple of times here,

shouldn't happen very often,

based on this chart back here.

very often, .. very seldom.

In fact, there's only one time that it was up around that point,

out of 100.

which means we've picked the 1% odds twice now.

Go with 140 again.

... 2 errors and we got it.

Go with that again.

We seem to be getting more (unclear) than I would have thought

at this high of a sample size.

Part of that is because we had an awful lot of errors in two

of those.

... for example,

is if I were going to go in expecting...

d and the actual c...

if I went in expecting 20% errors,

I wouldn't do the testing in the first place.

(um hm)

Then again, the point here is to maximize the payoff,

which is a little bit different situation than what we'd be...

If we testing,

doing compliance test samples,

ok, if we're going to be anywhere near that number of errors,

we wouldn't care how far they were.
SUBJECT I

1 YOU ARE THE SENIOR IN CHARGE OF THE AUDIT OF ROHR INDUSTRIES
2 SUBJECT I IS READING EXPERIMENT
3 TO MAKE THE COMPUTER GENERATION PROCESS MORE EXPLICIT TO YOU,
4 HERE IS A SAMPLE OF 100 ERROR RATE PERCENTAGES GENERATED BY THE
5 COMPUTER.
6 JUST TO UH, CLARIFY THIS,
7 I'M NOT QUITE SURE,
8 WHEN YOU'RE TALKING ABOUT, DOWN HERE,
9 YOUR NET PAYOFF;
10 IS THIS JUST LIKE...
11 (WELL FOR EACH REPEAT OF THE EXPERIMENT, I'M GOING TO
12 AWARD YOU
13 FICTITIOUS SUM OF MONEY, AND THAT FICTITIOUS SUM OF
14 MONEY IS A
15 FUNCTION OF THIS SYSTEM HERE.)
16 HH, OK.
17 SO IT'S LIKE,
18 KIND OF A REWARD.
19 IT'S JUST TRYING TO SAY WELL, THERE'S A SPECIFIC
20 ANSWER THAT
21 YOUR MODEL PREDICTS, OR..
22 (WELL, I'M JUST ASKING YOU TO DO AS WELL AS POSSIBLE IN THE
23 EXPERIMENT. UM, IT'S HARD TO SAY WHETHER THERE'S ANY
24 RIGHT
25 ANSWER, BUT SOME ANSWERS ARE BETTER THAN OTHERS.)
26 OK.
27 SO I CHOOSE THE SAMPLE SIZE AND YOU INPUT IT AND..
28 (WELL, UM, YEAH. THERE ARE Uh, THE CASE REFERS TO
29 OTHER MATERIAL)
30 ALL RIGHT, ITEMS A AND B.
31 OK, SO I NEED TO LOOK THROUGH THESE AND THEN,
32 BASED ON THIS INFORMATION,
33 COME UP WITH A SAMPLE SIZE.
34 OK, LET'S SEE,
35 IF I WERE JUST STARTING OUT WITH A NEW CLIENT,
36 AND AM CONCENTRATING ON THE PAYABLES AND PURCHASES
37 SYSTEM.
38 I KNOW THEY HAVE GIVEN THEM AN UNQUALIFIED OPINION,
39 AND THEY HAVE APPROXIMATELY 7000 TRANSACTIONS IN THE
40 POPULATION.
41 LET'S SEE--
42 EACH NEW COMPANY IS IDENTICAL EXCEPT FOR THE NUMBER OF
43 DISBURSEMENTS IN THE POPULATION, 7000, WHICH DO NOT
44 COMPLY, OK.
So I've already gone through and done a review of the internal controls at this point then, and we've written up a memorandum and flow charts. Ok, now let me look at what I've got here: financial statements; balance sheet and income statement, retained earnings, statement of changes, long term debt, stock redemption, income taxes. Company notes are attached, appendix a. memorandum contained in appendix b. Determine the specs. I guess the first thing that I would do is look over their financial statements to get an idea of the company itself, a little bit, try to get some background on it. A company that had $8,000,000 in assets last year. A million in payables, trade. Did you tell me in here... ok, they're a manufacturer of cabinets, I presume that they're publicly held then, they have sec reporting requirements. (that's right) number of shareholders, $100 par value, so we have preferred stock of 1224 shares issued & outstanding. That's cumulative, now we have non-cumulative of 1500 shares. Common stock; 6980 shares issued and 1032 in treasurer, non voting. 62,020 shares issued and outstanding, so we got a little over 70,000 shares. Ok, long-term debt, due after one year, ok, we have a mortgage note payable to the insurance company, have a bank, have a revolving credit agreement, converted to a term loan in 1980. Another note payable to the bank. A pension plan and retirement contracts. Capital leases lesser amount due within one year. Ok, so we're talking about substantial number of shareholders, I guess around 70,000. We've got a bank as far as determining who would be interested in their statements, and their results of operations.
The shareholders would be interested, and you have a bank out there who has made loans to us so, they obviously have some stake in the company. SEC reporting. I guess we have a number of outside users that are going to be very interested in our statements and our report. Has agreed to maintain working capital of no less than $1,000,000, certain limitations with respect to indebtedness. In addition, the loan agreement and mortgage deed contain among other things, provisions relating to restrictions on loans, advances, guarantees, etc.

So again, we also have a um, the insurance company has a number of restrictive covenants on our company. <Unclear>...terminated on May 31, converted to a 7 year term loan amortization schedule. The bank's prime rate. Ok, post interest with final payment of $387,000 in 87. Again, the loan agreement provides the company must maintain working capital of at least a million dollars, and stockholders equity of at least 2 million, with a ratio of shareholders equity, total liabilities of at least 1.0 to 3.25. Can repay it at any time without penalty, not payable to the bank with with a balance of <Unclear> income tax expense, pension plan, retirement contracts. Ok, so I guess, kind of a quick review, I guess it's kind of like for the risk that's involved here, the importance of trying to get an accurate result, and make sure, try to find out what kind of error rate there actually is in the system. Seems like there's a rather substantial, that we have a number of people on the outside that are going to be relying on these statements.
WE DO HAVE SOME REQUIREMENTS HERE FOR THEIR WORKING CAPITAL,
AND THEIR...
OK, NEXT I'D LIKE TO...
THEN I'D LIKE TO LOOK AT
I GUESS, THE REVIEW OF THE SYSTEM TO SEE WHAT ACTUAL SET UP IS,
IN THEORY.
SEE YOU KNOW, WHAT...
OK, SO WE GOT THE FLOW CHART;
ACCOUNTS PAYABLE, CASH DISBURSEMENTS AND PURCHASES.
WELL WE'LL LOOK AT PURCHASING HERE.
ROHR INDUSTRIES MEMO INVENTORY CONTROL.
WHEN AN ORDER IS TAKEN,
THE ENGINEERING DEPARTMENT IS GIVEN A COPY OF THE ORDER TO DETERMINE THE SPECs. THEY THEN FORWARD THE MATERIALS REQUIREMENTS TO THE INVENTORY CONTROL DEPARTMENT HEAD. BASED ON EOQ, AND MATERIALS NEEDS, INVENTORY CONTROL DETERMINES THE NEED TO ORDER MATERIALS.
WHEN A DETERMINATION IS MADE THAT AN ITEM SHOULD BE ORDERED,
INVENTORY CONTROL DEPARTMENT FORWARDS TO THE PURCHASING AGENT,
A REQUIREMENTS LIST FROM WHICH HE IS TO MAKE PURCHASE ORDERS.
THIS IS DONE DAILY.
OK, SO WE GET AN ORDER FROM OUTSIDE,
ENGINEERING DEPT. GETS A COPY OF THE ORDER TO DETERMINE THE SPECs,
WHAT THEY'RE GOING TO NEED,
THEN THEY FORWARD A MATERIALS REQUIREMENTS TO THE INVENTORY CONTROL DEPARTMENT HEAD.
MATERIAL NEEDS,
HE THEN DETERMINES THE NEED TO ORDER MATERIALS.
INVENTORY DEPARTMENT FORWARDS TO THE PURCHASING AGENT,
A REQUIREMENTS LIST FROM WHICH HE IS TO MAKE PURCHASE ORDERS.
OK.
SO ENGINEERING...
OK, SO WE HAVE 1,2,3,4,5,6,7 PART BILL OF MATERIALS.
ENGINEERING - FROM ORDER ENTRY SYSTEM.
OK BILL OF MATERIALS, CABINET PANELING, ETC.
ALL FILED ALPHABATICALLY AND NUMERICALLY.
FILE WITH ORDERS TO ASSIST IN COSTING AN ORDER.
FILE WITH ORDERS TO ASSIST IN COSTING AN ORDER, OK.
REQUEST FOR PURCHASE, LET'S SEE.
DOCUMENT, TO FILE.
SPECIAL ITEMS;
ANY ITEM RELATED TO A CONTRACT, OR A CONTRACT JOB UP TO 40% OF PURCHASES.
Bill of Materials.

Manual Operation.

168 Assets over $1000, the President has to approve.

169 Assets over $5000, the Board of Directors.

170 Purchase order over $10,000 for production materials, the President.

171 All other purchase orders, the Purchasing Agent has ...

172 Engineering determines materials requirements for special orders

173 and has bill of materials made out.

174 OK, purchasing.

175 Material needed for stock purchases and supplies,

176 for departmental expense,

177 it must have department approval.

178 What does the PA stand for?

179 (Purchasing Agent)

180 Purchasing Agent approval needed for stock purchases and supplies.

181 Department expense – it must have department approval.

182 Purchasing.

183 Purchasing, 8 part.

184 OK, purchase order;

185 got one going to the vendor.

186 Looks like one's going to accounting.

187 Operation.

188 Cost Accountant determines the cost distribution

189 and the determination of any variance.

190 OK, so he's costing it out.

191 distribution, determination of any variance.

192 It's filed.

193 It's filed on line, boom.

194 OK, we have one order going to Data Processing.

195 File on line, nope, file off line.

196 Numerically.

197 Off line, numerically.

198 OK, so we've got purchasing agent.

199 OK, inventory control.

200 Receiving gets one,

201 partial receiving,

202 have a partial receiving,

203 and another partial receiving.

204 Those go to receiving,

205 and they're filed on line.

206 Purchasing.

207 and file this one on line.

208 Only partial shipment.

209 Looks like a matching operation, yep.

210 OK, then receiving gets the goods from the vendor.

211 Freight bill and receiving report.

212 What's he do here,

213 match receiving report or quantity received with the purchase order.

214 And what's he do here,

215 so we do have a matching here.
MATCH THE RECEIVING REPORT OR QTY. RECEIVED WITH THE PURCHASE ORDER
LET'S SEE, NUMBER 4,
PURCHASE ORDER 4,
MATCHED TO THE RECEIVING REPORT OF THE VENDOR.
PURCHASE ORDER 7;
FREIGHT BILL, RECEIVING REPORT,
THEY SEND IT BACK TO PURCHASING.
DOES THIS INDICATE THAT BOTH OF THESE WERE SENT BACK OR,
THIS ONE JUST GETS FILLED DOWN HERE, AND THIS ONE SENT OVER HERE?
(UM, LET'S SEE, WELL THIS ONE, THIS WHOLE THING GETS SENT OVER HERE.
THIS THING GETS FILED DOWN HERE.)
OK.
IS THIS THE SAME SITUATION, OR ..
OK, THIS ONE SAYS PARTIAL RECEIVING,
PO 7,
AND NUMBER 4, I GUESS IS JUST THEIR COPY THEN, RIGHT,
THAT THEY'RE GOING TO KEEP OVER HERE.
OK,
(RIGHT)
OK, SO THEY FILED ONE AWAY IN THEIR FILES
AND THEY SEND THE OTHER ONE TO ACCOUNTING,
BACK TO PURCHASING.
OK.
OK, THIS IS JUST THE PURCHASING PART,
THIS DOES NOT INCLUDE,
THIS IS NOT ANY OF THE OTHER ACCOUNTING FUNCTIONS,
THIS IS NOT THE ACCOUNTING DEPARTMENT?
(THAT'S RIGHT, YOU'LL SEE THE ACCOUNTING DEPARTMENT IS THE 2ND FUNCTION)
OK, SO RECEIVING SENDS IT BACK TO PURCHASING,
PURCHASING MATCHES IT WITH THEIR COPY OF THE PO.
MATCH THE QTY.
IT'S ONLY A PARTIAL SHIPMENT.
LET'S SEE,
MATCH QTY.
SO THEY'VE GOT,
HERE WE GOT PO 8
MAYBE A COPY OF 5, 6, OR 7, DEPENDING WHETHER THE SHIPMENT WAS PARTIAL OR COMPLETE.
OK, USUALLY THE FREIGHT BILL,
WE'VE GOT OUR PURCHASE ORDER,
WE'VE GOT THE ONE SENT BACK FROM RECEIVING,
AND WE'VE GOT THE FREIGHT BILL NOW.
OK, IF IT'S ONLY A PARTIAL ORDER,
IT'S STORED ON LINE,
AND WE WAIT FOR THE REST OF IT TO COME IN.
IF IT'S A COMPLETE ORDER,
WE FILE OUR PURCHASE ORDER,
266 THE ONE IN PURCHASING OFF LINE, ALPHABETICALLY.
267 OK, THE OTHER PURCHASE ORDER AND THE FREIGHT BILL
268 FILED ON LINE.
269 UNMATCHED RECEIVER FILE REVIEWED EVERY MONTH.
269 HM, SO NOW WE HAVE THE VENDOR'S INVOICE BEING GIVEN TO
270 THE SECRETARY
270 IN PURCHASING.
271 OK, THE SECRETARY VENDORS INVOICE;
272 INVOICE IS SENT TO PURCHASING AGENT WHO REVIEWS IT,
273 HM,
274 SO APPARENTLY THE INVOICE HAS GONE STRAIGHT TO
275 PURCHASING WITHOUT
275 ANY ACCOUNTING CONTROL OVER IT YET,
276 OR THEY HAVEN'T HAD ANY...
277 OK, INVOICE IS SENT TO PURCHASING AGENT WHO REVIEWS
278 IT,
279 WHEN HE'S REVIEWING THE INVOICE AND HIS COPY OF THE
279 PURCHASE ORDER,
280 SENT BACK TO THE PURCHASING AGENT, REVIEWED,
281 GOES TO PAGE 2.
281 OK, RIGHT THERE.
282 (THIS IS THE CONNECTOR FROM HERE TO HERE)
283 OK.
284 LOOKING AT THE ACCOUNTS PAYABLE AND CASH
285 DISBURSEMENTS.
285 SO OUR INVOICE IS COMING DIRECTLY FROM PURCHASING.
286 NOW, LET'S SEE.
287 ACCOUNTING HAD RECEIVING A COPY OF THE PURCHASING
288 ORDER.
289 I GUESS FROM INVENTORY CONTROL.
289 THE COST GUY HAS ALREADY MADE THE COST DISTRIBUTION
290 AND
291 DETERMINED VARIANCES THAT ARE EXPECTED.
291 SO HE'S NOW MATCHING THE INVOICE WITH HIS, WE HAVE PO
292 5, 6, OR 7.
292 UPON APPROVAL FOR PAYMENT BY PURCHASING AGENT,
293 THE INVOICE AND THE PURCHASE ORDER ARE FORWARDED TO
294 THE COST
295 ACCOUNTANT
296 WHO STAMPS THE INVOICE WITH A PAYMENT STAMP,
296 WHO STAMPS THE INVOICE WITH A PAYMENT STAMP AND ENTERS
297 THE INVENTORY
297 ACCOUNT CODING.
298 HE THEN GIVES THE INVOICE TO THE ACCOUNTING CLERK FOR
299 PAYMENT.
299 OK, SO WE HAVE ONE COST ACCOUNTANT WHO'S MATCHING IT,
300 UPON APPROVAL,
301 AND HE'S GOT IT,
302 FORWARD IT TO THE COST ACCOUNT WHO STAMPS THE INVOICE
303 WITH PAYMENT,
304 SO HE'S MAKING THE ACCOUNTING CODE.
304 HE THEN GIVES THE INVOICE TO THE ACCOUNTING CLERK FOR
305 PAYMENT.
305 SO THE ACCOUNTING CLERK GETS IT.
THE ACCOUNTING CLERK TAKES ANY DISCOUNT AVAILABLE, RUNS A TAPE TO DETERMINE ACCURACY OF INVOICE SUMMARY, AND HAS PAYMENTS APPROVED BY THE CONTROLLER. THE CONTROLLER APPROVES ALL CASH DISBURSEMENTS. AUTHORIZED SIGNERS ARE THE PRESIDENT AND THE CONTROLLER.

OK, SO HE IS APPROVING IT BEFORE OK, WE HAVE ANY CHECKS MADE. THEN IT'S PASSED ON TO NCR OPERATOR. NOW, DATA PROCESSING HAS A CUT. SO THE INVOICE AND THE PURCHASE ORDER, 5, 6, AND 7. WHAT'S THE RB? (RECEIVING REPORT) OK, RECEIVING REPORT. PURCHASE ORDER, RECEIVING REPORT. INVOICE, PURCHASE ORDER, HE'S FILING ON LINE BY VENDOR AND BY DATE.

LET'S SEE HERE.

OK, SO WE HAVE THE UB, DATA PROCESSOR OPERATOR, THEN WE HAVE THIS DOTTED LINE THAT COMES DOWN TO VOUCHER, DOES THAT MEAN THEN THAT HE'S INPUTTED IN THE COMPUTER, AND THAT PRINTS UP A VOUCHER, OR? (I BELIEVE SO. WELL, THE NCR OPERATOR IS BASICALLY A POSTING MACHINE)

OK, OK.

OK, SO HE'S GOING TO MAKE UP A VOUCHER. SIMULTANEOUS POSTING, VOUCHER AND DISTRIBUTION REGISTER.

SO THAT WE'VE GOT ONE PERSON THAT IS RECORDING IT IN THE BOOKS, VOUCHER AND THE DISTRIBUTION REGISTER. AND HE ALSO IS IN CHARGE OF MAKING UP THE CHECK, DOES HE ALSO PRINT THE CHECK?

(YES) LET'S SEE, DISTRIBUTION TICKETS ARE KEPT FOR EACH ACCOUNT AND POSTED AT THE SAME TIME AS VOUCHER REGISTER. FOR EACH ACCOUNT, TAPE IS MADE AT END OF THE MONTH FOR EACH ACCOUNT WHICH IS USED FOR POSTING TO THE GENERAL LEDGER.

OK, VOUCHER, DISTRIBUTION REGISTER, DISTRIBUTION TICKET. WHICH WE COMPARED TO THE CHECK REGISTER APPARENTLY. WITH THE CHECKS POSTED, THE CHECK REGISTER IS COMPARED TO THE DISTRIBUTION TICKETS AND WE DON'T COMPARE IT TO THE VOUCHER AND DISTRIBUTION REGISTER.

LET'S SEE, BY ACCOUNT NUMBER, MONTHLY.
OK, this symbol right here, Ed.
(um, by date)
OK, by account number monthly.
OK, distribution tickets are filed by account number monthly.
OK, so this person, this operator is preparing a voucher,
proof of invoices, whatever,
he's posting that to two,
a voucher and distribution register,
those are separate items, correct?
(no that's one register)
that's one register, ok, ok.
so he does that,
and then the simultaneous posting we have listed here is at the
same time he records it on this distribution ticket.
(that's right)
ok, so we're doing that at the same time.
so the voucher and the distribution register are,
i presume,
is just recording the voucher amount?
by account number or distribution?
is it broken down, i guess?
and then at the same time we fill out a separate ticket for just
each account.
ok, so at the end of the month this will have the
total activity
of what we purchased for that one individual inventory account.
(right)
and then that is used to post to the general ledger, ok.
compare.
check register and distribution ticket.
so we make up a check,
he's comparing that to the distribution ticket.
mm, ok,
then he makes up an adding machine tape,
posts that to the general ledger.
posts the original,
makes up an adding machine tape for each account,
posted at the same time as voucher register.
The tape is made at the end of the month for each
account
which is used for posting.
so we make up this,
post it to the general ledger.
so we're posting inventory through the distribution tickets
and the check register is hitting the cash.
and making the payment of the actual payee.
ok, now what do we do here?
GOT THE CHECK,
ACCOUNTING CLERK RECEIVES THE CHECKS,
ATTACHES THE REMITTANCE ADVICES AND HAS THEM SIGNED,
SEE BELOW,
BEFORE SENDING THEM OUT.
OK, SO THE PRESIDENT OR THE CONTROLLER SIGN ALL THE
CHECKS MANUALLY
BEFORE SENDING THEM OUT.
NCR OPERATOR POSTS THE GROSS AMOUNTS AND THE CHECK
NUMBERS INTO
THE CHECK REGISTER.
THE CHECK GOES TO THE VENDOR,
AND LET'S SEE, WHAT ARE WE DOING HERE?
OK, SO THE VOUCHER GETS FILED,
OFF LINE, ALPHABETICALLY, BY YEAR.
THE INVOICE,
THE PURCHASE ORDER,
AND THE CHECK COPY ARE ALSO FILED.
OK.
SO NOW WHAT HAVE I NOTED FROM THIS REVIEW?
OK, I'VE NOTICED THAT THE INVOICE FROM THE VENDOR DOES
NOT GO DIRECTLY TO ACCOUNTING,
IT GOES TO PURCHASING.
ACCOUNTING IS UNABLE TO ESTABLISH A CONTROL OVER
IT.
INSTEAD OF Normally, YOU KNOW, SENDING IT TO
ACCOUNTING,
GETTING THE, GETTING IT IN THE SYSTEM
GETTING SOME CONTROL OVER IT AND THEN SENDING IT BACK
TO
PURCHASING FOR REVIEW,
IT'S GOING STRAIGHT TO PURCHASING,
LET'S SEE WHAT ELSE,
FREIGHT BILL,
PURCHASE ORDERING.
OK, WE HAVE ONE CLERK HERE,
OK, THE CONTROLLER APPARENTLY REVIEWS ALL INVOICES
BEFORE THEY'RE
PROCESSED,
AT THIS POINT,
DISCOUNT AVAILABLE.
PAYMENTS APPROVED BY CONTROLLER.
SO CONTROLLER HAS APPROVED THE INVOICE BEFORE IT'S
PROCESSED.
IT COMES INTO THE SYSTEM,
WE HAVE ONE CLERK HERE,
WHO GETS THE INVOICE,
DOES ALL THE POSTING.
HE POSTS IT TO THE VOUCHER AND DISTRIBUTION REGISTER,
HE MAKES UP THE CHECK, ETC.
HE'S THE ONE THAT RECORDS IT IN THE CHECK REGISTER,
HE'S THE ONE THAT POSTS IT TO THE LEDGER
HM, ALL OF THIS IS PERFORMED BY ONE PERSON, THAT CORRECT?

(UM, HM)

HE'S POSTING IT TO THE GENERAL LEDGER,

HE'S POSTING IT TO THE CHECK REGISTER,

WE DON'T HAVE A REAL GOOD, STRONG SEGREGATION OF DUTIES

IN THIS AREA.

DISTRIBUTION TICKETS

I'M NOT SURE,

LET'S SEE,

WE DON'T SEEM TO HAVE ANY,

WE'RE COMPARING THE CHECK REGISTER TO THE DISTRIBUTION TICKETS,

SO,

I'M NOT SURE,

IS THAT JUST LIKE HE'S ADDING UP ALL THE DISTRIBUTION TICKETS?

(NO, WHAT THAT SAYS IS THAT, UM, THE TOTAL DISTRIBUTION FOR A PARTICULAR CHECK HAS TO EQUAL THE AMOUNT OF THE CHECK THAT IT'S WRITTEN FOR).

OK, SO WE'RE ADDING UP,

IN OTHER WORDS,

IF WE DISTRIBUTED TO 3 DIFFERENT TICKETS,

WE ADD THOSE UP AND COMPARE THAT TO THE CHECK.

(RIGHT)

SO, ONE PERSON,

HE'S GOT A CONTROL,

SUPPOSEDLY HE'S COMPARING,

WE'VE MADE SURE WE'VE ENTERED THE SAME AMOUNT IN THE CHECK REGISTER AND IN THE DISTRIBUTION TICKETS,

WHICH HOPEFULLY GO TO THE GENERAL LEDGER,

OK.

SO,

NOW, DOES ACCOUNTING EVER GET TO,

SO BASICALLY,

YOU'VE GOT THE CHECKS BACK IN,

ATTACHING THE REMITTANCE ADVICE,

HAS THEM SIGNED,

SO THEN,

BEFORE THE CHECKS OUT,

SOMEBODY HAS REVIEWED THEM,

CAN YOU TELL ME

I DON'T KNOW IF THIS IS IN HERE OR--

WHEN THEY GET THESE SIGNED,

IN OTHER WORDS,

THE INVOICE;

LET ME CHECK THIS OUT HERE,

THE INVOICE JUST GETS FILED UP HERE

(RIGHT)

OK, I'M NOT SURE,
AT WHAT POINT,
THEN ARE WE Sending IT BACK TO ACCOUNTING TO BE FILED EVENTUALLY?
(YEah)
OK.
SO HE JUST...
WE FILE IT HERE AND THEN PASS IT ON ONCE WE'RE DONE?
(YEah, THAT'S RIGHT)
SO APPARENTLY,
AFTER THIS POINT,
NOBODY ACTUALLY COMPARES THE INVOICE AND THE CHECK AMOUNT.
WE JUST...
HE SUPPOSEDLY DOES ALL THAT
(UH Hn)
OK.
I GUESS THE MAJOR WEAKNESSES I'M SEEING,
IS THE INVOICE GOES TO PURCHASING FIRST
AND THAT WE HAVE ONE CLERK THAT'S POSTING, MAKING UP THE CHECK,
RECONCILING, POSTING TO THE GENERAL LEDGER,
THERE'S NOBODY REVIEWING HIS WORK, UM,
YOU KNOW, IT'S POSSIBLE THAT HE COULD MAKE UP FICTICIOUS CHECKS,
HOWEVER, THE ONE CONTROL IS THAT IT HAS TO BE SIGNED BY THE
PRESIDENT OR THE CONTROLLER,
FOR HE REVIEWS IT,
AND SINCE HE'S ALREADY REVIEWED THE INVOICE,
DEPENDING ON THE FLOW,
I DON'T KNOW,
DEPENDING ON HOW MUCH HE CAN RETAIN,
THAT HE'S LOOKED AT SOMETHING BEFORE.
SO I SEE A RATHER LARGE COMPANY THAT HAS A VERY POOR SEGREGATION
OF DUTIES HERE.
I GUESS I'M AT THE POINT NOW TO FIGURE OUT A SAMPLE SIZE.
SIDE 2 STARTS HERE OF TAPE.
TRUE STATE OF INTERNAL CONTROLS.
OK, I'M A LITTLE CONFUSED HERE,
I DON'T KNOW EXACTLY WHAT I'M SUPPOSED TO DO AT THIS POINT.
WE TALKED ABOUT...
OK, IT SAYS THAT I'M SUPPOSED TO PICK OUT THE SAMPLE SIZE
FOR THE COMPLIANCE TESTING.
BY FIRST HAVING YOU CHOOSE A SAMPLE SIZE.
SO AFTER EVALUATING ALL THIS INFORMATION,
AND KNOWING THAT THEY HAVE APPROXIMATELY 7000 CASH DISBURSEMENTS
IN THE POPULATION,
I'M JUST SUPPOSED TO SAY,
I WANT TO TEST SO MANY ITEMS,
I want to say,

or is that my decision also?

That's your decision.

Okay, that's my decision.

The true state of internal controls...

Sometimes after completion of the audit.

Now, the monitor will then inform you of this information, OK.

I won't worry about that I guess.

Well I guess my feeling are...

I see that the company has a number of,

there seems to be,

I would say a substantial risk here,

I don't feel that the internal controls are very strong,

and yet we do have a lot of,

as far as the company's concerned,

there is a number of outside users that are going to be prepared

to review these.

I know that I had another big 8 firm that reviewed this last.

Yeah,

I don't know what they did exactly, but,

we've given them an unqualified opinion,

so that leaves me to believe that they'd not have a lot of

problems.

I can't say that they didn't have any problem with the internal

controls themselves,

but at least their testing has proved that apparently the

process is working,

they're getting a good accounts payable balance,

and then spread to the inventory accounts etc.

I guess in my own mind,

I'm not going to place heavy reliance on the other accountants

at this point.

Seeing as we have very little familiarity with this company.

so seeing as this is the first year through,

I would probably,

I would definitely tend to probable test a little heavier

than I think I would in future years,

if we've done the testing ourselves for a number of years,

and looked at the results.

Um,

I think at this point,

Well, I guess with the population, 7000,
I WOULD,
I THINK I WOULD PROBABLY TEND TO USE A STATISTICAL,
RATHER THAN JUST PICKING OUT A NUMBER OF ITEMS.
(SO DO YOU WANT TO USE THE STATISTICAL APPROACH?)
CASH DISBURSEMENTS.
YES, I'LL TRY THE STATISTICAL SAMPLING I GUESS.
THIS IS CAPABLE OF DOING THIS RIGHT?
(YEAH)
NOW, WHAT EXACT INFORMATION I NEED.
(I'LL HELP YOU WITH THAT)
OK.
(HERE'S SOME INSTRUCTIONS ON HOW TO USE STATISTICAL SAMPLE TABLES.)
YOU NEED TO SPECIFY 3 VALUES TO USE THE SAMPLE SELECTION PROGRAM.
YOUR ESTIMATE OF THE POPULATION ERROR RATE,
HH,
UPPER CONFIDENCE LEVEL,
AND A LEVEL OF RISK.
DEFINITIONS OF IMPORTANT TERMS FOLLOW TO FACILITATE YOUR USE
OF THE PROGRAM.
SAMPLE SELECTION BASIS:
TRANSACTIONS WILL BE DRAWN FROM THE POPULATION RANDOMLY.
THAT IS EACH CASH DISBURSEMENT HAS AN EQUAL CHANCE OF BEING
SELECTED, REGARDLESS OF IT'S MAGNITUDE, OK.
CONFIDENCE LEVEL:
NUMBER OF TIMES IN 100 THAT THE SAMPLE WILL APPROXIMATELY
REPRESENT THE POPULATION LYING WITHIN THE CONFIDENCE LEVEL.
IT'S IN.
LEVEL OF RISK:
ONE MINUS THE LEVEL,
FOR EXAMPLE,
IF THE CONFIDENCE LEVEL IS 95%,
THE LEVEL OF RISK WOULD BE 5%.
POPULATION ERROR RATE:
The number of transactions in the population which do not
comply with the internal control procedures divided by
the total number of transactions in the population.
Oh, OK.
This number may be thought of as the probability that a randomly
selected transaction does not comply with internal control
procedures.
The upper limit on the population error rate auditor
is willing
to accept.
THE NUMBER OF TRANSACTIONS IN THE POPULATION WHICH DO NOT COMPLY
WITH INTERNAL CONTROL PROCEDURES DIVIDED BY ...
OK.
HERE IS A SAMPLE OF 100 ERROR RATES AND PERCENTAGES GENERATED BY
THE COMPUTER.
FOR EACH REPEAT OF THE EXPERIMENT,
ASSUME YOU ARE IN CHARGE...
EXCEPT FOR THE NUMBER WHICH DO NOT COMPLY..
THIS NUMBER IS DETERMINED AT RANDOM BY COMPUTER IN A MANNER WHICH
IS INTENDED TO REFLECT THE FREQUENCY OF ERRORS OBSERVED OVER ALL
CLIENTS AUDITED BY BIG 8 CPA FIRMS WITH SIMILAR CHARACTERISTICS.
OK.
WHAT EXACTLY IS THAT TELLING ME?
IS THIS,
WHEN WE'RE TALKING ABOUT THIS ERROR RATE IN THIS LAST PARAGRAPH
HERE,
IS THIS SAYING,
FOR EACH TIME WE DO THIS THAT...
(WELL, FOR ONE THING, IT COULD SAY, IS THAT OF YOUR LAST 100
AUDITS, ONE WAY TO INTERPRET IT IS OF YOUR LAST 100
AUDITS,
THESE ARE THE PERCENTAGE ERROR RATES THAT YOU OBSERVED.)
OH, OK.
SO, SEEING I REALLY DON'T HAVE A LOT OF FAMILIARITY WITH THIS
COMPANY,
I'M BASICALLY GOING TO HAVE TO USE MY EXPERIENCE WITH SIMILAR
CONTROLS,
AND PROCEDURES,
AND EXPERIENCE ERROR RATES WITHIN THE LAST NUMBER OF AUDITS, OK.
ESTIMATE OF POPULATION ERROR RATE, HM.
OBSERVED OVER ALL CLIENTS.
NOW, WHAT KIND OF CONTROLS?
WILL APPROXIMATELY REPRESENT THE POPULATION.
LEVEL OF RISK.
WELL, OK, MY, I GUESS,
SKIP THE POPULATION ERROR RATE FOR NOW, BUT,
I GUESS THE CONFIDENCE LEVEL I'M TENDING TOWARD IS PROBABLY A
95% RATE AT THIS POINT.
AGAIN, BEING A FIRST TIME AUDIT,
WITH THE CONTROLS,
WHICH APPEAR TO ME, NOT TO BE REAL STRONG.
AND THAT GIVES ME A LEVEL OF RISK OF 5%.
SO I'M SAYING, I'M TAKING A 5% RISK THAT THE RESULTS OF THIS
COULD BE INACCURATE, RIGHT?
(RIGHT)
OK, MY ESTIMATE OF THE POPULATION ERROR RATE, Huh...
THE OPPORTUNITY TO LEARN FROM EXPERIENCE, OK.
A MANUAl SYSTEM OF POSTING SIMULTANEOUSLY TO THE
GENERAL LEDGER.
LET'S SEE.
LOOKING AT EXPERIENCE,
8, 8, 8, 4, 6, 2, 2, 0, 14, 5, 9, 5, 5, 5, 1, 4, 17,
12, 3, 4.
WELL, LET ME SEE.
WELL, LET'S SEE, HOW 'BOUT IF WE,
WHAT RATIONALE DID I USE ON THIS,
POPULATION ERROR RATE.
WELL, I GUESS I WOULD,
I GUESS THERE'S ONLY 2 REAL THINGS I CAN DRAW HERE,
JUST KIND OF GOT THE FEEL,
I GUESS AFTER I'VE GONE THROUGH AND REVIEWED THE CONTROLS WITH
THE PEOPLE,
SUPPOSEDLY, I GUESS, SEEN THEM IN ACTION,
I GUESS IN OUR CONSIDERATION,
HOPEFULLY,
THIS SAME PERSON'S BEEN IN THE SAME POSITION THE ENTIRE YEAR,
AND WE HAVEN'T HAD A LOT OF PEOPLE CHANGING.
I GUESS, AND ALSO, JUST MY PAST EXPERIENCE,
UM, AND HOW CONSCIENTIOUS THIS PERSON IS,
I PRESUME, WELL,
I WOULD ESTIMATE — WELL,
I WOULD KINDA PICK A QUICK AVERAGE I GUESS —
I'D PROBABLY SAY
I GUESS 4.5,
I GUESS THAT'S BASICALLY JUST KIND OF A GUESS.
(ALRIGHT, AND YOU NEED YOUR UPPER CONFIDENCE BOUND)
MY UPPER CONFIDENCE BOUND.
THE UPPER LIMIT ON THE POPULATION ERROR RATE IS WILLING TO ACCEPT.
OK, POPULATION ERROR RATE.
THE NUMBER OF TRANSACTIONS IN THE POPULATION WHICH DO NOT COMPLY
WITH INTERNAL CONTROL PROCEDURES,
DIVIDED BY THE TOTAL NUMBER OF TRANSACTIONS IN THE POPULATION.
SO, WE'RE SAYING LIKE —
POPULATION ERROR RATE COULD BE LIKE...
(WELL, THESE ARE THE POPULATION ERROR RATES HERE)
OK, SO I'VE SAI.D, ESTIMATED IT AT 4.5,
NOW WE'RE SAYING WHAT'S...
(WHAT'S THE MAXIMUM TOLERANCE)
OK.
250 THIS NUMBER MAY BE THOUGHT OF AS THE PROBABILITY THAT

705 COMPLY WITH INTERNAL CONTROL PROCEDURES.
706 SO WE'RE SAYING HOW MUCH ERROR WILL I TOLERATE
707 TO RELY ON THESE CONTROLS
708 OR TO DETERMINE MY,
709 THAT MY INTENDED SUBSTANTIVE TESTING IS ADEQUATE..
710 I'M ESTIMATING THAT IT'S 4.5%.
711 HM.
712 SO, THE UPPER LIMIT ON THE POPULATION ERROR RATE
713 THE AUDITOR IS WILLING TO ACCEPT.
714 THEY'RE A BIG COMPANY,
715 7000 TRANSACTIONS.
716 UN, I GUESS I'LL SAY,
717 I GUESS 7%.
718 ...
719 (ALRIGHT, THE SAMPLE TABLE SAYS THAT YOU SHOULD SAMPLE
187

720 TRANSACTIONS, YOU WANT TO SAMPLE 187 TRANSACTIONS?)
721 HA HA, DO I WANT TO ... HAH.
722 OK, WELL, AT THIS POINT,
723 OK, WELL, I NEVER ENJOY TESTING 187 TRANSACTIONS.
724 (WELL, FOR WHAT IT'S WORTH, YOU WON'T BE DOING IT.)
725 I KNOW, THE COMPUTER - WELL,
726 WELL, I GUESS WE'RE GETTING INTO THE AMOUNT OF
CONFIDENCE I'LL
727 BE GETTING FROM TESTING THAT MANY
728 VERSUS -
729 DOING IT.
730 I'M JUST GOING TO TAKE THE COST
731 SO WHAT DO I HAVE TO DETERMINE,
732 187.
733 HOPEFULLY, I'VE ALREADY DECIDED THAT THOSE ARE THE
LIMITS THAT
734 I WANT.
735 NOW, I JUST HAVE TO DETERMINE WHETHER THE CONFIDENCE
IS WORTH THE
736 COST.
737 WELL, I GUESS MY,
738 MY EXPERIENCE IS THAT WE ALWAYS TEND TO AUDIT A LITTLE
HARDER
739 ON THE FIRST TIME SITUATION.
740 I GUESS I'M MORE RECEPTIVE TO DOING THAT AMOUNT OF
WORK WITH THIS
741 FIRST TIME.
742 ...
743 (OKIE DOKE, THIS IS YOUR FIRST SAMPLE REPORT THAT I'LL
SHOW YOU.
744 YOU SELECTED A SAMPLE SIZE OF 187, THE SAMPLE OUTCOME
- THERE
745 WERE 9 DEFECTIVE TRANSACTIONS. YOUR MANAGER CHOSE
THAT THE TRUE
STATE OF INTERNAL CONTROLS WAS B, IN FACT IT WAS. SO YOUR NET PAYOFF IS $62.60. YOUR CUMULATIVE PAYOFF IS $62.60, SINCE THIS IS THE FIRST TRIAL.) OK, NOW LET ME SEE IF I UNDERSTAND WHAT THIS ALL MEANS. THE INFERENCE MADE BY YOUR MANAGER BASED ON THE SAMPLE SIZE; THE TRUE STATE OF COMPLIANCE. NOW WHAT WE'RE SAYING IS THAT SO I PICK THE SAMPLE SIZE... (YOU PICK THE SAMPLE SIZE, YOU CONDUCTED THE COMPLIANCE TEST, AND OBSERVED A SAMPLE OUTCOME OF 9. YOU REPORTED THE SAMPLE SIZE AND THE SAMPLE OUTCOME TO YOUR MANAGER. YOUR MANAGER INFERRED...) HE'S ANALYZED THAT. (RIGHT... THE POPULATION ERROR RATE. (YOUR MANAGER INFERRED THAT THE TRUE STATE OF INTERNAL CONTROLS WAS THIS...) 3 AND 7%, WHICH IT WAS. (WHICH IT WAS. SO YOUR NET PAYOFF WAS $100, MINUS 20 CENTS PER TRANSACTION SAMPLED.) OK. YOU WIN ONE HUNDRED BUCKS IF YOUR MANAGER'S DECISION CORRESPONDS WITH THE TRUE STATE OF INTERNAL CONTROLS. IN ALL CASES, YOU WILL BE ASSESSED 20 CENTS PER TRANSACTION SAMPLED. OK. IF YOUR MANAGER ERBS BY ONE LEVEL IN EITHER DIRECTION, NO REWARD. YOU CAN LOSE $4000. IF IT IS A OR B, AND THE TRUE STATE OF INTERNAL CONTROLS IS D. OK. SO BASED ON MY SAMPLE SIZE AND THE SAMPLE OUTCOME, HE WAS ABLE TO DETERMINE THE ACTUAL. (RIGHT, NOW IF YOU WANT TO REREAD RIGHT HERE, IT JUST TELLS YOU ABOUT THE REST OF THE EXPERIMENT.) OK. YOU HAVE THE OPPORTUNITY TO LEARN FROM EXPERIENCE BECAUSE THE EXPERIMENT WILL BE REPEATED A LARGE NUMBER OF TIMES. FOR EACH REPEAT OF THE EXPERIMENT, ASSUME YOU ARE IN CHARGE OF
THE AUDIT OF ANOTHER COMPANY, BUT EACH NEW COMPANY IS IDENTICAL

SUBJECT I CONTINUES READING EXPERIMENT

... HERE IS A SAMPLE OF 100 ERROR RATES GENERATED BY THE COMPUTER.

OK, SO IT'S KINDA SHOWING US HOW IT JUST RANDOMLY PICKS THESE OUT.

(RIGHT, SO BASICALLY, YOU HAVE THE CHANCE TO RECLINE THAT

ORIGINAL ESTIMATE OF 187).

OK, I'VE GOT THE SAME COMPANY,

OR I HAVE A DIFFERENT COMPANY THAT'S SIMILAR TO THIS,

BUT NOW I HAVE THE EXPERIENCE OF THIS COMPANY,

WITH THAT SAMPLE SIZE AND THAT SAMPLE OUTCOME.

BUT THE FACT STILL REMAINS THAT THE ACTUAL ERROR RATE IS GOING

TO CHANGE

OR COULD CHANGE.

SO I HAVE TO, OK.

SAMPLE SIZE...

WELL LET ME SEE, WHAT DID I DO HERE,

I ENDED UP BLOWING $40, HM.

SAMPLE SIZE, A NEW COMPANY.

SAMPLE OUTCOME WITH 9 ERRORS IN 187.

ERROR RATE IS BETWEEN 3% AND 7%.

INTERNAL CONTROL PROCEDURES ARE SATISFACTORY,

EXCEPT FOR SOME MINOR DEPARTURES

AND SUBSTANTIVE TESTING WILL BE AT A HIGHER LEVEL.

OK.

SO WHAT DO I WANT TO DO WITH MY SAMPLE SIZE?

ALRIGHT, I PICKED A 95% CONFIDENCE LEVEL

AND I PICKED 5% RISK, WHICH I KINDA LIKE.

HM.

<UNCLEAR> YOUR PART OF THE LINE

GOES 95, 90, 85,

(YOU CAN USE ANY CONFIDENCE LEVEL YOU WANT)

IS THAT RIGHT?

92 OR 93?

(THAT'S RIGHT)

SO I HAVE A 95% CONFIDENCE LEVEL.

PROBABLY, HM.

187 DOWN TO 150.

WELL, COULD I SEE WHAT THE SAMPLE SIZE WOULD BE USING 93% AND 7%.

(SURE)

(SO THE ONLY THING YOU WANT TO CHANGE IS THE CONFIDENCE LEVEL

TO 93%)

(TABLE VALUE WOULD BE 151)

OK, NUMBER OF TRANSACTIONS,

POPULATION ERROR RATE.

THIS NUMBER MAY BE THOUGHT OF AS THE PROBABILITY...

SO, LET'S SEE.
So if I divided the 187 into 9 that would give me a sample error rate.

Is that correct?

So I had a 4.8% sample error rate.

And I said 4.5 (Right)

OK, 4.8% sample rate, hm.

4.8%

And population error rate was actually between 3 and 7.

Get the population error rate under control, that would be between 12 and 20, yuck.

I can't see that population error rate.

4.6, that seems to be in the ball park.

Alright. When,

Well, I guess, at this point I would probably lower the sample size down to 150.

But I would probably keep my expected error rate about the same.

(You want to lower the sample size?, you want to just pick 150)

Yeah, the same.

(Similar results)

HM, sample size 150,

Still got 9.

Is this going to change, this sample outcome?

(Yes)

OK.

Decision B, State B.

Payoff this trial, 70 bucks.

Made $8 this time over the last time, huh?

Sample size 151,

Maximum allowed.

OK, 150, so,

Tested at 150, and 187.

The manager still picked the good number here.

If you miss it by one level,

No reward or penalty will accrue to you.

So I'm at a 93% level approximately.

70 bucks this trial.

You will win $100 bucks if the manager's decision corresponds with the true state of internal controls.

Now,

I gotta be more efficient here.

Got 150 and I had 9 this time,

This could change at any time.

Is there any set number of times that we do this, oh?

Well there is a set number, it's a large number)
Uh, do I need to keep going faster or?

(no, um, I want you to do it at whatever pace you think is appropriate.)

If we get to the point where you get tired, but we're not at the end, we'll stop.

Now I got between 3 and 7.

9 divided by 150 equals, mm, 6%.

Getting a little closer here.

So we were between 6% when I did 150, between 7 and 12% there are significant departures from internal control procedures and extensive substantive testing is necessary.

Oh.

The question is if I lower my sample size, will the sample outcome change?

So basically, if I had said, if my expected error rate had been anywhere 3 and 7% on the last time, the results would have been the same, does that have any...

(I'm not sure I understand the question)

Ok,

(your sample size only impacts on the sample size selection)

Oh, ok, well, that's...

Ok, 150 items...

If I reduced that,

I'm trying to think of what the impact is if I changed my, uh,

the 4.5% expected error rate.

That's the population error rate though.

I had a 4.8%, the first time, sample error, and I had a 6% rate the second time.

But I had the exact same number of errors in both cases.

Exact same number of errors in both cases, mm.

So, if I were to raise my expected rate, and lower...

3%.

Ok, why don't we put,

I have to see what the sample size would be if I raise my 4.5%,

to let's make it 6.

Let's make it 6.

(6%)

Yeah,

(What confidence level do you want?)

I'd like to keep the 93%.
(OK)

... (YOU WOULD HAVE A SAMPLE SIZE OF 1001)

1001.

(YES)

OH MY.

THAT'S SAYING 93% SAMPLE RATE,

BUT AN EXPECTED ERROR RATE OF 6%.

(RIGHT, AND A TOLERANCE LEVEL OF 7%)

HH.

I'VE GOT MY TOLERANCE LEVEL DOWN.

I DON'T THINK I'D LIKE THAT.

(NO)

I THINK THAT'S A LITTLE TOO MANY HERE.

150, SAMPLE SIZE, 150.

150 ITEMS.

WELL LET'S SEE, UM.

I'M AT BAT NUMBER 3 HERE.

WHAT WOULD MY SAMPLE SIZE BE AT A 90% CONFIDENCE LEVEL?

927, HMPH.

I'VE ONLY DONE IT TWICE.

I'D KINDA LIKE TO STAY AROUND 93%.

I DON'T FEEL LIKE I WANNA.

I DON'T WANT TO GO ANY LOWER THAN THAT, I DON'T THINK.

MY SAMPLE ERROR WAS 4.8 AND THEN IT WAS 6.

SO IF I GET MUCH LOWER, IT MIGHT BE A LITTLE LOW.

SO IF I EXPECT THE ERROR RATE IS PROBABLY CLOSER TO 5 OR 6%,

I'LL LET MY MAXIMUM,

THAT I WOULD BE WILLING.

IF I KEEP THAT AT THE SAME,

THEN I END UP WITH A VERY LARGE SAMPLE SIZE.

SO, IF I ROSE THAT,

I WOULD PROBABLY GET A LOWER SAMPLE SIZE

BUT THEN I'M GIVING UP THE GOOD STUFF, UM.

WELL, LET'S SEE.

SHOULD I RAISE THE NEXT TOLERABLE ERROR RATE?

WELL, LET'S SEE.

TO BE MORE COST EFFICIENT,

I WOULD LIKE TO REDUCE MY SAMPLE SIZE.

BUT STILL GET A 93% CONFIDENCE LEVEL.

REDUCE MY SAMPLE SIZE, GET A 93% CONFIDENCE LEVEL

BUT I STILL THINK THERE'D PROBABLY BE A GOOD 6% ERROR RATE.

TALKING ABOUT POPULATION, SO I'VE CHANGED THAT.

WELL, I'LL TELL YOU, I THINK IF I WERE GOING TO DO IT AGAIN,

I'VE PROBABLY STICK WITH THE NUMBERS I HAD THE LAST TIME.

93% AND THE 4.5.

IN OTHER WORDS, I'D PROBABLY DO THIS SAMPLE AGAIN.

(150)

(100 WOULD HAVE A SAMPLE SIZE OF 1001)

1001.

(OK)

...
OK, WE HAD 12.

BUT THE STATE'S STILL B.

SO WHAT DOES THAT MEAN.

WHEN I DID 150 AND HAD 12 ERRORS.

WELL, THE MANAGER SAID IT WAS C.

POPULATION - 7 AND 12%.

INTERNAL CONTROL PROCEDURES.

AND EXTENSIVE SUBSTANTIVE TESTING IS NECESSARY.

BUT IN FACT, CONTROLS ARE STILL SATISFACTORY.

WHAT DOES THAT TELL ME.

HAD 9 AND THEN CAME UP WITH 12.

SO THAT DOES THAT TELL ME.

HAD 9 AND THEN CAME UP WITH 12.

SHOULD HAVE STOPPED WHILE I WAS AHEAD LAST TIME HUH?

12 WRONG.

WELL, LET ME SEE, WHAT DID I END UP WITH?

12 DIVIDED BY 50.

HM, DIVIDED BY 150.

I WAS UP TO 8% THAT TIME.

I THINK I'M GOING TO HAVE TO RAISE THAT.

SO LET'S SEE,

IF I RAISE MY EXPECTED ERROR RATE,

TOWARDS 6 OR 7%.

THAT'S WHAT I EXPECT IT TO BE.

WHAT, OK.

TRY ANOTHER SAMPLE SIZE HERE.

I'D LIKE TO KEEP THE 93% CONFIDENCE LEVEL,

I'D LIKE TO RAISE MY EXPECTED ERROR RATE TO 6%.

AND RAISE THE TOLERABLE MAXIMUM TO 8%.

(OK).

(309)

HM.

309 HUH? HM.

THAT SEEMS TOO MANY TO ME.

I EXPECT 6% ERROR RATE.

HAD 9 TWICE AND 12.

187 AND 150.

I HAD 12 ERRORS LAST TIME

AND 9 THE TWO BEFORE.

LESS THAN OR EQUAL TO 3%.

I'M STILL SEEING 4.

OK, SO I'M SAYING SO FAR, THIS KIND OF SYSTEM HAS

GIVEN ME 9

TO 12 ERRORS.

IF I'D USED 187,

AND I HAD GOTTEN 12,

I PROBABLY WOULD HIT IT UP ON A B AGAIN.

NOW LET'S SEE HERE.

IF I USE 150,

I COME OUT WITH 4 ERRORS,

GIVE ME 5 ERRORS.

IT GIVES ME 3.3%.

SO I'VE GOT 5 ERRORS.

IF I GET 10 ERRORS,
THAT GIVES ME 6.6.
IF I GET 11 ERRORS, 7.3.
SO I'M LOCKED IN BETWEEN 5 AND 10 ERRORS.
COME UP WITH THAT IN B.
SO LET'S SEE,
5 TO 10 ERRORS WOULD GIVE ME PROBABLY THAT IT'S GOING TO
BE A B,
OR ADEQUATE.
I USED 187,
5 ERRORS,
GIVES ME 2.6,
SO THAT WOULD HAVE BEEN A 3.
DIVIDED BY 150,
equal TO <UNCLEAR>
3.2.
SO IF I'M USING 187,
6 ERRORS WOULD GIVE ME INTO THE B CATEGORY.
WHAT ARE YOU CALCULATING?)
I'M TRYING TO SEE THE RANGE OF ERRORS.
I WAS JUST TAKING AN ESTIMATED NUMBER OF ERRORS
DIVIDING IT INTO MY SAMPLE SIZE.
(THAT WAS 9 DIVIDED BY 187)
YEAH.
TRY 12,
DIVIDED BY 187, GIVES ME 6.
12 DIVIDED...
3.2, GIVES ME ABOUT 6 ERRORS.
12 DIVIDED BY 187,
GIVES ME 6, 13, LOOKS LIKE 6 TO 13 ERRORS.
DON'T TELL ME THAT WITH 187,
WHILE IF I'M USING 4 IT WAS 5 TO 10 ERRORS
WOULD TELL ME IT WOULD BE A B SITUATION.
SO, LET'S SEE,
6 TO 13 ERRORS.
VERSUS 5 TO 10.
WE GOT 9, 9, AND 12.
SEEM TO BE UP IN...
I DON'T KNOW IF IT'S GOING TO CHANGE,
BUT SO FAR WE'VE BEEN IN THE 9 - 12 CATEGORY.
187,
OH BOY,
SO, IF I USE ANYWHERE BETWEEN THESE 2 NUMBERS,
I'M WITHIN 93 TO 95 % CONFIDENCE.
HM.
I HIT TOO MANY THAT TIME AND IT THREW ME OFF.
I HIT TOO FEW ON THAT ONE AND IT WILL THROW ME OFF.
I USE 160 AS A SAMPLE SIZE,
- 5 TO 11.
TEN MORE WILL GIVE ME 11.
I DON'T KNOW WHAT THE REAL SITUATION COULD BE.
LET'S TRY 160.
(OK.)
HH, DOUBLE A'S.
1076 PAYOFF THIS TRIAL $68.
1077 ALRIGHT, NOW WE'VE GOT 4 OF THEM.
1078 160 OVER 3.
1079 HIT IT 3 TIMES NOW.
1080 WELL LET'S SEE,
1081 LOOKING AT MY RANGE NOW,
1082 160,
1083 WELL, 160.
1084 I'LL TRY 160 AGAIN.
1085 ---
1086 WELL, DID IT AGAIN.
1087 OK, SO WE HIT A AGAIN.
1088 THERE WERE ONLY 2 ERRORS.
1089 WE GOT THE SAME SCORE.
1090 160, SO.
1091 I'LL TRY 160 AGAIN.
1092 ---
1093 WE HIT B.
1094 WE GOT 8 THAT TIME.
1095 WE'RE JUST MOVING ALONG HERE, AREN'T WE?
1096 WELL, I'VE USED 160 3 TIMES NOW,
1097 AND I HIT IT ON THE NOSE EACH TIME.
1098 I USED 150 TWICE,
1099 GOT IT ONCE AND BLEW IT ONCE.
1100 AND I'VE USED 187 AND LUCKED OUT.
1101 I'LL TRY 160 AGAIN.
1102 ---
1103 WHEN, WE HAD 15 THAT TIME.
1104 PAYOFF THAT TIME 32, MINUS.
1105 DECISION WAS C BUT THE STATE WAS STILL B.
1106 SAMPLE OUTCOME - 15,
1107 15 HUH?
1108 5 ERRORS DIVided BY 160 WAS 3.1.
1109 SO 5 ERRORS WOULD GIVE ME B, 211.
1110 THAT GIVES ME B,
1111 BELOW THAT GIVES ME A.
1112 C GIVES ME 12.
1113 10.5%.
1114 NOW LET'S SEE 18, 11.2.
1115 19 DIVided BY 160 GAVE ME 8.
1116 SO MY C RANGE IS 12 TO 19 ERRORS.
1117 I HAD 15.
1118 THE ACTUAL STATE WAS B THOUGH.
1119 THAT SEEMS A LITTLE WEIRD.
1120 I HAD 150, WELL, SEEMS TO ME
1121 THAT IN THIS PARTICULAR CASE,
1122 WE HAD A FLUKE,
1123 OR WE,
1124 IT SEEMS THAT THE SAMPLE I SELECTED WAS SKEWED,
1125 IT HAD TOO MANY ERRORS.
1126 A NORMAL SITUATION WOULD HAVE CALLED FOR,
OR THE ACTUAL POPULATION, I GUESS, THE ACTUAL STATE HERE,
B, SAYS IT'S BETWEEN 3 AND 7X.
WHICH MEANS THAT I WOULD HAVE NORMALLY HAD BETWEEN 5 &
11 ERRORS.
IN A NORMAL SAMPLE.
I TEND TO THINK THAT THAT'S A FLUKE,
THAT THAT'S MORE THAN SELECTION THAN THE ACTUAL NUMBER THAT
WAS PICKED;
WHERE WHEN I WAS USING 150,
I ENDDED UP WITH 12.
THE ACTUAL RANGE WAS,
THE ACTUAL STATE OF BEING IN B WOULD HAVE BEEN 5 TO 10
ERRORS.
I ENDDED UP WITH 12,
WHICH LED ME TO BEING WRONG.
IF I'D USED B,
WELL, LET'S.
I HAD TO DO EXTRA WORK THERE.
I'VE ENDDED UP WITH 15,
I DON'T THINK THAT I'M GOING TO COVER MYSELF WHEN I
HIT 15.
I THINK I'LL TRY 160 AGAIN.
---
A AND A.
WELL, OK, FROM WHAT I'M SEEING NOW,
THE RANGE OF ERRORS SEEMS TO BE PRETTY WELL COINCIDING
WITH THIS
AREA,
160.
I HAD ONE SITUATION WHERE I HAD AN UNUSUAL NUMBER OF
ERRORS PICKED OUT.
WHEN I PICK
IF I PICK 150,
TO GET A B SITUATION,
I HAVE 5 TO 10 ERRORS.
IF I PICK 160, I GET 5 TO 11,
THAT GIVES ME LIKE ONE EXTRA ERROR.
IT SEEMS SO FAR TO BE COMING UP WITH THE TRUE STATE.
SO MY QUESTION IS,
DO TEN ITEMS,
THEY DON'T MAKE MUCH DIFFERENCE.
TIMewise, I DON'T THINK THAT 10 ITEMS WOULD TAKE UP A
LOT OF COST.
TEN EXTRA INVOICES OR TEN EXTRA CHECKS TO TEST THEM.
THE QUESTION IS DO I GET ENOUGH CONFIDENCE,
OR EXTRA ASSURANCE FROM,
BY TAKING THOSE 10 EXTRA ONES THAT IT'S GONNA HELP ME
ZERO IN
ON THE RIGHT - THE TRUE STATE.
SO FAR THE RANGE HAS BEEN 9,9,8, OR 12.
WE HIT SOME REAL LOW ONES - 3, 2, 8, 15, AND 2.

SO IN EITHER CASE,

IF I HIT 15 OR IF I HIT 12,

I STILL WOULD HAVE BEEN WRONG.

THE ONLY SITUATION IS IF I HIT ON 11.

BUT I THINK - TO ME,

IT WOULD BE WORTH DOING 10 EXTRA INVOICES

TO AVOID COMING UP WITH A CONCLUSION THAT, YOU KNOW,

YOUR INTERNAL CONTROLS ARE INADEQUATE,

THOUGH I SUPPOSE THE POSSIBILITY ALWAYS WORKS THE

OTHER WAY TOO.

I'LL STICK WITH 160.

(YOU'VE BEEN DOING THIS A LONG TIME, LET ME ASK YOU,

HOW MUCH

LONGER ARE YOU WILLING TO DO IT)

CAN YOU TELL ME WHAT TIME IT IS ACTUALLY?

(YEAH, IT'S 10:43)

OH BOY, PROBABLY TILL AROUND 11:00 I THINK.

(OK, LET'S CALL IT QUILS AROUND 11:00)

(I APPRECIATE ALL THE TIME YOU'RE TAKING)

---

DID IT AGAIN,

I KINDA LIKE 160,

AS IF YOU COULDN'T TELL.

(OK, YOU WANT IT AGAIN?)

YEAH.

---

C AND B.

WE HIT 13.

WELL LET'S LOOK BACK HERE,

HOW MANY TIMES HAVE WE DONE THIS.

OK, WE'VE USED IT 1, 2, 3, 4, 5, 6, 7 TIMES NOW,

AND WE'VE BEEN WRONG TWICE.

THAT GIVES ME 2 OUT OF 5, ER, 2 OUT OF 7 I'VE BEEN

WRONG.

BOTH TIMES IT'S BEEN A LITTLE HIGHER.

THE NUMBER OF ERRORS THAT HAVE COME OUT IN MY SAMPLE

HAVE

BEEN HIGHER THAN WHAT THE POPULATION SAYS.

WHAT CAN I DO ABOUT THAT.

WELL, DO I WANT TO EXPAND MY SAMPLE SIZE,

SO I CAN BUMP THAT UP A LITTLE?

I DON'T KNOW WHY.

LET'S SEE.

160.

187.

WELL, 12, 13, AND 15 HAVE BEEN WRONG.

WELL, IF I GET IN A SITUATION WHERE I'VE EXPANDED MY

"B SCOPE"

TOO FAR, YOU KNOW.

IF I'M SAYING 5 TO 13, OR 14, OR 15 ERRORS,

WOULD GIVE ME THIS B ASSUMPTION,
I would tend to think that of the actual 11, 12, 13 errors, I don't want to expand my sample size so much that it would probably throw you down into C, so I don't think - I don't want to expand my sample size so much that it would probably throw you down into C, when I get these 11, 12, 13 errors, it's telling - you know, I got a B situation, because I think in reality, you'd probably be more towards a C. where the controls wouldn't be right. I don't know, 13, but then you're testing more. 2 out of 7 times I've been wrong. 2 out of 7 times. both times I've concluded that I couldn't rely on the controls. significant departures, extensive substantive testing. so in the end, it's costing me more probably, doing substantive work, than if I'd done interim work. I seem to get more errors on my sample than what the population would show. which tends to say, if I increase my size, I can bump that down a little. I'll try - let me look at 170 here. what kind of range does that give me. 4 divided by 170, 5 divided by 170 gives me 2.9. 2 out of 7 times it's been wrong. it keeps pulling out 13, 12, and 15. when I raise my size, try 160 again. ... b and b again. sample outcome 11. that was a good limit there. so now I've got 2 out of 8 wrong. I'll try 160 again. ... did it again. I'd have to say at this point, that I'd probably keep using 160 for a while.
BIBLIOGRAPHY


Ungson, G. R., Braunstein, D. M. and Hall, P. D.,
"Managerial Information Processing: A Research Review,"

White, P., "Limitations on Internal Reports of Internal Events: A Refutation of Nisbett and Wilson and of Bem,"

