An Examination of Semiotic Theories of Accounting Accruals.

Harlan Lynn Etheridge

Louisiana State University and Agricultural & Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_disstheses

Recommended Citation

https://digitalcommons.lsu.edu/gradschool_disstheses/5230

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Historical Dissertations and Theses by an authorized administrator of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
An examination of semiotic theories of accounting accruals

Etheridge, Harlan Lynn, Ph.D.
The Louisiana State University and Agricultural and Mechanical Col., 1991
AN EXAMINATION OF SEMIOTIC THEORIES OF ACCOUNTING ACCRUALS

A Dissertation

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Doctor of Philosophy in

The Department of Accounting

by

Harlan Lynn Etheridge
B.S., McNeese State University, 1981

December 1991
ACKNOWLEDGEMENTS

This study reflects the knowledge, advice, and influence of a multitude of individuals. I am grateful to my dissertation committee: Dr. Bart Hartman, Dr. Barbara Apostolou, Dr. Ye-Sho Chen, Dr. Daryl Guffey, and Dr. Ram Sriram. The exceptional guidance they provided eased the arduous dissertation process. I am especially indebted to Dr. Bart Hartman, chairman of my dissertation committee, for the encouragement and counsel he provided throughout my doctoral program.

My appreciation is extended to Dr. Vince Brenner for his assistance in obtaining a neural network for use in my dissertation and for his support during my studies at LSU. I am indebted to Dr. Charles Monlezun for his invaluable help with the statistics of this study. Many thanks are due to Ihssan Alkadi, Ghassan Alkadi, and the other staff at the System Network Computer Center for their technical assistance.

I extend my deepest gratitude to my family for their love and encouragement provided throughout my doctoral studies. I owe much to my parents, Joyce and Thomas H. Etheridge. Their commitment to education and knowledge instilled in me the desire to learn. It is this desire that guided me through the doctoral program.

Finally, I thank the Lord God for all the blessings He has given me. Without His help, none of my accomplishments, including this dissertation, would have been possible.
To my parents, Joyce and Thomas H. Etheridge
# TABLE OF CONTENTS

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

**DEDICATION** ................................................... iii

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

**ABSTRACT** ...................................................... x

**Chapter**

1. **INTRODUCTION** .......................................... 1

   Formulation of the Research Questions .................. 3
   Semiotics ...................................................... 5
   Research Questions .......................................... 7

   Importance of the Research Questions .................. 7

   Method ......................................................... 9

   Objectives, Scope, and Limitations ..................... 16

   Contributions of the Study ............................... 18

   Summary ...................................................... 19

2. **LITERATURE REVIEW AND THEORY** .................. 20

   Literature Review .......................................... 20
   Syntactic Information Theory ............................ 20
   Semantic Information Theory ............................. 24
   Information Content of Accounting Accruals and Cash Flows ........................................ 28
   Discussion of the Literature ............................. 31
   Syntactic Information Theory ............................ 31
   Semantic Information Theory ............................. 32
   Information Content of Accounting Accruals and Cash Flows ........................................ 32

   Contribution of this Research to the Literature ........ 34

   Theory ........................................................ 34
   Theories of Accounting Accruals ........................ 34
   Data, Information, and Knowledge ...................... 35
   Semiotic Concepts .......................................... 35
   Theory of the Functions of Accounting Accruals .......... 46
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Independent Variables</td>
<td>13</td>
</tr>
<tr>
<td>2. Research Design</td>
<td>14</td>
</tr>
<tr>
<td>3. Categories of Data Type</td>
<td>57</td>
</tr>
<tr>
<td>4. Categories of Financial Statement Time Frame</td>
<td>57</td>
</tr>
<tr>
<td>5. Dependent Variables</td>
<td>60</td>
</tr>
<tr>
<td>6. Research Design</td>
<td>60</td>
</tr>
<tr>
<td>7. Data Items</td>
<td>71</td>
</tr>
<tr>
<td>8. Grouping of Data Items by Data Types</td>
<td>72</td>
</tr>
<tr>
<td>9. Sample Selection</td>
<td>73</td>
</tr>
<tr>
<td>10. Neural Network Performance Test Results</td>
<td>76</td>
</tr>
<tr>
<td>11. Neural Network Performance Measures and Error Metrics--Quarterly Cash-Flow Data</td>
<td>76</td>
</tr>
<tr>
<td>15. Critical Values for Hypothesis Tests</td>
<td>81</td>
</tr>
<tr>
<td>16. Estimated Error Metric Means</td>
<td>84</td>
</tr>
<tr>
<td>17. Values of Estimated Error Metric Means</td>
<td>85</td>
</tr>
<tr>
<td>18. Means Examined</td>
<td>86</td>
</tr>
</tbody>
</table>
19. Contrasts and Test Statistics ............................................................ 87
20. Results of Hypothesis Tests ............................................................. 93
ABSTRACT

Two theories of accounting accruals form the core of this study. The theory of the functions of accounting accruals is induced from observation of accounting accruals. One type of accounting accrual reflects incomplete transactions under a system of cash receipts and cash disbursements. The function of these accruals is to convey data about economic events associated with incomplete cash transactions. These accruals are referred to as syntactically redundant accruals. The second type of accounting accrual presents messages in a different format than their counterparts in a system of cash receipts and cash disbursements. The function of these accruals is to enable accounting data users to perceive messages that may be unclear when signaled under a system of cash receipts and cash disbursements. These accruals are referred to as semantically redundant accruals. The theory of the functions of accounting accruals states that accounting accruals are syntactically redundant or semantically redundant in function. The theory of the pragmatic information of accounting accruals states that accounting accruals contain pragmatic information because of their functions. A hypothesis associated with these theories asserts that the pragmatic information of accounting accruals decreases as the financial statement time frame increases.

Cash-flow and accrual data were obtained from COMPUSTAT and provided to a neural network to generate forecasts of cash flows from operations. The forecasted cash flows were used to calculate an error
metric that functioned as a measure of pragmatic information. A variant of the Bonferroni procedure was used to detect differences between means of error metrics and to test research hypotheses. The results indicated that syntactically and semantically redundant accounting accruals contain pragmatic information. Hypothesis tests, however, provided evidence that only annual accounting accruals contain pragmatic information. Hypothesis tests also indicated that no difference exists between the pragmatic information contained in syntactically redundant accounting accruals and the pragmatic information contained in semantically redundant accounting accruals. Finally, the results of hypothesis tests presented no evidence of a negative relationship between the pragmatic information of accounting accruals and the financial statement time frame.
CHAPTER 1
INTRODUCTION

Accounting involves communication. Accountants communicate information to third parties in various ways including financial statements, auditor reports, and management reports. Despite the importance of the process of communication to accounting, accounting researchers have not focused on the development of formal theories of the accounting communicative process. Accounting researchers either have centered their research efforts on individual perception and processing of financial information or have concentrated on the examination of the information content of accounting data. However, the emphasis placed on the communication aspect of accounting by the Financial Accounting Standards Board (FASB) in the Statements of Financial Accounting Concepts (SFACs) and other promulgations suggests more research is desirable on communication in the accounting process. For example, SFAC No. 1 (FASB 1978, 4) states that:

Financial statements . . . are a principal means of communicating accounting information . . .

Financial reporting includes not only financial statements but also other means of communicating information that relates, directly or indirectly, to the information provided by the accounting system . . .

---

1 See, for examples, Schroeder, Driver, and Streufert (1967); Driver and Mock (1975); and Libby (1981).

2 Examples of this area of research include Harmon (1984); Schaefer and Kennelley (1986); Wilson (1986, 1987); Bowen, Burgstahler, and Daley (1987); Bublitz, Frecka, and McKeown (1987); and Ismail and Kim (1989).
Formal theories of communication enabling the description and analysis of languages are used extensively in fields such as psychology and linguistics. Accounting researchers began using these theories, especially syntactic information theory\(^3\) (i.e., Shannon and Weaver 1949) and semantic information theory\(^4\), in the late 1960s and early 1970s. However, accounting researchers directed their efforts more towards employing communication theories to analyze financial statements than to developing specific theories of the accounting communicative process.

Because accounting may be categorized as a type of language (Kleerekoper 1963; Guthrie 1972; Jain 1973; Belkaoui 1980), examination of the communication features of accounting and development of theories that explain and describe these features should be a part of the evolution of the accounting discipline. Theories that describe the accounting communicative process should foster refinement of the means of communication in accounting--accounting reports. In particular, theories of the accounting communicative process should encourage the examination of the components of accounting reports. This examination of accounting reports may lead to a more thorough understanding of the nature of the accounting communicative process. Accounting accruals are one of the components of accounting reports that could be examined with communication theory.

\(^3\) Examples include Lev (1969a, 1969b, 1970); Lee and Bedford (1969); Ronen and Falk (1973); Abdel-khalik (1974); Babich (1975); Gorelik (1975); Belkaoui (1976); and Pendlebury (1980).

\(^4\) See, for examples, Nakano (1972); Haried (1972); Jain (1973); Oliver (1974); and Belkaoui (1980).
Many studies examined the information content of accrual accounting data versus cash flows or funds flows. However, these studies did not have the advantage of theory support. Bedford (1965) commented that the controversy surrounding the benefits of accrual accounting data versus the benefits of cash-flow data would continue until a general theory of accruals was available. Bedford (1965, 29) asserted that:

... there appears to be no satisfactory explanation of the nature of the accrual concept, nor are sufficient reasons set forth explaining why it should be used, and there exists no overall authoritative statement indicating where the concept should be used, or how the accrual is to be computed.

The purpose of this study is to develop and test semiotic theories of accounting accruals. Semiotics encompasses syntactic information theory, semantic information theory, and pragmatic information theory. This study uses concepts from syntactic information theory and semantic information theory to explain the functions of accounting accruals. Pragmatic information theory is used to define and examine the value of accounting accruals.

Formulation of the Research Questions

Scrutiny of studies examining the information content of accrual earnings (or the incremental information content of accrual accounting data) and the information content of cash flows or funds flows reveals inconsistencies in the results. Two possible rationales exist for this

---

5 Fawcett and Downs (1986, vii) comment that "... the paucity of recognizable theory in some disciplines is due to investigators' failure to be explicit about the theoretical components of their studies." It is possible that past studies examining the information content of accounting accruals and/or cash flows were based on unstated theories of accruals and/or cash flows. However, no previous study has explicitly expressed a formal theory of accounting accruals.
situation. First, a lack of theory leads to inconsistency of variables tested, resulting in restricted generalizability across variables. With a guiding theory, the use of multiple constructs (and variables) may be preferable to the utilization of the same construct (or variable) across studies. The absence of a theory to direct research, however, reduces the ability of researchers to compare and contrast the results of the various studies. Second, some of the statistical methods used are not suitable in situations where the data contains a high level of noise. The conversion of accrual accounting data to cash-flow data is always less than perfect and consequently, the resulting cash-flow data contains noise. Databases, such as COMPUSTAT, frequently used in these studies also are sources of noise. The traditional statistical techniques employed in the accrual versus cash-flow information content studies may yield inaccurate and/or non-interpretable results in such situations.

No previous research has proposed a formal theory of accounting accruals. Theories of accounting accruals based upon semiotics are central

---

6 The use and testing of multiple constructs reduces the threat from mono-operation bias and thus, enhances construct validity across studies (Cook and Campbell 1979).

7 See Cook and Campbell (1979) for a discussion of the limitations of archival data.

8 As the noise level in data increases, the size of the error term of models developed with traditional statistical techniques such as regression analysis and ANOVA increases. As the error term enlarges, the potential for heteroscedasticity grows. If heteroscedasticity is present in the data, the estimated coefficients that result from the use of these models will be unbiased but will not be the most efficient estimators of the true coefficients. More importantly, moderate levels of heteroscedasticity can seriously affect the alpha level and power of tests using traditional statistical techniques and will invalidate the use of the tests (Berenson, Levine, and Goldstein 1983, 66).
to this study. However, before the research questions are expressed, concepts essential to semiotics are defined.

Semiotics

_Syntactic information theory_ is concerned with the probability of occurrence of events. The greater the probability of an event occurring, the smaller the syntactic information contained in the report of its occurrence. Conversely, the smaller the probability of an event occurring, the larger the syntactic information contained in the report of its occurrence (Bostwick 1968).

_Semantic information theory_ involves the relationship between signs and referents. Consequently, semantic information theory is used to investigate and analyze the meaning of data. It is impractical to measure the semantic information of data in real-world situations, e.g., object-languages such as English or French. However, the concepts of semantics can be used to describe object-languages.

_Pragmatic information theory_ is concerned with the value of data to individuals. The amount of pragmatic information in a message is dependent upon the individual receiving the message. The pragmatic information contained in a message must be considered in the context of the individual receiving the message.

Whenever there exists a series of possible events, the expected syntactic information of a message related to this series of events is called _entropy_. Entropy is zero when the probability of occurrence of any event (of all possible events) is 1 and increases to a maximum when the

---

9 See Cherry (1978) for a discussion of appropriate uses of measures of semantic information content.
probabilities of all possible events are equal. The amount of entropy also increases as the number of possible events increases.

Syntactic noise is described as distortion of the communication signals. Semantic noise, in contrast, can be defined as distortion of the message transmitted. In other words, a message is transmitted, but is misunderstood by the receiver. Noise requires the need for additional effort in the determination of the original message, and noisy signals never reduce the amount of uncertainty surrounding a communication to zero (Nakano 1972).

Conditional entropy is the "... average amount of remaining uncertainty after receiving a noisy signal..." (Nakano 1972, 695). The average amount of syntactic information that can be communicated on a noisy channel is the difference between the entropy and the conditional entropy of a given information source and is called equivocation (Nakano 1972).

Redundancy is a state where the information or expected information of a message is lower than its capacity for information due to data in the message that conveys no information. It seems that, at first glance, redundancy serves no useful purpose. However, the presence of redundancy in a noisy channel makes the interpretation of a message possible (Hyvärinen 1988).

Two types of redundancy are differentiated: syntactic redundancy and semantic redundancy. Syntactic redundancy is the existence of coding

---

10 This definition reflects the traditional philosophy of redundant data in syntactic information theory. In this study, the concept of syntactic redundancy is expanded to encompass data that acquire the information content of the original distorted data. Semantic redundancy also is used in this study.
(data) beyond that needed to convey a message. The excess data does not change the meaning of the original message. Semantic redundancy is the existence of data beyond the signal of a message that presents the message in a different format.

Research Questions

The primary research questions addressed in this study are:

1. Do syntactically redundant accounting accruals possess pragmatic information?

2. Do semantically redundant accounting accruals possess pragmatic information?

3. Do syntactically redundant accounting accruals possess a level of pragmatic information different from the level of pragmatic information possessed by semantically redundant accounting accruals?

4. Does the pragmatic information contained in accounting accruals diminish as the financial statement time frame increases?

Importance of the Research Questions

Business scholars in disciplines other than accounting question the value of accounting accruals. For example, Copeland and Weston (1988, 362-363) remark:

Investors should care only about the cash-flow implications of various corporate decisions. However, corporations report accounting [accrual] definitions of earnings, not cash flow, and frequently the two are not related.

Copeland and Weston (1979, 226) also comment that the results of several accounting studies (Sunder 1973; Sunder 1975; Kaplan and Roll 1972) indicate that "... investors actually discount cash flows and disregard changes in eps [accounting accruals] which do not reflect real economic events."
While Copeland and Weston (1988) presume that accounting accruals are unrelated to future cash flows, this study provides a theory relating accounting accruals to future cash flows. This research suggests that the value of accounting accruals is related to the enhancement of the accuracy of prediction of prospective cash flows. Consequently, this study does not dispute the contention that data items related to prospective cash flows have value, but maintains that accounting accruals are related to future cash flows and can be used to enhance the prediction of cash flows.

Accounting data have different values in different situations. If accounting accruals have no value to users of accounting data in a specific situation, then those accruals should be examined to determine if they have a purpose in that situation. For example, one accounting data item may have value to investors but no value to managers of a firm. That data item should be included in financial reports for investors but possibly may be eliminated from financial reports for managers. The value of accounting accruals can be examined only in the context of a specific situation. The situation examined by this study is external financial reporting. Implications of the results of this study for financial reporting only apply to this context.

Research questions 1, 2, and 3 address the issue of whether specific types of accounting accruals have value. The significance of this matter is noted by Copeland and Weston (1988, 363):

Does an efficient market look at the effect of managerial decisions on [accrual] earnings per share (eps) or cash flow? This is not an unimportant question . . .

Financial reports are issued for periods of various durations. Quarterly and annual financial reporting periods are standard time frames
in external financial reporting. Internal financial reports can encompass periods ranging from less than one day to several decades. This study hypothesizes a negative relationship between the value of accounting accruals and the financial reporting period. The relationship is negative because the noise introduced into cash-flow reports by periodic reporting decreases in relation to the total signal as the time period covered by the reports increases. Conversely, the noise introduced into cash-flow reports by periodic reporting increases in relation to the total signal as the time period covered by the reports decreases. Any redundant data (accruals) introduced into financial reports contains less pragmatic information as the financial report time frame increases because of the reduction in the noise-to-signal ratio (accruals have less value); however, this redundant data contains increasing pragmatic information as the financial report time frame decreases because of the increase in the noise-to-signal ratio (accruals have greater value). For example, an accounting accrual may have value in a quarterly financial report but may be without value in an annual financial report. Because this study hypothesizes that the value of accounting accruals declines with increases in the financial reporting period, a determination must be made of the value of accruals in the context of time frames relevant to external financial reporting. Research question 4 considers this matter.

**Method**

Quarterly and annual accounting data from the balance sheet, income statement, and statement of changes in financial position (SCFP) were gathered from COMPUSTAT for a sample of firms over a three year period. Both quarterly and annual data from the first year were used to train a
neural network that forecasted future cash flows. Quarterly and annual data from the second and third years were used to predict future cash flows and determine the accuracy of the cash-flow predictions. The sample included all manufacturing firms (standard industrial codes 2000 to 3999) but was limited to firms for which all quarterly and annual data from the first quarter of 1986 to the last quarter of 1988 were available. The sample also was restricted to firms for which cash flows from operations\textsuperscript{11}, accrual elements of earnings, and accrual balance sheet items were available for the entire period under examination.

The data used in this study included the cash, cash flows from operations\textsuperscript{12}, accrual components of earnings from operations, and the accrual components of the balance sheet\textsuperscript{13}. Balances in the cash account were gathered from the balance sheet. Cash flows from operations was obtained from the SCFP for all companies in the sample. Syntactically redundant accruals and semantically redundant accruals from the income statement were gathered from the SCFP\textsuperscript{14}. The accrual components of the

\textsuperscript{11} This requirement restricted the sample to firms that used the cash basis of reporting on the SCFP.

\textsuperscript{12} A system of cash receipts and cash disbursements contains two items of interest to this study—cash and cash flows. Because this study focuses on operations, cash and cash flows from operations were used and collectively described as cash-flow data.

\textsuperscript{13} Only balance sheet items associated with operations were used, e.g., accounts receivable, inventory, and accounts payable. Most balance sheet items associated with either financing or investing activities, e.g., equipment and long-term notes payable, were excluded because they are not related to cash flows from operations.

\textsuperscript{14} APB No. 19 required that noncash items (accruals) be used in determining income reported on the SCFP. Consequently, accrual data was obtained from the body of the SCFP. If insufficient accrual data was reported on a firm's SCFP, additional accrual data was derived from the firm's balance sheet and income statement.
balance sheet also were gathered and classified as either syntactically or semantically redundant accruals¹⁵.

Syntactically redundant accounting accruals are used to signal messages about economic activity that are not reported under a system of cash receipts and cash disbursements. This economic activity is continuous and cannot be entirely represented by discrete cash transactions. Consequently, periodic reporting under a system of cash receipts and cash disbursements can be viewed as a source of syntactic noise because some messages are not signaled. The purpose of syntactically redundant accounting accruals is to reduce this syntactic noise. Bedford (1965, 30) comments on this situation:

... [cash] transactions are often inadequate for the task of revealing business activity. Extensive use of the accrual process suggests that transactions are, to accountants, merely means for recording activities. The implication is that transactions are not the only activities which accountants should record. In fact, the interpretation may be made that exchange transactions are merely expedient points useful in recording economic activity and that, as methods for measuring activity improve, it may become possible and appropriate to de-emphasize the importance of the transaction concept or event to drop it.

Examples of syntactically redundant accounting accruals include accounts receivable and accounts payable. Accounts receivable indicates an amount of money owed to the firm while accounts payable represents an amount of money owed by the firm. Neither of these items are reported under a system of cash receipts and cash disbursements.

Semantically redundant accounting accruals restate messages reported under a system of cash receipts and cash disbursements in order to

---

¹⁵ Although these accruals are the result of many periods of business activity, they represent potential cash inflows and potential cash outflows. Accordingly, balance sheet accruals associated with operations were used in the prediction of cash flows from operations.
provide a better representation of underlying economic activity. A system of cash receipts and cash disbursements treats economic activity as having no effect on the firm beyond the current period. Semantically redundant accounting accruals correct this flaw by reporting economic activity over a more appropriate time period. Examples of semantically redundant accounting accruals include inventory and depreciation. Both inventory and depreciation expense the cost of assets in periods when the assets generate revenues. The costs of these assets are reported as cash disbursements under a system of cash receipts and cash disbursements.

A 2 X 3 factorial design was used in this research. The first independent variable, data type, had three instances--cash and cash flows from operations (cash-flow data), cash-flow data plus syntactically redundant accounting accruals, and cash-flow data plus semantically redundant accounting data. The second independent variable, financial statement time frame, had two instances--quarterly and annual. See Table 1 for a description of the independent variables. Cash flows from operations one quarter into the future was the dependent variable for the quarterly financial statement time frame. Cash flows from operations one year into the future was the dependent variable for the annual financial statement time frame.

---

18 Pragmatic information of accounting accruals was operationalized in this research as the ability to predict future cash flows. Previous accounting studies have defined information content of accounting data as the generation of abnormal rates of return or of stock price changes. However, these definitions of information content assume that accounting data items have syntactic information content, semantic information content, and pragmatic information content. This study isolated pragmatic information by disregarding the syntactic and semantic information of accounting accruals. Accrual data may have been communicated to users before the release of financial statements through sources other than financial statements. This earlier communication would result in the lack
The 2 X 3 factorial research design used in this study was a function of the independent variables—2 financial statement time frames (quarterly and annual) X 3 data types [cash and cash flows from operations (cash-flow data), cash-flow data plus syntactically redundant accounting accruals, and cash-flow data plus semantically redundant accounting data]. Table 2 presents a diagram of the research design employed in this study. Cash-flow data is represented by $CF$, syntactically redundant accruals are represented by $SYN$, and semantically redundant accruals are represented by $SEM$. 

---

TABLE 1

INDEPENDENT VARIABLES

<table>
<thead>
<tr>
<th>Data Type</th>
<th>1. cash-flow data (CF)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2. cash-flow data plus syntactically redundant data (CF + SYN)</td>
</tr>
<tr>
<td></td>
<td>3. cash-flow data plus semantically redundant data (CF + SEM)</td>
</tr>
<tr>
<td>Financial Statement Time Frame</td>
<td>1. quarterly</td>
</tr>
<tr>
<td></td>
<td>2. annual</td>
</tr>
</tbody>
</table>
A neural network was used to structure the relationships between the independent variables and the dependent variables and to predict future cash flows. To utilize the neural network, the data set was divided into two subsets: (1) a training sample and (2) a holdout sample. The training sample was used to teach the network the relationships between the independent and dependent variables and consisted of the 1986 quarterly and annual data. The holdout sample was used to test the research questions and consisted of the 1987 through 1988 quarterly and annual data. The quarterly data and annual data were examined separately. The quarterly data was used to investigate the amount of pragmatic information in accounting accruals over a short time frame while the annual data was used to investigate the amount of pragmatic information in accounting accruals over a longer time frame. The neural network first was trained with the 1986 quarterly data and once the training was completed, the quarterly data was provided to the network for the first quarter of 1987. The neural network predicted the cash flows from operations for the second quarter of 1987 based on this data. Data from the second quarter of 1987 was then given to the network and was used to predict the cash flows from operations for the third quarter of
1987. This sequence was performed for all quarters in 1987 through 1988\textsuperscript{17}. The neural network was then retrained with the 1988 annual data. When this training was complete, the neural network was given the 1987 annual data and predicted cash flows from operations for the year 1988. This procedure was performed for all three data types.

After the neural network predicted future cash flows from operations based on the independent variables, a measure of the accuracy of the predictions (error metric) was constructed. This measure was defined as the absolute value of the difference between the actual cash flows and the forecasted cash flows divided by the forecasted cash flows\textsuperscript{18}:

\[ \frac{|(\text{actual cash flows} - \text{forecasted cash flows})|}{\text{forecasted cash flows}}. \]

\textsuperscript{17} Data from the fourth quarter of 1988 were not used to predict cash flows in the first quarter of 1989 because the actual cash flows from the first quarter of 1989 were not available on the COMPUSTAT version used in this study.

\textsuperscript{18} Similar measures of predictive accuracy have been used in several previous forecasting studies (Waymire 1986; Hassell and Jennings 1986; and Brown, Hagerman, Griffin, and Zmijewski 1987). Using the absolute value of the measure enables the calculation of the magnitude of the sum of the differences between the forecasts and the actual future cash flows. Simply summing the differences results in positive differences offsetting negative differences and thus reduces the actual magnitude of the sum of the differences. Dividing the numerator by forecasted cash flows results in the expression of the error metric as a percentage of forecasted cash flows. Without the divisor, it is probable that the mean of the annual measures will be different from the mean of the quarterly measures.
Because the predictions of future cash flows between the various cells of the research design were not independent, ordinary least squares regression and ANOVA could not be used to estimate and investigate the error metric means. Because of the lack of independence, a generalized least squares (GLS) technique was used in this research. Generalized least squares does not require the assumption of independence of observations. Differences between means were examined using a GLS modification of the Bonferroni procedure. The modification of the Bonferroni procedure was accomplished by replacing the mean squared error in the test statistic with the estimated variance/covariance matrix of estimated means. The use of the modified Bonferroni procedure facilitated the determination of the pragmatic information of the two types of accounting accruals.

Objectives, Scope, and Limitations

This study has several objectives. The first is to explain a theory of the functions of accounting accruals. The second is to explain a theory

---

19 For example, the CF data for one firm was related to the CF + SYN, and CF + SEM data for that same firm. Consequently, the predictions of future cash flows across cells were related. The same sample also was used in all cells of the research design. This use of the same sample in all cells resulted in a lack of independence across cells. Berenson, Levine, and Goldstein (1983, 65-66) state:

It must be assumed that the subjects or experimental units in each of the c independent sample groups (i.e., levels of the factor) are randomly and independently drawn so that an observed value in any one group has no effect or influence on any other observed value in that group or any of the other groups. This assumption of independence cannot be relaxed. If it is violated, the ANOVA procedure . . . is invalid.

of the pragmatic information of accounting accruals. The third is to examine the pragmatic information of accounting accruals. The final objective is to examine the pragmatic information of accounting accruals in the context of different financial statement time frames.

The scope of the theoretical section is generic. The theory of the functions of accounting accruals explains the general purposes of broad categories of accounting accruals but does not explain the functions of individual accounting accruals. The theory of the pragmatic information of accounting accruals delineates why syntactically redundant and semantically redundant accounting accruals have pragmatic information. This theory does not attempt to define the pragmatic information of individual accounting accruals. This portion of the research does not attempt to determine if the theorized categories of accounting accruals have pragmatic information in all situations, but attempts to measure the pragmatic information of the broad categories of accounting accruals in quarterly and annual external financial statements.

Limitations include the fact that pragmatic information of cash-flow data and accounting accruals is operationalized as the ability to predict future cash flows. This measure of pragmatic information assumes that users of financial statements value the ability to predict future cash flows. Another limitation is the fact that the data items measured for pragmatic information may be communicated to users from sources other than financial statements. In this situation, these data items would have no syntactic information to users when reported in financial statements. The fact that quarterly accounting data usually is unaudited and annual accounting data usually is audited may confound the results of this study. The annual
accounting data may be more precise than the quarterly accounting data. Another limitation is related to managerial income smoothing. Managers may use accounting accruals to smooth quarterly accounting income and thus alter the relationship between accounting accruals and future cash flows. The variability of the quarterly accounting data also would be diminished. Finally, the data items investigated may have no meaning to users. Consequently, the data items would have no semantic information in this situation.

Contributions of the Study

Platt (1985) states that the use of theory in research assists in the inference of causation in the relationship between explanatory and outcome variables. The lack of an underlying theory makes results difficult to interpret because the associations between variables may be either causal or spurious. If a casual relation does not exist, then the relationship between variables may change over time. If the relationships between variables shift with time, similar studies separated in time may yield dissimilar results. This research describes theories of accounting accruals that may aid the inference of causation in future studies examining the effects of accounting accruals on other variables such as future cash flows and stock prices.

Three major categories of information exist: syntactic information, semantic information, and pragmatic information. This study investigates linkages between these three categories of information which may result in

---

21 See Platt (1985) for a discussion of shifting relationships between financial distress and variables used to predict financial distress.
a more thorough understanding of the informational characteristics of accounting data.

Summary

This chapter presented an overview of the theories, research problems, objectives, scope, limitations, and method that are relevant to this study. The contributions of the study also were discussed. Chapter 2 reviews the previous research and contrasts the purposes of those studies with the purpose of this study. Chapter 2 also presents an elaboration of the theories motivating this study. Chapter 3 contains a statement of the research hypotheses and an explanation of the method that is used to test the hypotheses. A discussion of the data analysis procedures is contained in Chapter 4. Chapter 5 summarizes the study and presents implications of the study, limitations of the study, and suggestions for future research.
CHAPTER 2

LITERATURE REVIEW AND THEORY

Several areas of research are important to this project. Accounting research related to syntactic and semantic information theory is of interest because these theories provide the theoretical framework for this study. Research related to the information content of accounting accruals and cash flows is also of interest because this research examines the pragmatic information of accounting accruals.

This chapter also develops two theories of accounting accruals upon which this research is based. The topics discussed in the theory section of this chapter are semiotic concepts, the theory of the functions of accounting accruals, the theory of the pragmatic information of accounting accruals, and the relationship between financial statement time frame and the pragmatic information of accounting accruals.

Literature Review

Syntactic Information Theory

One of the earliest applications of syntactic information theory in business research was undertaken by Theil (1967). He developed the concept that uses the decomposition of financial statement items as a basis for applying syntactic information theory to accounting. The decomposition concept involves describing all financial statement items as a proportion of a financial statement total (e.g., total assets, total liabilities, two times the total assets, total sales, total expenses). Thus, proportions replaced
probabilities in Theil's (1969) application of syntactic information theory to accounting. Theil suggested the use of assets information, liabilities information, balance sheet information, sales information, cost and expenses information, and industry-wide time horizon information. These measurement concepts form the basis for much of the later accounting literature concerned with syntactic information theory.

Lev (1969a, 1969b, 1970) extended Theil's ideas. In his 1969a monograph, Lev discussed the application of basic syntactic information theory concepts and Theil's concepts to differentiate between bankrupt and non-bankrupt firms and to determine the accuracy of budgets. Lev was able to distinguish between bankrupt and non-bankrupt firms a majority of the time (from 57 to 89 percent). He also demonstrated the usefulness of syntactic information theory concepts in the determination of budget accuracy. Lev is unique among accounting researchers who examined syntactic information theory concepts in that he discussed business applications of continuous syntactic information theory (peak demand forecasts of an electrical utility) in addition to discrete syntactic information theory.

Lee and Bedford (1969) defined accounting classification and measurement in terms of syntactic information theory. They proposed that the amount of information provided by a research study could be derived through the use of syntactic information theory concepts. A priori

---

1 These ideas are explained in his 1969 paper, "On the Use of Information Theory Concepts in the Analysis of Financial Statements" (which was based in part on his 1967 book: Economics and Information Theory).

2 Lee and Bedford (1969) suggest two entropy measures—"a priori" entropy and "a posteriori" entropy.
entropy is a measure of initial uncertainty regarding the possible outcomes of the research and a posteriori entropy is the unresolved uncertainty regarding the outcomes. The difference between these two entropy measures would be the information obtained from the research.

Ronen and Falk (1973) investigated the relationship between entropy and the perceived value of expected information. Specifically, they examined whether the entropy measure suggested by Lev (1969a) for the calculation of information loss from aggregation reflected user-perceived information loss and user-perceived need for information. They conducted three experiments in which cost information was presented to subjects. The entropy in each situation was calculated and compared to the user-perceived value of information. The results indicated that when the expected value of information monotonically increased with information, a high level of correlation existed between entropy and user-perceived value of information. However, the authors concluded that individual characteristics can affect the relationship between entropy and the perceived value of information. Consequently, the application of entropy to measure information loss from aggregation may be limited.

Jain (1973) described accounting in terms of linguistics. He stated that research in linguistics and psycholinguistics supports the theory that alternate accounting methods affect decision-making. He described accounting as a language, asserted that concepts from linguistics could be applied to accounting, and concluded (Jain 1973, 104):

3 The user-perceived value of information was operationalized by having subjects indicate how much they would be willing to pay for information in each situation.
Accounting rules are considered analogous to financial grammar and based on this analogy, the effect of grammatical structure on perception of listeners is studied to support the hypothesis that accounting methods affect decision making.

Abdel-khalik (1974) conducted a field study to determine if syntactic information theory concepts as extended by Theil (1969) and Lev (1969a, 1969b, 1970) were useful in the context of decisions based on accounting information. The study required business loan officers to make (1) loan recommendations and (2) estimates of the probability of firm default. These recommendations and estimates were based on financial statements with three levels of data aggregation. The results indicated no strong association between the entropy measure and the means of the task variables. Abdel-khalik (1974, 280-281) concluded that:

(a) The measure [entropy] is not temporal; it does not measure what the information will be at a given point in time.

(b) The measure is only defined for closed systems. Although some physical systems may be closed, all systems, including accounting, are open systems.

(c) The entropy's definition of information does not predict the behavior of the system when the information entropy reaches a particular level.

Babich (1975) discussed the concepts of syntactic information theory and the extensions by Theil (1969) and Lev (1969a, 1969b, 1970). He compared syntactic information theory analysis of financial statements with ratio analysis and concluded that syntactic information theory analysis was inferior to ratio analysis. The basis for his conclusions was that ratio analysis reveals both the magnitude and direction of change of financial items while syntactic information theory analysis reveals only the magnitude of the change.
Gorelik (1975) investigated the nature of information through the use of a systems approach. He discussed the various definitions of information and the methods of measuring information. He concluded that information is related to uncertainty (syntactics) and value (pragmatics).

Belkaoui (1976) used syntactic information theory to discriminate between firms acquired by other companies and firms not acquired by other companies. The information measures used were the disaggregation measures proposed by Theil (1969) and Lev (1969a, 1969b, 1970). The results (Belkaoui 1976, 49) indicated that "...the information measures for the firms taken over are larger than for the firms in the control group..." Belkaoui concluded that the information measures he used were good predictors of corporate takeovers.

Pendlebury (1980) investigated the use of syntactic information theory in accounting for groups of companies. The concepts used in his research were basically those proposed by Theil (1969) and Lev (1969a, 1969b, 1970). He attempted to determine the best level of data aggregation for groups of companies and concluded that a data decomposition index could be used to accomplish the purpose.

Semantic Information Theory

Nakano (1972) was the only accounting researcher to utilize the concept of redundancy as the basis for his paper. Nakano (1972) discussed the use of semantically redundant financial data (price-adjusted financial statements and physical and environmental data) to reduce the social loss from the use of historical cost financial statements. He concluded that semantic redundancy could reduce social loss from the use of historical cost financial statements alone, but also stated that the
benefits and costs of using semantically redundant data need to be considered before introducing it into a situation.

Haried (1972, 376) investigated the "semantic problems of external accounting communication." A process called semantic differentiation was used to locate the meaning of various accounting terms, e.g., asset, liability, depreciation, in semantic space. Sophisticated subjects and unsophisticated subjects completed questionnaires composed of semantic scales. Factor analysis identified seven factors important to meaning in accounting contexts: objectivity, evaluation, control, activity, time, stability, and necessity.

Haried (1973) investigated whether certain accounting terms had similar meanings to different groups and whether synonymous accounting concepts had similar meanings. Sixteen accounting concepts including cash flow, depreciation, goodwill, and retained earnings were tested. Both sophisticated and unsophisticated subjects were used in the study. The results indicated that no differences were found in the overall meaning of concepts between groups. Haried did find that certain between-concept differences existed.

Lebar (1982) conducted a semantic analysis of the management analysis section of Form 10-K, the financial press release, and the

---

4 Semantic space is defined as a "Euclidian multidimensional region" (Haried 1972, 378). Orthogonal vectors in semantic space are represented by seven-point semantic scales that are defined by two contrasting adjectives, e.g., beneficial--adverse, tangible--intangible. The semantic scales are used to measure both distance and direction from the origin in semantic space.

5 Sophisticated subjects included CPAs and graduate accounting students. Unsophisticated subjects included investment club members and undergraduate engineering students.
president's letter in the annual report to determine if the language used in these items was consistent with the objectives of corporate disclosure. Extensional language was defined as (Lebar 1982, 177) "... language that tends toward full description, qualification, specificity, and objectivity." The converse of extensional language was defined as intensional language and was described as (Lebar 1982, 177) "... unqualified, generalized, and evaluative ..." Lebar (1982) commented that the use of extensional or intensional language can affect the types and quality of written information. Methods developed by semanticians were used to identify extensional and intensional language in the Form 10-K, press release, and annual report. Ten randomly selected New York Stock Exchange companies formed the sample of firms for the study. The results of the study indicated that (1) the three documents examined in the study addressed different topics, (2) in some cases, the three documents contradicted each other, (3) the Form 10-Ks had the greatest amount of extensionality and the least amount of intensionality, and (4) the annual reports had the greatest amount of intensionality and the least amount of extensionality.

Oliver (1974) used the semantic differential technique to determine if accounting concepts had different meanings to members of seven different professional groups. Unlike Haried (1972), Oliver focused on basic accounting concepts, e.g., accounting, income determination,

---

6 Lebar (1982, 177) defines the objectives of corporate disclosure as "... completeness, qualification, specificity, verification, and references."

7 These groups are the American Institute of Certified Public Accountants, the American Association of Collegiate Schools of Business, the Financial Analysis Federation, the National Association of Securities Dealers, Corporate Financial Executives, the Investment Bankers' Association, and the American Banking Association.
disclosure, matching, and cost. The results indicated that when all groups were analyzed simultaneously, significant differences existed for the meanings of six of eight concepts. However, when the groups were restructured to represent only accountants and users of accounting information, significant differences existed for five of eight concepts. When accounting academicians (American Association of Collegiate Schools of Business) were removed from the accounting group, significant differences existed for only two of eight concepts. The author concluded that CPAs share similar concepts of meanings with the five user groups. Accounting academicians, however, have different perceptions of meanings of the accounting concepts examined.

Belkaoui (1980) used multidimensional scaling techniques to analyze the perceptual differences of accounting concepts (e.g., the entity assumption, the going concern assumption, the period assumption) across three groups affiliated with accounting: accounting academicians, public accountants, and accounting students. Sociolinguistics was utilized in the construction of the hypothesis tested: "... perceptions of accounting concepts ... are a function of the professional accounting group affiliation" (Belkaoui 1980, 364). Data collected from a questionnaire was analyzed on three dimensions: conjunctive, relational, and disjunctive. The results indicated that there was a consensus on the meaning of the accounting concepts between the groups only in the relational dimension.

---

8 The conjunctive dimension is concerned with concepts that are perceived as sharing common perceptual characteristics. The relational dimension involves concepts that are linked by a fixed relationship. The disjunctive dimension consists of concepts that differ on one or more perceptual characteristics.
Houghton (1988) re-examined the results of Haried (1972, 1973). Houghton questioned Haried's (1972) method of factor analysis and expressed concern over the seven factors determined by Haried (1972) and the semantic differentials used to describe the factors. Houghton re-examined Haried's (1973) data and stated (Houghton 1988, 269) "... his [Haried's] conclusions about connotative meaning are not well founded." Houghton determined that Haried's (1972, 1973) seven-dimensional structure was not appropriate for use with accounting data and that the three-dimensional structure of Osgood, Suci, and Tannenbaum (1957) was appropriate for use with Haried's (1973) data. Houghton also noted that the re-examination of Haried's (1973) data indicated that significant between-group and between-concept differences existed.

Information Content of Accounting Accruals and Cash Flows

Harmon (1984) tested the association between stock market price reaction and (1) three measures of income and (2) six measures of cash flows. The purpose of his study was to determine whether income measures or funds flow measures were more highly associated with market reaction. The results indicated that earnings variables as a group ranked higher in their association with market reaction than funds flow variables. Consequently, funds flows may not be as useful as earnings to investors for decision making.

Schaefer and Kennelley (1986) investigated the explanatory power of cash-flow data over historical cost (accrual earnings) data. Previous studies, which used crude measures of cash flows, found that cash flows did not improve the explanation of security returns over earnings. Their
study used three measures of cash flows: one crude measure and two refined measures. The refined cash-flow measures did not show improvement in explanatory power over the crude cash-flow measure. In fact, none of the cash-flow measures added much explanatory power, perhaps because of significant collinearity between historical cost (accrual earnings) data and cash-flow data.

Rayburn (1986) explored the association of operating cash flows and accrual data with security returns over a twenty-year period. Both current accrual data and noncurrent accrual data were tested. Cross-sectional regression equations were constructed for each year examined and used to determine the level of association between cumulative abnormal returns (dependent variable) and operating cash flows, change in working capital, depreciation, and change in deferred taxes (independent variables). Current accruals were defined as change in working capital and noncurrent accruals were defined as depreciation and change in deferred taxes. The results indicated that operating cash flows and aggregate accruals were associated with cumulative abnormal returns. Current accruals also were associated with security returns. Noncurrent accruals were associated with security returns when holdout errors were used in the regression models and were not associated with security returns when errors generated by a random-walk process were used in the regression models.

Wilson (1986) examined the information content of both current and noncurrent accruals and the relative information content of total accruals and cash from operations. The relative information content of noncurrent accruals and working capital from operations and of current accruals and cash from operations was also examined. Current accruals were defined as
cash from operations minus working capital from operations. Noncurrent accruals were defined as working capital from operations minus earnings. Wilson found that cash and total accruals had information content beyond earnings and that total accruals had information content beyond the cash component of earnings. Another conclusion was that the noncurrent accrual component of earnings either did not have information content beyond working capital from operations or was known before the announcement of earnings.

Wilson (1987) attempted to ascertain if the release of the components of earnings (accruals and funds flows) have information content beyond the announcement of earnings. Accruals were separated into two categories: noncurrent accruals and total accruals. Funds flows were defined as working capital from operations and cash from operations. Cash from operations had incremental information content over earnings announcements while working capital from operations contained no incremental information content. Wilson concluded that cash and non-cash components of earnings, taken together, possessed information content beyond earnings. However, it could not be determined if either component had incremental information content alone.

Bowen, Burgstahler, and Daley (1987) examined the incremental information content of accrual earnings and cash flows. The questions they addressed were (1) Do cash flows have information content beyond accrual earnings? and (2) Do accrual earnings have information content beyond cash flows? The results indicated that (1) unexpected working capital from operations contained no information content beyond unexpected accrual earnings, (2) unexpected cash from operations and unexpected cash
flows after investment had information content beyond accrual earnings, and (3) unexpected accrual earnings possess information content beyond cash flows.

Ismail and Kim (1989) attempted to determine if funds flows and cash flows provide information beyond accrual earnings in explaining market risk. The results of their study revealed that (1) the addition of either funds flow or cash-flow-based risk measures improved the explanatory power of accrual earnings and (2) accrual earnings did not possess explanatory power beyond that of funds flow or cash-flow measures of risk.

Discussion of the Literature

Syntactic Information Theory

The studies involving syntactic information theory introduced accounting to concepts of semiotics. The entropy measures introduced by Theil (1969) and Lev (1969a, 1969b, 1970) provided alternative means of evaluating firm performance. Several studies investigating these measures, however, found that their usefulness may be limited9. The usefulness of the entropy measures may be restricted due to the fact that the measures are not calculated with probabilities but use financial statement proportions instead. Therefore, the entropy measures proposed by Theil (1969) and Lev (1969a, 1969b, 1970) may not be measuring entropy as defined by Shannon and Weaver (1949). Abdel-khalik (1974) recognized the limitations of Theil's (1969) and Lev's (1969a, 1969b, 1970) application of syntactic

---

9 See, for example, Ronen and Falk (1973); Abdel-khalik (1974); and Babich (1975).
information theory to accounting and commented that it should be designated as data decomposition theory.

**Semantic Information Theory**

Studies that applied semantic information theory to accounting focused the attention of the accounting profession on the meaning of accounting terms and concepts. The early studies in this area (Haried 1972, 1973; Oliver 1974; Belkaoui 1980) yielded conflicting results. Haried (1973) found that sophisticated and unsophisticated subjects attached similar meanings to accounting terms. Oliver (1974) found that seven different professional business groups assigned different meanings to basic accounting concepts. Belkaoui (1980) determined that seven professional groups agreed on the meanings of basic accounting concepts in only one of three dimensions. The fact that these studies measured the meanings of different items and used different methods of investigation and different groups of subjects may account for the disparity of results. A recent study (Houghton 1988), re-examined two of these early studies (Haried 1972, 1973) and concluded that the seven-dimensional structure used in these two early studies was not appropriate for use with accounting data.

**Information Content of Accounting**

**Accruals and Cash Flows**

Accounting research investigating the information content of accounting accruals and cash flows fixed the attention of the accounting profession on accounting data use. These studies explored new areas in accounting and did not have the advantage of theory support. Another complication is introduced by the definition of information content used in these studies. Most tested the ability of accounting accruals and cash
flows to explain changes in security prices or returns. For accounting data to influence security prices and/or security returns, the data must be used by investors in the selling or buying of securities. A data item used rationally in this manner must contain syntactic information (surprise value), semantic information (meaning), and pragmatic information (value). All data items that possess less than all three informational characteristics will exhibit no information content in these studies. An accounting data item may have been communicated to investors before the release of the financial statements. Accordingly, this data item will have no syntactic information but may contain semantic and pragmatic information. Another accounting data element may have no meaning (semantic information) to investors but may contain syntactic and pragmatic information. Both of these data elements would have exhibited no information content in past studies examining the information content of accounting accruals and/or cash flows.

Tailoring accounting reports to meet the needs of users requires knowledge of all the informational characteristics of accounting data. For example, if one data item lacks semantic information (meaning) but contains syntactic information (surprise value) and pragmatic information (value), ways of altering the data item or educating users on its meaning can be developed so that the data item can be used by users of accounting reports. Previous studies that examined the information content of accounting accruals and cash flows focused on market reaction to accounting data and did not examine the individual informational characteristics of accounting data. Research focusing on a single
informational element of accounting data may provide guidance for revision of accounting reports.

Contribution of this Research to the Literature

This research is most similar to the studies that examine the information content of accounting accruals and/or cash flows. This study, however, seeks solutions to some of the difficulties encountered in previous inquiries. First, theories of the functions and value of accounting accruals are central to this study. These theories may provide greater linkage between accounting accruals and the information contained in accounting accruals. Second, this study only attempts to examine the pragmatic information of accounting accruals. The investigation of only the pragmatic information of accrual data indicates the value of accruals to users of external financial statements. Syntactic information and semantic information of accounting accruals are not explored in this study. These informational characteristics of accounting accruals should be examined separately in other studies.

Theory

Theories of Accounting Accruals

Two theories are used to explain the pragmatic information of accounting accruals. Data, information, and knowledge are defined and differentiated before either of these theories is discussed. Semiotic concepts also are explained in detail before the presentation of the theories of accounting accruals.
Data, Information, and Knowledge

The words data, information, and knowledge, while sometimes used interchangeably, have distinct meanings in the realm of semiotics. Data is a necessary, but not sufficient, condition for information. Data is defined as (Caspari 1968, 8):

... facts which are obtained through empirical observation. They are predicates which may form the antecedent conditions of conditional statements. Both knowledge and information are built upon data.

Knowledge is described as "... a group of well-confirmed law-like generalizations which relate data to their environment" (Caspari 1968, 8). Information is a product of the application of knowledge to data. Information may be defined as "... the resultant of the coordination of data with knowledge when the data are screened, edited and evaluated for use . . ." (Caspari 1968, 8). Thus data, information, and knowledge are different, although related, items.

Semiotic Concepts

Properties of Information

Most broad definitions of information enumerate three qualitative characteristics of information: (1) surprise value, (2) meaning, and (3) value. \(^{10}\) Surprise value indicates that the content of a message or data item should be unknown to the receiver before the message is received. This characteristic is known as syntactic information. Syntactic information theory is described as (Hyvärinen 1968, 2):

---

\(^{10}\) These concepts of information can be traced to the theory of signs developed by Morris (1938, 1946). The study of all three qualitative characteristics of information is encompassed by the field of semiotics.
... a study of information carrying capacities of communication channels and the design of appropriate coding systems for an efficient data transfer with high reliability. No attention is paid to whether information has meaning or practical significance.

The characteristic meaning requires that a data item be understandable to the receiver. This quality is known as semantic information. Semantic information is concerned with the relationship between signs or symbols and the objects or ideas that they represent.

Finally, the characteristic value means that an item of data must be useful to the receiver\textsuperscript{11}. This characteristic of information is known as pragmatic information. Pragmatic information theory is concerned more with the characteristics of the receiver of the data than the actual data.

Use of Semiotics

Syntactic information theory is concerned with the coding of data, and consequently, is appropriate for use only in situations concerned with the structure (syntax) of data. Some researchers (Bar-Hillel and Carnap 1953) have indicated that the mapping of syntactic information theory concepts onto situations examining semantic information is possible. Cherry (1978) addressed this issue and concluded that syntactic information theory should not be used in situations examining the meaning or value of information\textsuperscript{12}.

\textsuperscript{11} In SFAC No. 1, the FASB states that information must be useful for a specific purpose, decision making; while in SFAC No. 2, the qualitative characteristics enumerated by the Board encompass these concepts of information.

\textsuperscript{12} Cherry (1978) noted that the only situation in which syntactic information theory may be mapped onto semantic theory involves meta-language. Meta-language is the language used by linguists and logicians to describe object-languages, e.g., English or French. Cherry (1978, 243) states "It is the semantic theory of this meta-language upon which the statistical (syntactic) theory may be mapped."
Semantic information is concerned with the meaning of data. Meaning however is difficult to define. Odgen and Richards (1949) discussed the triadic nature of meaning and described it as involving (1) a person having thoughts, (2) a sign, and (3) a referent (that which is represented by the sign). Semantic information theory is not concerned with coding or structure (syntactics) or value (pragmatics) and only should be used in examining the relationship between signs and their referents.

Pragmatic information is concerned with the relation of signs to users. Cherry (1978, 223) remarked that the study of pragmatic information includes "... all personal, psychological factors which distinguish one communication event from another, all questions of purpose, practical results, and value to sign users" (emphasis added). Signs can have different values to different individuals and consequently, can have different levels of pragmatic information across different individuals. It is therefore impossible to separate the value of information from individuals. Cherry (1978, 227) commented:

The pragmatic properties of any message depend upon the past experiences of the sender or the recipient, upon their present circumstances, their states of mind, and upon all matters personal to them as individuals.

13 Cherry (1978, 341) states that the word meaning is ambiguous and is used to represent several different concepts:

... (1) Translation ("What does the Latin amo mean?") (2) Near-synonyms, as in dictionaries ("Hit" means "strike.") (3) Significance ("What does £ mean?") (4) Value ("Life has little meaning to me now").

Quine (1960) identified two domains of meaning: (1) the theory of meaning, and (2) the theory of reference. The theory of meaning is concerned with the logical truth and the logical equivalence of statements. The theory of reference concerns "... extra-linguistic truthfulness and reference, whether a statement is true 'in fact' and experience" (Cherry 1978, 226).
Thus, pragmatic information theory is concerned with the effects of information upon individuals and should not be used to examine either the syntactics or semantics of information.

**Syntactic Information Theory**

Syntactic information theory is concerned only with syntactic information. As Hyvärinen (1968, 2) notes:

> On this level of information we are mainly interested in the number of possible symbols, words or some other suitable elements of information, their durations, statistical and deterministic constraints of successive symbols imposed by the rules of adopted language or coding system.

Expressed in simpler terms, syntactic information theory is concerned with the expectations and probabilities of informational elements. Therefore, syntactic information theory is related to probability theory, the science of uncertainty. Both syntactic information theory and probability theory are concerned with the probabilities of occurrences and measures of uncertainty (Drechsler 1969). The greater the probability of an event occurring, the smaller the syntactic information contained in the report of its occurrence. Conversely, the smaller the probability of an event occurring, the larger the syntactic information contained in the report of its occurrence (Bostwick 1968). These concepts are related to the surprise value of data items discussed earlier. The larger the surprise value contained in a data item, the larger the amount of syntactic information.

The message about the occurrence of an event is known as a definite message and the syntactic information contained in this message is defined as $\log \frac{1}{p}$, where $p$ is the probability of the event occurring (Lev 1969a). If a message contains a revised probability of the occurrence of an event, then this message is referred to as a nondefinite message and the
syntactic information contained in this message is defined as \( \log \frac{q}{p} \) where 

\( q \) is the revised probability and \( p \) is the initial probability (Lev 1969a)*.

Whenever there exists a series of possible events, the expected syntactic information of a message related to this series of events is called entropy. Entropy for a definite message (a message indicating which of the events occurred) is defined (Lev 1969a) as:

\[
(2.1) \quad \sum_{i=1}^{n} p_i \log \frac{1}{p_i}
\]

The expected syntactic information (entropy) of a nondefinite message (a message that indicates changes in the probabilities of occurrence of the events) is defined (Lev 1969a) as:

\[
(2.2) \quad \sum_{i=1}^{n} q_i \log \frac{q_i}{p_i}
\]

Entropy is zero when the probability of occurrence of any event (of all possible events) is 1 and increases to a maximum when the probabilities of

---

* One reason for using logarithmic measures is that the amount of syntactic information required to halve the number of possibilities in any situation is equal to \( \log 2 \). Any base of logarithms may be used, however \( \log_2 \) (base 2) is the most commonly used for the reason that \( \log_2 2 = 1 \); therefore, the amount of syntactic information required to halve the number of possibilities is equivalent to one unit or one bit (binary unit) (Staniland 1966); the result of using natural logarithms to measure syntactic information results in nits (Lev 1969a). Other reasons for using logarithms are due to their additive properties and the fact that they reflect the monotonically decreasing function of syntactic information—see Figure 1 in Appendix 5.
all possible events are equal—see Figure 2 in Appendix 5. For example, if a 100 percent probability of rain existed and someone told you that it was raining, that message would contain no syntactic information because you expected it to rain with total certainty and the message did not cause a revision of your expectations. Also, no syntactic information would be contained in a message if there is a zero probability of rain and the message indicated that it was not raining. The amount of entropy also increases as the number of possible events increases.

15 This concept is related to uncertainty. As the probabilities of all possible events converge, the amount of uncertainty associated with the occurrence of any event increases and consequently, the amount of entropy increases. Conversely, entropy decreases with a divergence of probabilities because of the reduction in uncertainty. If uncertainty is zero, the expected syntactic information content of a message (entropy) is zero because the content of the message is known with absolute certainty.

16 If someone told you that it did not rain when a high probability of rain existed, that message would contain a high amount of syntactic information because your beliefs would be drastically revised. The entropy in this situation, however, would be small. This small measure of entropy is a consequence of the fact that entropy is a measure of the expected syntactic information in a message and is concerned with the syntactic information and probabilities of all possible messages. Although the message about the lack of rain would contain a high amount of syntactic information, the probability of its occurrence would be small and, thus, its contribution to entropy would be small. Conversely, a message indicating rain would have low syntactic information and a high probability of occurrence and also would make a small contribution to entropy.

17 As the number of possible events increases, the uncertainty of occurrence of any one event increases. The increase in uncertainty results in an increase in entropy. For example, the amount of uncertainty, and thus entropy, associated with a specific play in a football game would increase if the number of rules of play were reduced. The entropy would increase because rules of play reduce uncertainty by restricting the number of acceptable plays. Consequently, if fewer rules were observed, the amount of entropy would increase because the number of possible events (plays) would increase.
Semantic Information Theory

Semantic information theory is concerned with the relationship between signs and referents and is used to investigate and analyze the meaning of data. Semantic information theory should be used only to examine the semantic information of simple declarative statements and is not concerned with the users of the statements. Cherry (1978, 238) explains:

The theory is in no way concerned with distinction between the "meaning of a sentence to a recipient" and the "intended meaning of a speaker,"... In fact, the theory is not concerned with communication at all--only with the semantic information "contained in" statements.

It is not practical to apply semantic information theory to real-world situations, e.g., object-languages such as English or French. However, the concepts of semantics can be used to describe object-languages.

Noise

Noise in syntactic information theory is the uncertainty of transmitted information given a non-zero probability of the occurrence of errors in transmission (Hyvärinen 1968). Expressed in simpler terms, syntactic noise is described as distortion of communication signals. Applied to accounting, this concept encompasses errors and deliberate manipulations

---

An example of the proper use of semantic information theory is given by Cherry (1978). The situation involves books and whether the books are (1) fiction (F) or nonfiction (not F), and (2) written in English (E) or written in a non-English language (not E). Example of statements to which semantic information theory can be applied include:

(1) $F_a \& E_a$ : book $a$ is fiction and written in English, and

(2) $F_a \& \text{not } E_a$ : book $a$ is fiction and not written in English.
of the financial statements (Nakano 1972). Semantic noise, in contrast, can be defined as distortion of the message transmitted. In other words, a message is transmitted, but is misunderstood by the receiver. According to Nakano (1972), semantic noise may be the result of the receiver's (1) lack of experience in communication, (2) insufficient knowledge of the source of the transmission, or (3) lack of understanding of the communication environment.

Noise requires the need for additional effort in the determination of the original message, and noisy signals never can reduce the amount of uncertainty surrounding a communication to zero (Nakano 1972). This research investigates both syntactic and semantic noise.

Conditional Entropy and Equivocation

Conditional entropy is the "... average amount of remaining uncertainty after receiving a noisy signal..." (Nakano 1972, 695) and is calculated as follows:

$$\sum_{i,j} P(X_i,Y_j) \log \frac{1}{P(X_i|Y_j)}$$

where $P(X_i,Y_j)$ is the probability that $X_i$ has been transmitted and $Y_j$ has been received, and $P(X_i|Y_j)$ is the conditional probability that $X_i$ has been transmitted when $Y_j$ is received (Nakano 1972). Thus the average amount

---

The concept of syntactic noise is related to qualitative characteristics of accounting information in SFAC No. 2. Representational faithfulness is described as "... correspondence or agreement between a measure or description and the phenomenon it purports to represent" (FASB 1980, 27). Verifiability of accounting information ensures that similar reporting of economic events by independent measurers using the same method of measurement would result. To be neutral, accounting information must faithfully report events without bias. These three characteristics, in turn, compose the more general qualitative characteristic of reliability.
of syntactic information that can be communicated on a noisy channel is the difference between the entropy and the conditional entropy of a given information source and is called equivocation (Nakano 1972).

Redundancy

Two types of redundancy are differentiated: syntactic redundancy and semantic redundancy. Syntactic redundancy is the existence of coding (data) beyond that needed to convey a message. The excess data does not change the meaning of the original message. See Appendix 1 for a discussion of the syntactic redundancy in accrual-based financial statements. Semantic redundancy is the existence of data that presents a message in a format different than the original message.

Syntactic redundancy is a state where the syntactic information or expected syntactic information of a signal is lower than its capacity for syntactic information due to data in the signal that conveys no syntactic information. It would seem that syntactic redundancy serves no useful purpose. However, the presence of syntactic redundancy in a noisy channel makes the interpretation of a signal possible (Hyvärinen 1968).

The concept of syntactic redundancy is expanded by Nakano (1972, 697):

In [syntactic] information theory, "redundancy" means existence of interdependence among the signals or messages to be transmitted (hence existence of predictability of a forthcoming signal or message on the basis of the already received ones). This interdependence is created by the use of excessive or redundant encoding rules above the minimum necessary for the communication.

---

20 This definition reflects the traditional philosophy of redundant data in syntactic information theory. In this study, the concept of syntactic redundancy is expanded to encompass data that acquire the syntactic information content of the original distorted message.
Syntactic redundancy is related to the Fundamental Theorem of (Syntactic) Information Theory (Hyvärinen 1968, 97):

It is possible to transmit [syntactic] information through a noisy channel with an arbitrarily small probability of error if, and only if, the rate of information is less than the channel capacity.

Thus, the introduction of an appropriately conceived syntactic redundancy strategy to the communication of financial information should improve the reliability of the communication\(^2\).

Another advantage of syntactic redundancy is that it offers the receivers of a communication the possibility of strategy (Staniland 1966). Non-redundant data is either perceived or not perceived. Data containing syntactic redundancy gives the receiver a choice of possibilities in extracting syntactic information from a signal. The choice of possibilities in the interpretation of a signal enables success in the most efficient manner (Staniland 1966). Syntactic redundancy is expressed in the following manner (Hyvärinen 1968):

\[
R = 1 - \frac{A}{C_{\text{max}}}
\]

where \(R\) represents syntactic redundancy, \(A\) represents entropy and \(C_{\text{max}}\) represents the channel capacity (the maximum amount of syntactic information that can be conveyed by a method of communication).

\(^2\) Since accounting data are subject to distortion by syntactic noise, it appears that there should be uncertainty as to whether the correct signals have been received by users of financial statements. Some of the uncertainty that results from syntactic noise is removed through the performance of an audit. However, even an audit cannot remove all syntactic noise. To reduce the uncertainty regarding the signals transmitted in financial statements to zero, all transactions and events would have to be examined. Since this is not cost effective, an audit is structured to reduce the remaining syntactic noise to a tolerable level.
Semantic redundancy is the result of multiple transmissions of the same message in one or more signals. The redundant messages have the same meaning as the original message but have a different format. The purpose of presenting the message in different ways is to reduce semantic noise and enable the receivers of the message to interpret the message correctly.

Pragmatic Information

The value and usefulness of information is inherently associated with the users of that information and it is ultimately the value and usefulness of information that is of concern to the users. This fact is recognized in SFAC No. 2 (FASB 1980, ix):

The purpose of this Statement is to examine the characteristics that make accounting information useful...

... The usefulness of information must be evaluated in relation to the purposes to be served...

The characteristics of information that make it a desirable commodity can be viewed as a hierarchy of qualities, with usefulness for decision making of most importance (emphasis added).

The ultimate test of accounting data is the test for value or usefulness and involves the examination of accounting data for pragmatic information.

---

22 Cherry (1978) notes that semantic redundancy requires that additional signs be used in communication.

23 Uncertainty that results from semantic noise must be reduced through the use of methods different than those used to reduce uncertainty from syntactic noise. The reduction of semantic noise usually involves the training of users to correctly interpret information or the addition of additional data that presents the original message in a different format. An example of presenting a message in a different format is the use of price-level adjusted data (semantically redundant data) to compensate for uncertainty that results from the effects of inflation (semantic noise) on historical-cost data (Nakano 1972).
Pragmatic information, however, cannot be separated from individual users. The same accounting data can have different levels of pragmatic information for different users thus necessitating the assumption of an average user to test accounting data for pragmatic information. This assumption is inherent to the section of SFAC No. 1 (FASB 1978, 17-18) that describes the type of accounting data that is useful to users:

... financial reporting should provide information to help investors, creditors, and others assess the amounts, timing, and uncertainty of prospective net cash inflows. ...

In this statement, the FASB is making an assumption about the type of information that is useful to the average user of accounting data. This study also assumes that data about future cash flows is useful to the average user of accounting data.

Accounting researchers also have recognized the importance of cash flows to users of financial data. Watts and Zimmerman (1986) note that one of the assumptions underlying many capital market studies is that earnings is a surrogate for cash flows. In fact, they also note that current earnings may be more highly associated with future cash flows than current cash flows indicating that the accrual components of earnings have pragmatic information. This study examines the pragmatic information of accounting accruals.

Theory of the Functions of Accounting Accruals

The theory of the functions of accounting accruals categorizes accounting accruals by their function. Consequently, this theory is a
descriptive theory. More specifically, the theory of the functions of accounting accruals is a classification theory because it describes mutually exclusive categories of accounting accruals.

The theory of the functions of accounting accruals is induced from observation of accounting accruals. One type of accounting accrual reflects incomplete transactions under a system of cash receipts and cash disbursements. The function of these accruals is to convey data about economic events associated with incomplete cash transactions. These accruals are referred to as syntactically redundant accruals. The second type of accounting accrual presents messages in a different format than their counterparts in a system of cash receipts and cash disbursements. The function of these accruals is to enable accounting data users to perceive messages that may be unclear when signaled under a system of cash receipts and cash disbursements. These accruals are referred to as semantically redundant accruals. The theory of the functions of accounting

24 According to Fawcett and Downs (1986, 4-5), descriptive theories "... describe or classify specific dimensions or characteristics of individuals, groups, situations or events by summarizing the commonalities found in discrete observations."

25 A classification theory is a type of descriptive theory that

... states that the dimensions or characteristics of a given phenomenon are structurally interrelated. The dimensions may be mutually exclusive, overlapping, hierarchical, or sequential. Classification theories frequently are referred to as typologies or taxonomies (Fawcett and Downs 1986, 5).

26 Some cash flow data is so highly aggregated in the statement of changes in financial position and the statement of cash flows that the argument may be made that some semantically redundant accounting accruals are also syntactically redundant (the information of the original data is lost through aggregation.) This research, however, assumes that the two categories of accounting accruals are mutually exclusive.
accruals states that accounting accruals are either syntactically redundant or semantically redundant in function.

Because a system of cash receipts and cash disbursements focuses on discrete transactions, economic events that have not resulted in completed cash transactions at the end of a reporting period are not disclosed. The messages associated with these incomplete transactions are not reported (signaled) under a system of cash receipts and cash disbursements in the period associated with the economic messages. Periodic reporting under a system of cash receipts and cash disbursements introduces syntactic noise into financial reports because some of the economic messages are not disclosed (signaled)^{27}.

Conceptually, syntactically redundant accounting accruals are considered redundant because they enable financial statement users to receive signals that are distorted by syntactic noise. In this case, noise is introduced into the financial reporting process by periodic reporting under a system of cash receipts and cash disbursements. Syntactically redundant accruals are used to recover some of the information lost to noise.

^{27} The messages that are not signaled are those that are associated with the incomplete cash transactions. Bedford (1965, 30) described this situation:

Essentially, the accrual process is an accounting for economic activities between transactions or for activities leading up to a transaction. That is, economic activity may be initiated by a purchase transaction and may be terminated by a sales transaction, but accountants must rely upon the accrual process to reveal economic activity between the initial transaction and the final transaction. Alternatively, economic activity may be initiated by the expectation of a future transaction, such as an employee's starting to work with the expectation that he will be paid later, and the accrual process must be used to reveal the economic activity prior to the final exchange transaction.
This concept is explained by an example. Imagine the periodic accounting reporting process as a communication process similar to a batch processing operation. Individual messages (transactions) are accumulated (and aggregated) until the passage of time requires the transmission of the signal (financial statements). In a system of cash receipts and cash disbursements, some transactions are not complete at the transmission date (reporting date) and thus the messages associated with those incomplete transactions are not signaled. An example would be the delivery of inventory items before the financial reporting date with payment made after the financial statement cut-off date. Because payment is made after the financial reporting date, the recording (signaling) of the economic event occurs after the report cut-off date. A second example would be the sale of inventory items before the reporting date with payment received after the reporting date. A system of cash receipts and cash disbursements would result in distortions of the signals associated with these incomplete transactions.

A purpose of this research is to examine the information contained in syntactically redundant accounting accruals. As stated earlier, it is the pragmatic information of accounting data that is of ultimate interest to accounting. This study examines the pragmatic information of syntactically redundant accruals.

The semantic redundance of accruals provides signals of economic messages in different formats than the original signals. A system of cash receipts and cash disbursements may sometimes lead to a misinterpretation of the meaning of accounting signals. Examples of messages that may be misinterpreted under a system of cash receipts and cash disbursements
Include (1) the purchase of long-lived machinery and (2) prepayment of insurance. In both cases, a system of cash receipts and cash disbursements results in both items being reported as cash disbursements in the purchase period. In both examples, future economic benefits are expected to result from the transactions and thus the economic events are not discrete but continuous. A system of cash receipts and cash disbursements, however, treats these economic events as discrete and does not reflect the fact that these economic events affect future cash flows and future performance of the firm. Thus, the users of financial statements may not interpret the messages signaled under a system of cash receipts and cash disbursements correctly. If the messages are misunderstood, semantic noise is present. Semantically redundant accounting accruals provide the same message as their counterparts under a system of cash receipts and cash disbursements; however, the message is provided in a format that is more easily comprehended. Semantically redundant accruals contain pragmatic information if they reduce semantic noise associated with periodic reports issued under a system of cash receipts and cash disbursements.

A purpose of this research is to examine the pragmatic information of semantically redundant accounting accruals. The fact that syntactic and semantic information are subsets of pragmatic information (Cherry, 1978) makes these examinations possible. The examination of the pragmatic information of syntactically redundant and semantically redundant accruals is consistent with the uses of informational concepts discussed earlier:

---

28 See Bedford (1965) for a discussion of the use of discrete transactions to reflect continuous economic events.
syntactic and semantic information theories are not being used to examine pragmatic information.

**Theory of the Pragmatic Information of Accounting Accruals**

The theory of the pragmatic information of accounting accruals explains why accounting accruals contain pragmatic information. This theory is an explanatory theory. Three axioms are used to deduce the theory of the pragmatic information of accounting accruals:

- **Axiom 1:** The prediction of future cash flows can be enhanced by adding syntactically redundant accounting accruals to cash-flow data.

- **Axiom 2:** The prediction of future cash flows can be enhanced by adding semantically redundant accounting accruals to cash-flow data.

- **Axiom 3:** Data that enhances the prediction of future cash flows when added to cash-flow data contains pragmatic information.

The theory of the pragmatic information of accounting accruals states that syntactically redundant and semantically redundant accounting accruals have pragmatic information because accounting accruals enhance the prediction of future cash flows when added to cash-flow data.

Axiom 1 and Axiom 2 are derived from the Financial Accounting Standards Board's (FASB's) position in the Statement of Financial Accounting Concepts (SFAC) No. 1 that accounting accruals assist financial

---

29 "Explanatory theories specify relations between dimensions or characteristics of individuals, groups, situations, or events" (Fawcett and Downs 1986, 6).

30 Cash-flow data is defined in this study as cash and cash flows from operations.

31 This axiom equates predictive ability with value.
statement users in forecasting future cash flows. The FASB (1978, 21) states that investors' and creditors'

... interest in an enterprise's future cash flows and its ability to generate favorable cash flows leads primarily to an interest in information about its earnings rather than information directly about its cash flows. Financial statements that show only cash receipts and payments during a short period, such as a year, cannot adequately indicate whether or not an enterprise's performance is successful. Information about enterprise earnings and its components measured by accrual accounting generally provides a better indication of enterprise performance than information about current cash receipts and payments.

Axiom 3 ensues from the FASB's position in SFAC No. 1 that data that enables financial statement users to forecast future cash flows is useful (provides pragmatic information) to users. The FASB (1978, 17-19) notes that:

... financial reporting should provide information to help investors, creditors, and others assess the amounts, timing, and uncertainty of prospective cash inflows . . . .

... Investors, creditors, and others need information to help them form rational expectations about . . . prospective cash receipts and assess the risk that the amounts or timing of the receipts may differ from expectations, including information that helps them assess prospective cash flows to the enterprise in which they have invested or to which they have loaned funds.

SFAC No. 1 suggests that any data that can be used to construct forecasts of future cash flows that are more precise than forecasts produced with only cash-flow data contains pragmatic information. The theory of the pragmatic information of accounting accruals states that both syntactically redundant and semantically redundant accounting accruals combined with cash-flow data provide for more precise estimation of future cash flows than only cash-flow data and, therefore, syntactically redundant and semantically redundant accounting accruals contain pragmatic information.
Accounting Accruals and the Financial Statement Time Frames

The use of accounting accruals also is associated with time. Syntactically redundant accruals relate financial statements issued earlier in time to financial statements issued later. Semantically redundant accruals result in the distribution of the signal of an economic event over time. The longer the period of time covered by financial statements, the more similar reports of cash receipts and cash disbursements and accrual-basis financial statements become. This result is due to the fact that the noise introduced into reports of cash receipts and cash disbursements by periodic reporting decreases in relation to the total signal as the time period covered by the financial reports increases. Conversely, the noise introduced into reports of cash receipts and cash disbursements by periodic reporting increases in relation to the total signal as the time period covered by the financial reports decreases. Any redundant data (accruals) introduced into the accrual financial statements has less information as the time period covered by the statements increases because of the reduction in the noise-to-signal ratio; however, this redundant data has increasing information as the time period covered decreases because of the increase in the noise-to-signal ratio. This study examines the effect of the time frames of financial statement data upon the pragmatic information of syntactically redundant and semantically redundant accruals.

Summary

This chapter discussed previous accounting inquiries into (1) syntactic information theory, (2) semantic information theory, and (3) the information content of accounting accruals and/or cash-flow data. A
discussion of previous studies also was presented. The contribution of this study to the literature was discussed. This chapter also contained a discussion of the theories of accounting accruals that are a major focus of this research.
CHAPTER 3

METHOD

This chapter discusses the research hypotheses derived from the theory of the functions of accounting accruals and the theory of the pragmatic information of accounting accruals. A detailed explanation of the method that is used to investigate the research hypotheses also is included in this chapter. Included in the discussion of method are the following topics: independent variables, dependent variables, research design, data collection, forecasting of future cash flows, and statistical techniques.

Testing the Theories of Accounting Accruals

The theory of the functions of accounting accruals and the theory of the pragmatic information of accounting accruals are the focus of the empirical portion of this study. Research will either support the theories or will indicate that revision or abandonment of the theories is necessary. Because the theory of the pragmatic information of accounting accruals relies on the theory of the functions of accounting accruals for the categorization of accounting accruals by function, direct testing of the theory of pragmatic information of accounting accruals will result in indirect testing of the theory of the functions of accounting accruals. This indirect testing approach is used in this study. Therefore, the

---

1 Direct tests of the theory of the functions of accounting accruals would require the determination of probabilities associated with a variable of interest, e.g., future cash flows or earnings per share, for accounting
research questions, research hypotheses, variables, research design, data manipulations, and statistical tests have designs based on the theory of the pragmatic information of accounting accruals.

**Research Questions**

The research questions addressed in this study are:

1. Do syntactically redundant accounting accruals possess pragmatic information?

2. Do semantically redundant accounting accruals possess pragmatic information?

3. Do syntactically redundant accounting accruals possess a level of pragmatic information different from the level of pragmatic information possessed by semantically redundant accounting accruals?

4. Does the pragmatic information contained in accounting accruals diminish as the financial statement time frame increases?

**Independent Variables**

The research questions suggest the independent variables used in this research. Research questions 1, 2, and 3 consider the pragmatic information of syntactically redundant accounting accruals and semantically redundant accounting accruals. These research questions require that data type be an independent variable in this research. The categories of data type based on research questions 1, 2, and 3 are listed in Table 3 and include cash and cash flows from operations (cash-flow data), cash-flow accruals and cash-flow data and the determination and quantification of the meanings associated with accounting accruals and cash-flow data. Because of the extreme difficulty associated with these tasks, the indirect test approach is used in this research.
Research question 4 considers the effect of financial statement time frames on the pragmatic information of accounting accruals. Accordingly, financial statement time frame is an independent variable. Because the most common financial statement time frames in external financial reporting are quarterly and annual time frames, the categories of financial statement time frame used in this research are quarterly and annual\(^2\). The categories of financial statement time frame are listed in Table 4.

**TABLE 3**

<table>
<thead>
<tr>
<th>CATEGORIES OF DATA TYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. cash-flow data</td>
</tr>
<tr>
<td>2. cash-flow data plus syntactically redundant data</td>
</tr>
<tr>
<td>3. cash-flow data plus semantically redundant data</td>
</tr>
</tbody>
</table>

**TABLE 4**

<table>
<thead>
<tr>
<th>CATEGORIES OF FINANCIAL STATEMENT TIME FRAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. quarterly financial statement time frame</td>
</tr>
<tr>
<td>2. annual financial statement time frame</td>
</tr>
</tbody>
</table>

\(^2\) One of the contributions of this study is to provide evidence on the value of accounting accruals in the external reporting environment. Use of a financial statement time frame longer than one year, e.g., five years, in conjunction with a quarterly financial statement time frame would probably increase the probability of finding differences in the pragmatic information content of accounting accruals across financial statement time frames compared to using an annual financial statement time frame with a quarterly financial statement time frame. The use of a financial statement time frame longer than one year, however, would not generate evidence on the value of accounting accruals in annual financial statements and would not contribute as much to the understanding of external financial reporting as the use of an annual financial statement time frame.
Dependent Variables

This study examines the effect of the independent variables upon the pragmatic information of accounting accruals. The pragmatic information of accounting accruals is calculated by subtracting the pragmatic information of cash-flow data from operations from the pragmatic information of accrual accounting data plus cash-flow data. If the result is positive, accounting accruals contain pragmatic information. Otherwise, accounting accruals do not contain pragmatic information. The pragmatic information of accounting accruals, however, cannot be examined directly. The pragmatic information of accounting accruals and the pragmatic information of cash-flow data must be inspected through the use of a proxy. The proxy that represents pragmatic information in this study is the enhancement of the precision of forecasts of cash flows from operations. This proxy is used because the FASB has indicated in SFAC No. 1 that data that enhances the accuracy of the prediction of future cash flows is useful to financial statement users. Consequently, data that facilitates the prediction of future cash flows has pragmatic information to financial statement users. The greater the precision of forecasts of future cash flows, the more pragmatic information the data contains. Comparing the precision of cash-flow forecasts based on cash-flow data and the precision of cash-flow forecasts based on accounting accruals plus cash-flow data enables the determination of the pragmatic information of accounting accruals. The ability of forecasters to use accounting accruals to forecast future cash flows with greater accuracy than with cash-flow data only is measured with an error metric that is defined as the absolute
value of the difference between the actual cash flow and the forecasted cash flow divided by the forecasted cash flow:

\[ \frac{|(actual\ cash\ flow - forecasted\ cash\ flow)|}{forecasted\ cash\ flow}. \]

Both quarterly and annual accounting data can be used to forecast quarterly and annual cash flows. Quarterly and annual accounting data in various combinations across time also can be used to forecast quarterly and annual cash flows. Testing the forecasting qualities of all possible combinations of quarterly and annual accounting data would be extremely tedious and best suited for research attempting to identify combinations of data having maximum pragmatic information. This research attempts only to determine if accounting accruals contain pragmatic information in specific situations. The forecasting characteristics of all possible combinations of accounting data are not examined. This research focuses on the qualities of quarterly accounting data that improve the forecasting of cash flows from operations one quarter into the future and the qualities of annual accounting data that improve the forecasting of cash flows from operations one year into the future. Thus the dependent variable for the quarterly financial statement time frame is cash flows from operations one quarter into the future and the dependent variable for the annual financial statement time frame is cash flows from operations one year into the future.

---

1 Similar measures of predictive accuracy have been used in previous forecasting studies (Waymire 1986; Hassell and Jennings 1988; and Brown, Hagerman, Griffin, and Zmijewski 1987). Using the absolute value of the measure facilitates the calculation of the magnitude of the sum of the differences between the predictions and the actual future cash flows. Simply summing the differences results in positive differences offsetting negative differences and thus reduces the actual magnitude of the sum of the differences. Dividing the numerator by forecasted cash flows permits the comparison of quarterly measures with annual measures. Without the divisor, it is probable that the mean of the annual measures will be different from the mean of the quarterly measures.
future. The dependent variable for each financial statement time frame is examined across all data sets. The dependent variables used in this study are listed in Table 5.

TABLE 5
DEPENDENT VARIABLES

<table>
<thead>
<tr>
<th>Financial Statement Time Frame</th>
<th>Dependent Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>quarterly</td>
<td>cash flows from operations one quarter into the future</td>
</tr>
<tr>
<td>annual</td>
<td>cash flows from operations one year into the future</td>
</tr>
</tbody>
</table>

Research Design

The research questions and independent variables of this research indicate that a 2 X 3 factorial design be used. Table 6 presents a diagram of the research design employed in this study. Cash-flow data is represented by \( CF \), syntactically redundant accruals are represented by \( SYN \), and semantically redundant accruals are represented by \( SEM \).

TABLE 6
RESEARCH DESIGN

<table>
<thead>
<tr>
<th>Quarterly Data</th>
<th>Data Type</th>
<th>Annual Data</th>
<th>Data Type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CF</td>
<td>CF + SYN</td>
<td>CF + SEM</td>
</tr>
<tr>
<td>CF</td>
<td>CF + SYN</td>
<td>CF + SEM</td>
<td></td>
</tr>
</tbody>
</table>
Research Hypotheses

Based on the research questions, independent variables, dependent variables, and the error metric, the anticipated results and the research hypotheses are developed. The anticipated results are necessary to formulate the research hypotheses and to determine the proper form of the hypothesis tests. The research hypotheses related to this study are separated into two categories: (1) the research hypotheses associated with the data sets and (2) the research hypotheses associated with the financial statement time frames.

Research Hypotheses Associated With Data Sets

The anticipated results related to the data sets are based on the FASB position in SFAC No. 1 that accounting accruals have value (pragmatic information) in the prediction of future cash flows. This FASB position, together with research questions 1 and 2, leads to the following null and alternative research hypotheses:

- **H_{01}**: The data set composed of cash-flow data plus syntactically redundant accounting accruals will have an error metric greater than or equal to the error metric of the data set composed of only cash-flow data.

- **H_{11}**: The data set composed of cash-flow data plus syntactically redundant accounting accruals will have an error metric smaller than the error metric of the data set composed of only cash-flow data.

- **H_{02}**: The data set composed of cash-flow data plus semantically redundant accounting accruals will have an error metric greater than or equal to the error metric of the data set composed of only cash-flow data.

- **H_{12}**: The data set composed of cash-flow data plus semantically redundant accounting accruals will have an error metric smaller than the error metric of the data set composed of only cash-flow data.
The theory of the pragmatic information of accounting accruals raises the question of how the pragmatic information of syntactically redundant accounting accruals differs from the pragmatic information of semantically redundant accounting accruals. Different functions of accounting accruals may possibly result in different levels of pragmatic information. The following null and alternative research hypotheses address the question of whether the pragmatic information of syntactically redundant accounting accruals differs from the pragmatic information of semantically redundant accounting accruals.

\( H_0^3 \): The data set composed of cash-flow data plus syntactically redundant accounting accruals will have an error metric no different from the error metric of the data set composed of cash-flow data plus semantically redundant accounting accruals.

\( H_a^3 \): The data set composed of cash-flow data plus syntactically redundant accounting accruals will have an error metric different from the error metric of the data set composed of cash-flow data plus semantically redundant accounting accruals.

Research Hypotheses Associated With Financial Statement Time Frames

The anticipated results related to the financial statement time frames are based upon the supposition that the noise from periodic reporting under a system of cash receipts and cash disbursements declines as a percentage of the total financial report signal as the financial reporting time frame increases. The following null and alternative research hypotheses are derived from this supposition and research question 4:

\( H_0^4 \): The difference between the accrual accounting data error metrics and the error metric of cash-flow data will either increase or remain unchanged as the financial statement time frame increases.

\( H_a^4 \): The difference between the accrual accounting data error metrics and the error metric of cash-flow data will decrease as the financial statement time frame increases.
Data Collection

Quarterly and annual accounting data from the balance sheet, income statement, and the statement of changes in financial position (SCFP) were gathered from COMPUSTAT for a sample of firms for the period 1986 through 1988. The quarterly data was used to investigate the amount of pragmatic information in accounting accruals over a short time frame while the annual data was used to investigate the amount of pragmatic information in accounting accruals over a longer time frame. The sample included all manufacturing firms (SIC codes 2000 to 3999) but was limited to firms for which all quarterly and annual data from the first quarter of 1986 to the last quarter of 1988 were available. The sample also was restricted to firms for which cash flows from operations, accrual components of earnings, and balance sheet data were available for the entire period under examination.

The data used in this study included the cash, cash flows from operations, accrual components of earnings from operations, and accrual components of the balance sheet. Balances in the cash account were gathered from the balance sheet. Cash flows from operations was obtained from the SCFP for all companies in the sample. Syntactically redundant accruals and semantically redundant accruals from the income statement

---

4 A system of cash receipts and cash disbursements contains two items of interest to this study—cash flows and cash. Because this study focuses on operations, cash and cash flows from operations are used in this study. Collectively, these two data items are designated cash-flow data.

5 Only balance sheet items associated with operations were used, e.g., accounts receivable, inventory, accounts payable. Most balance sheet items associated with either financing or investing activities, e.g., equipment, long-term notes payable, were excluded.
were identified from the SCFP. The accrual components of the balance sheet also were gathered and classified as either syntactically or semantically redundant accruals.

**Forecasting Future Cash Flows**

A neural network was used to structure the relationships between the independent variables and the dependent variables and to forecast cash flows. The neural network paradigm used in this study was backpropagation with a momentum term.

**Neural Network Advantages**

Neural networks have advantages over traditional statistical techniques. Neural networks have no assumptions about data that must be satisfied before proceeding with data analysis. Neural networks learn from data and adjust the relationships between variables if the variable relationships shift over time.

Models constructed with traditional statistical techniques become cumbersome if too many variables are included. Eliminating some variables

---

6 APB No. 19 required that noncash items (accruals) used in determining income be reported on the SCFP and therefore, accrual data were obtained from the body of the SCFP. If insufficient accrual data were reported on a firm’s SCFP, additional accrual data were derived from the firm’s balance sheet and income statement.

7 Although these accruals are the result of many periods of business activity, they represent potential future cash inflows and potential future cash outflows (FASB 1978, 20). Accordingly, they were used in the prediction of future cash flows from operations.

8 See Appendix 2 for a discussion of neural networks.

9 Backpropagation neural networks operate in a fashion similar to the stochastic approximation method and can be considered an approach to solving the least squares problem (White 1989). Appendix 3 discusses the backpropagation paradigm.
from the model however can result in model misspecification error. Using a neural network reduces the possibility of model misspecification error—all relevant data can be used by the neural network. Also, the difficulty of selecting the most appropriate model to use is eliminated because neural networks do not structure relationships among variables in predefined manners—they create models based on the underlying data.

Many statistical techniques do not perform well with noisy data. Neural networks can operate effectively with a high amount of noise in the data. In fact, Wasserman and Schwartz (1988, 11) state that "Neural networks can abstract the 'ideal' from a non-ideal training set."

Another advantage of neural networks is economy of effort. Because each independent variable in neural network processing can be considered a different dimension in Euclidean space, individual forecast models can be developed for each firm by the neural network by providing each firm a unique identification number. The identification numbers can be processed by the neural network with the rest of the data. An example is given to illustrate the effort-saving qualities of neural networks. Assume 300 firms compose the sample in a study. The purpose of the study is to develop forecast models of each firm for 3 data sets over 8 different forecast periods. To develop an individual forecast model of each firm for each forecast period and each data set in this study, data would be presented to a neural network 24 times (3 data sets X 8 forecast periods). Traditional regression analysis would require substantially more effort to develop forecast models for each firm in all forecast periods and all data sets. A total of 7200 (300 firms X 8 forecast periods X 3 data sets) regression models would be required to provide the desired results.
Training the Neural Network and Forecasting Cash Flows

Neural networks require a training set of data to learn the relationships between data items. The backpropagation paradigm is a type of supervised learning. Supervised learning paradigms require that neural networks be given a feedback value in order to learn the relationships between variables. Quarterly and annual data from 1986 is used to train the neural network in this study.

The neural network was trained first with quarterly data. After training with quarterly data was complete, data from the first quarter of 1987 was presented to the network. The neural network used this data to forecast cash flows from operations for the second quarter of 1987. Then the network forecasted cash flows from operations for the third quarter of 1987 with second quarter 1987 data. This cycle was repeated until forecasts of cash flows from operations were obtained for the second quarter of 1987 through the fourth quarter of 1988. Once all quarterly forecasts were made, the neural network was retrained with 1986 annual data. After training was complete, 1987 annual data was used to forecast cash flows from operations for 1988. This procedure was repeated for all data types.

Measuring Pragmatic Information

After the neural network predicted future cash flows from operations based on the independent variables, the error metric was constructed for all forecast periods and all data types. The error metric represents the precision of the forecasts of cash flows from operations and was used as a proxy for pragmatic information of accounting data in this study. An
inverse relationship exists between the error metric and pragmatic information: the smaller the error metric, the larger the pragmatic information of accounting data and the larger the error metric, the smaller the pragmatic information of accounting data.

Because the predictions of future cash flows between the various cells of the research design were not independent, ordinary least squares regression and ANOVA could not be used to estimate and investigate the error metric means. The generalized least squares technique (GLS) is appropriate to use when data is not independent and identically distributed (Berenson, Levine, and Goldstein 1983). For this reason, GLS techniques were used in this research.

Using the generalized least squares requires knowledge of the variance/covariance matrix of the data to calculate the weights used in the GLS procedure. The variance/covariance matrix of the sample used in this study, however, was unknown. Berenson, Levine, and Goldstein (1983) recommend a two-stage approach to GLS in this situation. The first stage is to use the data to estimate the variance/covariance matrix. The second stage entails using the estimated variance/covariance matrix obtained in the first stage to estimate the regression coefficients in the generalized least

---

10 For example, the CF data for one firm was related to the CF + SYN, and CF + SEM data for that same firm. Consequently, the predictions of future cash flows across cells were related. Berenson, Levine, and Goldstein (1983, 85-86) state:

It must be assumed that the subjects or experimental units in each of the c independent sample groups (i.e., levels of the factor) are randomly and independently drawn so that an observed value in any one group has no effect or influence on any other observed value in that group or any of the other groups. This assumption of independence cannot be relaxed. If it is violated, the ANOVA procedure . . . is invalid.
squares equations. A technique similar to this latter approach was used in this study.

Differences between individual means were examined using a GLS modification of the Bonferroni procedure\textsuperscript{11}. The modification of the Bonferroni procedure entailed replacing the mean squared error with the estimated variance/covariance matrix of the estimated means in the calculation of the test statistic. Use of the modified Bonferroni procedure facilitated the determination of the pragmatic information of the two types of accounting accruals.

\textsuperscript{11} See Kirk (1982) for a discussion of the Bonferroni procedure.
CHAPTER 4
DATA ANALYSIS

This chapter describes the data analysis procedures used in this study and the results of hypothesis testing. First, this chapter contains a discussion of the sample selection procedure used. Second, the forecasting of cash flows and the calculation of the error metrics are delineated. Third, the method of multiple comparisons used in this study is described. Finally, the results of hypothesis testing and a discussion of the results are presented.

Sample Selection

Firms used in this study were gathered from the 1988 annual and quarterly COMPUSTAT industrial tapes and included only those with standard industrial codes (SIC) between 2000 and 3999. The data items collected for the years 1986 through 1988 (including all quarters) included cash, cash flows from operations, and accounting accruals. These data items are listed in Table 7. Appendix 5 summarizes the decision model used to select data items.

In Table 8, the data items are categorized into one of three data types: cash and cash flows from operations (cash-flow data), syntactically redundant data, or semantically redundant data. This categorization was based upon the decision model described in Appendix 6. Accruals that signal messages not reported under a system of cash receipts and cash disbursements were classified as syntactically redundant while accruals
that restate messages that are reported under a system of cash receipts and cash disbursements were classified as semantically redundant.

To ensure that missing data items would not influence the results of the study, firms missing any data items in the period of interest initially were eliminated from the sample. However, this restriction eliminated virtually every firm. Therefore, the missing data item restriction was relaxed by reducing the number of data items examined for missing or zero values\(^1\) and the new restriction resulted in a sample size of 214 firms. Examination of this sample disclosed that data values of some firms were unrealistic\(^2\) and these firms were eliminated from the sample thus providing a final sample of 198 firms. See Table 9 for a summary of the sample selection process. Many firms had zero values in sale of property, plant, and equipment and sale of investments—loss (gain) for all time periods of interest. Therefore, this data item was dropped. Both annual and quarterly data items for the final sample were validated with COMPACT DISCLOSURE.

---

\(^1\) Debt in current liabilities, income taxes payable, equity in net loss (earnings), and sale of property, plant, and equipment and sale of investments—loss (gain) were not examined for missing or zero values.

\(^2\) All of the data values of these firms are near zero, e.g., $100.
<table>
<thead>
<tr>
<th>DATA ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. cash and short-term investments³</td>
</tr>
<tr>
<td>2. operating activities—net cash flow</td>
</tr>
<tr>
<td>3. receivables—total</td>
</tr>
<tr>
<td>4. inventories—total</td>
</tr>
<tr>
<td>5. debt in current liabilities</td>
</tr>
<tr>
<td>6. accounts payable and accrued liabilities</td>
</tr>
<tr>
<td>7. income taxes payable</td>
</tr>
<tr>
<td>8. deferred taxes⁴</td>
</tr>
<tr>
<td>9. equity in net loss (earnings)</td>
</tr>
<tr>
<td>10. sale of property, plant, and equipment and sale of investments—loss (gain)</td>
</tr>
<tr>
<td>11. depreciation and amortization</td>
</tr>
<tr>
<td>12. accounts receivable—decrease (increase)</td>
</tr>
<tr>
<td>13. inventories—decrease (increase)</td>
</tr>
<tr>
<td>14. accounts payable and accrued liabilities—increase (decrease)</td>
</tr>
<tr>
<td>15. income taxes payable—increase (decrease)</td>
</tr>
</tbody>
</table>

³ Certain disaggregated items, e.g., cash, accounts receivable—trade, and inventory—finished goods, that are available on the annual tape are not available on the quarterly tape. To ensure consistency between annual and quarterly data, aggregated data items were collected for both annual and quarterly time frames when disaggregated items were not available on the quarterly tape.

The level of aggregation used resulted in the mingling of cash and short-term investments, an accrual, and the combination of separate accruals into one data item, e.g., accounts payable and accrued liabilities, and depreciation and amortization.

⁴ This data item is classified as a liability on COMPUSTAT.
<table>
<thead>
<tr>
<th>Data Type</th>
<th>Data Item</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cash-flow data</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cash and short-term investments</td>
</tr>
<tr>
<td></td>
<td>Operating activities—net cash flow</td>
</tr>
<tr>
<td></td>
<td>Receivables—total</td>
</tr>
<tr>
<td></td>
<td>Accounts payable and accrued liabilities</td>
</tr>
<tr>
<td></td>
<td>Income taxes payable</td>
</tr>
<tr>
<td></td>
<td>Equity in net loss (earnings)</td>
</tr>
<tr>
<td><strong>Syntactically redundant data</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Accounts receivable—decrease (increase)</td>
</tr>
<tr>
<td></td>
<td>Accounts payable and accrued liabilities—increase (decrease)</td>
</tr>
<tr>
<td></td>
<td>Income taxes payable—increase (decrease)</td>
</tr>
<tr>
<td></td>
<td>Deferred taxes</td>
</tr>
<tr>
<td><strong>Semantically redundant data</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inventories—total</td>
</tr>
<tr>
<td></td>
<td>Debt in current liabilities</td>
</tr>
<tr>
<td></td>
<td>Inventories—decrease (increase)</td>
</tr>
<tr>
<td></td>
<td>Depreciation and amortization</td>
</tr>
<tr>
<td></td>
<td>Sale of property, plant and equipment and sale of investments—loss (gain)</td>
</tr>
</tbody>
</table>
TABLE 9
SAMPLE SELECTION

<table>
<thead>
<tr>
<th>Sample Selection Phase</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screening for missing or zero values for all data items</td>
<td>1</td>
</tr>
<tr>
<td>Screening for missing or zero values for subset of all data</td>
<td>214</td>
</tr>
<tr>
<td>items</td>
<td></td>
</tr>
<tr>
<td>Editing for unrealistic data values</td>
<td>198</td>
</tr>
</tbody>
</table>

Forecasting Cash Flows

The neural network software package used to forecast cash flows is NeuralWorks Professional II/Plus. Twenty-two neural network paradigms are available on this software and eight, including backpropagation, are suitable for forecasting. Although neural network literature describes backpropagation as one of the best forecasting paradigms in neural computing, a comparison was made of the forecasting capabilities of the networks suitable for prediction. The networks were trained and tested with 1986 annual cash-flow data. The sample was split into a training set of 150 firms and a test (holdback) set of 48 firms. The network performance test results are summarized in Table 10 and indicate that the

---

5 One of these networks, digital network architecture, was not evaluated because it requires the use of a proprietary computer chip. Another network paradigm, backpropagation with recirculation, was not considered because it is an unsupervised type of neural network and not suited for the type of forecasting performed in this study.

6 Consultation with a quantitative business analysis professor led to the decision to use a holdback sample size of 20% - 25% of the total sample.

7 Pearson's R coefficient measures how well a network is performing. A coefficient between .9 and 1 indicates excellent performance. The root mean square error is a measure of the network error. A low root mean
backpropagation paradigm performed as well or better than other networks. Consequently, the backpropagation paradigm was used for forecasting cash flows in this study.

Training the neural network begin with the 1986 quarterly data training sample consisting of 150 firms. Data was presented to the network for 200,000 iterations. If the Pearson's R coefficient and the root mean squared error indicated that the network was performing adequately\(^8\), training ceased and the network was tested with the 1986 quarterly data holdout sample. If network results were unacceptable, training was continued for another 200,000 iterations. Again, if the Pearson's R coefficient and the root mean squared error provided evidence that the network was performing adequately, training was stopped and the network was tested with 1987 annual data. If the network did not exhibit tolerable performance after 400,000 iterations through the test data, the network was discarded\(^9\) and a new backpropagation network architecture was constructed and trained\(^10\). This process continued until an acceptable

---

\(^8\) A Pearson's R coefficient of .8 or higher and a root mean square error of .1 or lower is evidence of acceptable network performance.

\(^9\) Experience with the data sets used in this study indicated slight performance improvement after 400,000 iterations.

\(^10\) The neural network software used in this study allows the construction of thousands of variations of backpropagation networks. Items that may be modified include the backpropagation learning law, the transfer function, the error function, the summation function, the output function, the noise function, the number of hidden layer processing elements, the number of hidden layers, the learning coefficient, the momentum coefficient, and the epoch.
network architecture was found for the data set. This procedure was performed for all data types.

Once an appropriate network architecture was established for a data type, the network was trained and forecasts of cash flows from operations were generated for the second quarter of 1987 through the fourth quarter of 1988. For example, the network was trained with 1986 quarterly cash-flow data and then cash-flow data from the first quarter of 1987 was used to forecast cash flows from operations for the second quarter of 1987. The network was then trained with quarterly cash-flow data from the first quarter of 1986 through the first quarter of 1987. Next a forecast of cash flows from operations for the third quarter of 1987 was generated with cash-flow data from the second quarter of 1987. This process was repeated through the third quarter of 1988. Tables 11, 12, and 13 summarize the performance measures and error metrics of neural networks used for forecasting quarterly cash flows.
### TABLE 10
NEURAL NETWORK PERFORMANCE TEST RESULTS

<table>
<thead>
<tr>
<th>Network Paradigm</th>
<th>Pearson's R Coefficient</th>
<th>Root Mean Squared Error</th>
<th>Mean Error Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backpropagation</td>
<td>.9942/.8899</td>
<td>.002</td>
<td>.715249</td>
</tr>
<tr>
<td>Backpropagation with functional link net sine inputs</td>
<td>.9967/.9130</td>
<td>.006</td>
<td>.788276</td>
</tr>
<tr>
<td>Delta-Bar-Delta</td>
<td>.9878/.8957</td>
<td>.004</td>
<td>.715345</td>
</tr>
<tr>
<td>Directed Random Search</td>
<td>.9634/.8809</td>
<td>.02</td>
<td>.763797</td>
</tr>
<tr>
<td>Extended Delta-Bar-Delta</td>
<td>.7269/.8751</td>
<td>.003</td>
<td>.714100</td>
</tr>
<tr>
<td>Kohonen self-organizing map into backpropagation</td>
<td>.8578/.8561</td>
<td>.0002</td>
<td>1.780147</td>
</tr>
</tbody>
</table>

### TABLE 11
NEURAL NETWORK PERFORMANCE MEASURES AND ERROR METRICS--QUARTERLY CASH-FLOW DATA

<table>
<thead>
<tr>
<th>Forecast Time Period</th>
<th>Pearson's R Coefficient--Training Data/Forecast Data</th>
<th>Root Mean Squared Error</th>
<th>Mean Error Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd qtr 1987</td>
<td>.8363/.9147</td>
<td>.004</td>
<td>.695239</td>
</tr>
<tr>
<td>3rd qtr 1987</td>
<td>.9858/.9777</td>
<td>.03</td>
<td>.600607</td>
</tr>
<tr>
<td>4th qtr 1987</td>
<td>.9854/.9605</td>
<td>.002</td>
<td>.860867</td>
</tr>
<tr>
<td>1st qtr 1988</td>
<td>.8452/.9627</td>
<td>.002</td>
<td>.809923</td>
</tr>
<tr>
<td>2nd qtr 1988</td>
<td>.8631/.9219</td>
<td>.003</td>
<td>.713670</td>
</tr>
<tr>
<td>3rd qtr 1988</td>
<td>.8114/.9565</td>
<td>.003</td>
<td>.697360</td>
</tr>
<tr>
<td>4th qtr 1988</td>
<td>.8618/.9484</td>
<td>.002</td>
<td>.701197</td>
</tr>
</tbody>
</table>
### TABLE 12

NEURAL NETWORK PERFORMANCE MEASURES AND ERROR METRICS—QUARTERLY CASH-FLOW DATA PLUS SYNTACTICALLY REDUNDANT DATA

<table>
<thead>
<tr>
<th>Forecast Time Period</th>
<th>Pearson's R Coefficient—Training Data/Forecast Data</th>
<th>Root Mean Squared Error</th>
<th>Mean Error Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd qtr 1987</td>
<td>.9097/.9566</td>
<td>.007</td>
<td>.656695</td>
</tr>
<tr>
<td>3rd qtr 1987</td>
<td>.9664/.9193</td>
<td>.001</td>
<td>.608605</td>
</tr>
<tr>
<td>4th qtr 1987</td>
<td>.7938/.9790</td>
<td>.0005</td>
<td>.668911</td>
</tr>
<tr>
<td>1st qtr 1988</td>
<td>.9728/.9661</td>
<td>.0009</td>
<td>.827604</td>
</tr>
<tr>
<td>2nd qtr 1988</td>
<td>.9447/.8832</td>
<td>.0007</td>
<td>.692804</td>
</tr>
<tr>
<td>3rd qtr 1988</td>
<td>.9757/.9616</td>
<td>.0005</td>
<td>.852700</td>
</tr>
<tr>
<td>4th qtr 1988</td>
<td>.9896/.9477</td>
<td>.0003</td>
<td>.708030</td>
</tr>
</tbody>
</table>

### TABLE 13

NEURAL NETWORK PERFORMANCE MEASURES AND ERROR METRICS—QUARTERLY CASH-FLOW DATA PLUS SEMANTICALLY REDUNDANT DATA

<table>
<thead>
<tr>
<th>Forecast Time Period</th>
<th>Pearson's R Coefficient—Training Data/Forecast Data</th>
<th>Root Mean Squared Error</th>
<th>Mean Error Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd qtr 1987</td>
<td>.9941/.8070</td>
<td>.01</td>
<td>.597449</td>
</tr>
<tr>
<td>3rd qtr 1987</td>
<td>.9325/.9621</td>
<td>.003</td>
<td>.594701</td>
</tr>
<tr>
<td>4th qtr 1987</td>
<td>.8773/.9613</td>
<td>.002</td>
<td>.621899</td>
</tr>
<tr>
<td>1st qtr 1988</td>
<td>.9857/.9484</td>
<td>.003</td>
<td>.839102</td>
</tr>
<tr>
<td>2nd qtr 1988</td>
<td>.9972/.9428</td>
<td>.008</td>
<td>.657273</td>
</tr>
<tr>
<td>3rd qtr 1988</td>
<td>.9148/.9316</td>
<td>.003</td>
<td>.658987</td>
</tr>
<tr>
<td>4th qtr 1988</td>
<td>.9351/.9567</td>
<td>.001</td>
<td>.710101</td>
</tr>
</tbody>
</table>
Once all the quarterly forecasts were obtained, networks were established for forecasting annual cash flows. Annual 1986 training data sets for all data types were used to train neural networks. The same method of constructing networks used with quarterly data was used with annual data. Annual 1986 data was used to train the networks and annual 1987 data was used to forecast annual 1988 cash flows. A summary of the performance measures and error metrics of neural networks used in forecasting annual cash flows is found in Table 14.

### TABLE 14

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Pearson’s R Coefficient--Training Data/Forecast Data</th>
<th>Root Mean Squared Error</th>
<th>Mean Error Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cash-Flow Data</td>
<td>.9549/.9635</td>
<td>.001</td>
<td>.620029</td>
</tr>
<tr>
<td>Cash-Flow plus Syntactically Redundant Data</td>
<td>.9753/.9475</td>
<td>.0001</td>
<td>.529143</td>
</tr>
<tr>
<td>Cash-Flow plus Semantically Redundant Data</td>
<td>.9660/.9534</td>
<td>.015</td>
<td>.529456</td>
</tr>
</tbody>
</table>
Hypothesis Tests

The Bonferroni procedure\(^\text{11}\) is an a priori method that can be used for testing multiple hypotheses concerning planned contrasts. In the usual ANOVA setting, if the hypothesized value of the contrast is \(L_0\), the test statistic for this method is:

\[
B = \frac{\sum_{j=1}^{p} c_j \bar{Y}_j - L_0}{\sqrt{MS_e \sum_{j=1}^{p} \frac{c_j^2}{n_j}}}
\]

where \(c_j\) is the contrast coefficient for the \(j\)th group, \(\bar{Y}_j\) is the estimated mean of the \(j\)th group, \(p\) is the number of groups, \(n_j\) is the number of observations for the \(j\)th group, and \(MS_e\) is the mean squared error (Kirk 1982). Critical values are obtained from the t-distribution and depend upon the number of contrasts being considered (Winer, Brown, and Michels 1991, 160). When the number of observations in each group are equal, this test statistic can be expressed in matrix form as:

\[
B = \frac{l' \hat{\mu} - l' \hat{\mu}}{\sqrt{MS_e \frac{1}{n} l' l}}
\]

where \(l\) is the column vector of contrast coefficients, \(\mu\) and \(\hat{\mu}\) are the column vectors of group means and estimated group means respectively, and \(n\) is the number of observations in each group.

When observations within a replication, i.e., in this study, a firm, are not independent, the test statistic must be modified to reflect this lack of independence.

\(^{11}\) This method also is known as Dunn's multiple comparison procedure.
independence (Dhrymes 1974, 112-113). Assuming independence between replications, i.e., firms, this modification is accomplished by replacing the mean squared error with the estimated covariance matrix for a vector of observations for a firm. The new statistic is:

\[
B' = \frac{l'\hat{\mu} - l'\mu}{\sqrt{\frac{1}{n} l'\hat{\Sigma} l}}
\]

where \( \hat{\Sigma} \) is the estimated covariance matrix. The \( B' \) test statistic was used in this study to test the null hypotheses. Critical values for \( B' \) test statistics were obtained from the standard normal distribution, since these statistics are asymptotically normal with mean 0 and variance 1\(^{12} \). Thus, the critical values for testing "c" one-tailed (two-tailed) hypotheses simultaneously with an overall experimentwise error rate of \( \alpha \) are:

\[
Z_{a/c} \quad (Z_{a/2C})
\]

respectively, where \( Z_{\alpha} \) is an upper Z-score from the standard normal distribution.

Testing of hypotheses 1, 2, and 4 required the use of a one-tailed test while testing of hypothesis 3 required the use of a two-tailed test. The critical values for both one-tailed tests and two-tailed tests at overall \( \alpha = .05 \) and overall \( \alpha = .01 \) are listed in table 15. These critical values are based on 12 contrasts.

\(^{12} \) If \( \hat{\Sigma} \) is replaced by \( D \) in \( B' \), asymptotic normality is a direct application of the central limit theorem and the theory of linear transformations of vector random variables (Seber 1977, 8-11; Mood, Graybill, and Boes 1974, 195). Asymptotic normality is guaranteed when estimating \( D \) with \( \hat{\Sigma} \) since \( \hat{\Sigma} \) is a consistent estimator of \( D \) (Fuller and Baltese 1973).
TABLE 15
CRITICAL VALUES FOR HYPOTHESIS TESTS

<table>
<thead>
<tr>
<th></th>
<th>One-Tailed Tests</th>
<th>Two-Tailed Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \alpha = .05 )</td>
<td>( \alpha = .01 )</td>
</tr>
<tr>
<td></td>
<td>2.64</td>
<td>3.17</td>
</tr>
</tbody>
</table>

To provide insight into the results of tests of the four primary research hypotheses, eight secondary research hypotheses also were tested. The null and alternative forms of the two secondary research hypotheses related to the first primary research hypothesis are:

\( H_0^{1a} \): The data set composed of both annual cash-flow data and annual syntactically redundant accounting accruals will have an error metric greater than or equal to the error metric of the data set composed of only annual cash-flow data.

\( H_a^{1a} \): The data set composed of both annual cash-flow data and annual syntactically redundant accounting accruals will have an error metric less than the error metric of the data set composed of only annual cash-flow data.

\( H_0^{1b} \): The data set composed of both quarterly cash-flow data and quarterly syntactically redundant accounting accruals will have an error metric greater than or equal to the error metric of the data set composed of only quarterly cash-flow data.

\( H_a^{1b} \): The data set composed of both quarterly cash-flow data and quarterly syntactically redundant accounting accruals will have an error metric less than the error metric of the data set composed of only quarterly cash-flow data.

The null and alternative forms of the two secondary research hypotheses associated with the second primary research hypothesis are:

\( H_0^{2a} \): The data set composed of both annual cash-flow data and annual semantically redundant accounting accruals will have an error metric greater than or equal to the error metric of the data set composed of only annual cash-flow data.
H₄₂ₐ: The data set composed of both annual cash-flow data and annual semantically redundant accounting accruals will have an error metric less than the error metric of the data set composed of only annual cash-flow data.

H₀₂ₐ: The data set composed of both quarterly cash-flow data and quarterly semantically redundant accounting accruals will have an error metric greater than or equal to the error metric of the data set composed of only quarterly cash-flow data.

H₄₂ₐ: The data set composed of both quarterly cash-flow data and quarterly semantically redundant accounting accruals will have an error metric less than the error metric of the data set composed of only quarterly cash-flow data.

Primary research hypothesis three has two affiliated secondary research hypotheses. The null and alternative forms of these two secondary research hypotheses are:

H₀₃ₐ: The data set composed of both annual cash-flow data and annual syntactically redundant accounting accruals will have an error metric no different from the error metric of the data set composed of both annual cash-flow data and annual semantically redundant accounting accruals.

H₄₃ₐ: The data set composed of both annual cash-flow data and annual syntactically redundant accounting accruals will have an error metric different from the error metric of the data set composed of both annual cash-flow data and annual semantically redundant accounting accruals.

H₀₃ₐ: The data set composed of both quarterly cash-flow data and quarterly syntactically redundant accounting accruals will have an error metric no different from the error metric of the data set composed of both quarterly cash-flow data and quarterly semantically redundant accounting accruals.

H₄₃ₐ: The data set composed of both quarterly cash-flow data and quarterly syntactically redundant accounting accruals will have an error metric different from the error metric of the data set composed of both quarterly cash-flow data and quarterly semantically redundant accounting accruals.

The null and alternative forms of the two secondary research hypotheses related to primary research hypothesis four are:
$H_0^{4a}$: The difference between the annual accrual accounting data error metrics and the annual cash-flow data error metric will either increase or remain unchanged as the financial statement time frame increases.

$H_{4a}$: The difference between the annual accrual accounting data error metrics and the annual cash-flow data error metric will decrease as the financial statement time frame increases.

$H_0^{4b}$: The difference between the quarterly accrual accounting data error metrics and the quarterly cash-flow data error metric will either increase or remain unchanged as the financial statement time frame increases.

$H_{4b}$: The difference between the quarterly accrual accounting data error metrics and the quarterly cash-flow data error metric will decrease as the financial statement time frame increases.

The SAS procedure CORR was used to calculate the estimated means of error metrics and the estimated variance/covariance matrix. Table 16 describes the estimated error metric means used in this study and Table 17 lists the values of these estimated means.

Linear combinations of the estimated means listed in Table 17 were used to test the research hypotheses and are listed in Table 18. The IML procedure in SAS was used to calculate $B'$ test statistics associated with each null hypothesis. Testing of the null hypotheses was accomplished by comparing the $B'$ test statistics with the critical values. Table 19 lists the comparisons made and the $B'$ test statistics.
### TABLE 16

**ESTIMATED ERROR METRIC MEANS**

<table>
<thead>
<tr>
<th>Data Type ((Y_j))</th>
<th>Forecast Period ((Y_{ij}))</th>
<th>Quarterly</th>
<th>Annual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>87-2</td>
<td>87-3</td>
<td>87-4</td>
</tr>
<tr>
<td>CF</td>
<td>(Y_{11})</td>
<td>(Y_{12})</td>
<td>(Y_{13})</td>
</tr>
<tr>
<td>CF + SYN(^{**})</td>
<td>(Y_{21})</td>
<td>(Y_{22})</td>
<td>(Y_{23})</td>
</tr>
<tr>
<td>CF + SEM(^{***})</td>
<td>(Y_{31})</td>
<td>(Y_{32})</td>
<td>(Y_{33})</td>
</tr>
</tbody>
</table>

\(^{*}\) Cash-flow data  
\(^{**}\) Cash-flow data plus syntactically redundant accounting accruals  
\(^{***}\) Cash-flow data plus semantically redundant accounting accruals
<table>
<thead>
<tr>
<th>Estimated Error Metric Mean</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y11</td>
<td>0.69523878</td>
</tr>
<tr>
<td>Y12</td>
<td>0.60060714</td>
</tr>
<tr>
<td>Y13</td>
<td>0.66086721</td>
</tr>
<tr>
<td>Y14</td>
<td>0.80992254</td>
</tr>
<tr>
<td>Y15</td>
<td>0.71367041</td>
</tr>
<tr>
<td>Y16</td>
<td>0.69736036</td>
</tr>
<tr>
<td>Y17</td>
<td>0.70119715</td>
</tr>
<tr>
<td>Y18</td>
<td>0.62002919</td>
</tr>
<tr>
<td>Y19</td>
<td>0.59744901</td>
</tr>
<tr>
<td>Y20</td>
<td>0.59470074</td>
</tr>
<tr>
<td>Y21</td>
<td>0.62189926</td>
</tr>
<tr>
<td>Y22</td>
<td>0.83910217</td>
</tr>
<tr>
<td>Y23</td>
<td>0.65727296</td>
</tr>
<tr>
<td>Y24</td>
<td>0.65988680</td>
</tr>
<tr>
<td>Y25</td>
<td>0.71010059</td>
</tr>
<tr>
<td>Y26</td>
<td>0.52945555</td>
</tr>
<tr>
<td>Y27</td>
<td>0.65689544</td>
</tr>
<tr>
<td>Y28</td>
<td>0.60860458</td>
</tr>
<tr>
<td>Y29</td>
<td>0.66891080</td>
</tr>
<tr>
<td>Y30</td>
<td>0.82760424</td>
</tr>
<tr>
<td>Y31</td>
<td>0.69280392</td>
</tr>
<tr>
<td>Y32</td>
<td>0.65270038</td>
</tr>
<tr>
<td>Y33</td>
<td>0.70803031</td>
</tr>
<tr>
<td>Y34</td>
<td>0.52914275</td>
</tr>
<tr>
<td>Mean</td>
<td>Symbol</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>Quarterly cash-flow data ([(Y_{11} + Y_{12} + \ldots + Y_{17})/7])</td>
<td>(\mu_1)</td>
</tr>
<tr>
<td>Annual cash-flow data ((Y_{18}))</td>
<td>(\mu_2)</td>
</tr>
<tr>
<td>Quarterly and annual cash-flow data ([(\mu_1 + \mu_2)/2])</td>
<td>(\mu_3)</td>
</tr>
<tr>
<td>Quarterly cash-flow plus syntactically redundant data ([(Y_{21} + Y_{22} + \ldots + Y_{27})/7])</td>
<td>(\mu_4)</td>
</tr>
<tr>
<td>Annual cash-flow plus syntactically redundant data ((Y_{28}))</td>
<td>(\mu_5)</td>
</tr>
<tr>
<td>Quarterly and annual cash-flow plus syntactically redundant data ([(\mu_4 + \mu_5)/2])</td>
<td>(\mu_6)</td>
</tr>
<tr>
<td>Quarterly cash-flow plus semantically redundant data ([(Y_{31} + Y_{32} + \ldots + Y_{37})/7])</td>
<td>(\mu_7)</td>
</tr>
<tr>
<td>Annual cash-flow plus semantically redundant data ((Y_{38}))</td>
<td>(\mu_8)</td>
</tr>
<tr>
<td>Quarterly and annual cash-flow plus semantically redundant data ([(\mu_7 + \mu_8)/2])</td>
<td>(\mu_9)</td>
</tr>
</tbody>
</table>
### TABLE 19

**CONTRASTS AND TEST STATISTICS**

<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>Contrast</th>
<th>Test Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₀₁</td>
<td>H₃ - H₄ ≤ 0</td>
<td>5.8280⁹</td>
</tr>
<tr>
<td>H₀₁₄</td>
<td>μ₂ - μ₃ ≤ 0</td>
<td>4.8141⁹</td>
</tr>
<tr>
<td>H₀₁₅</td>
<td>μ₁ - μ₄ ≤ 0</td>
<td>2.1880¹¹</td>
</tr>
<tr>
<td>H₀₂</td>
<td>μ₃ - μ₄ ≤ 0</td>
<td>4.3998⁹</td>
</tr>
<tr>
<td>H₀₂₄</td>
<td>μ₂ - μ₅ ≤ 0</td>
<td>5.0885⁹</td>
</tr>
<tr>
<td>H₀₂₅</td>
<td>μ₁ - μ₅ ≤ 0</td>
<td>0.7267¹¹</td>
</tr>
<tr>
<td>H₀₃</td>
<td>μ₃ - μ₆ ≤ 0</td>
<td>-0.9785¹¹</td>
</tr>
<tr>
<td>H₀₃₄</td>
<td>μ₄ - μ₅ = 0</td>
<td>0.0171¹¹</td>
</tr>
<tr>
<td>H₀₃₅</td>
<td>μ₇ - μ₄ = 0</td>
<td>-2.4941¹¹</td>
</tr>
<tr>
<td>H₀₄</td>
<td>[μ₁ - (μ₄ + μ₇)/2] - [μ₂ - (μ₅ + μ₆)/2] ≤ 0</td>
<td>-3.4735¹¹</td>
</tr>
<tr>
<td>H₀₄₄</td>
<td>(μ₁ - μ₄) - (μ₂ - μ₃) ≤ 0</td>
<td>-3.9286¹¹</td>
</tr>
<tr>
<td>H₀₄₅</td>
<td>(μ₁ - μ₄) - (μ₃ - μ₅) ≤ 0</td>
<td>-2.4749¹¹</td>
</tr>
</tbody>
</table>

⁹ Significant at α = .01.
¹¹ Not significant.

**Results of Hypothesis Tests**

Primary null hypothesis 1 (H₀₁) states that the error metric of the data set composed of cash-flow data and syntactically redundant accounting accruals will be larger than the error metric of the data set composed of only cash-flow data. The hypothesis test provides strong support against this null hypothesis. This result indicates that syntactically redundant accounting accruals possess pragmatic information. Rejection H₀₁ provides direct support for the theory of the pragmatic information of accounting
accruals and indirect support for the theory of the function of accounting accruals.

Examination of two secondary null hypotheses associated with $H_{01}$ yield further insight on the pragmatic information of syntactically redundant accounting accruals. The first secondary null hypothesis ($H_{01a}$) is concerned with the pragmatic information of syntactically redundant accounting accruals in annual financial statements while the second secondary null hypothesis ($H_{01b}$) focuses on the pragmatic information of syntactically redundant accounting accruals in quarterly financial statements. The results of the hypothesis tests provide strong evidence that syntactically redundant accounting accruals in annual financial statements have pragmatic information while no there is no evidence to support the existence of pragmatic information of these accruals in quarterly financial statements. These results imply that annual syntactically redundant accounting accruals have value in external financial reporting and that quarterly syntactically redundant accounting accruals do not have value in external financial reporting.

The second primary null hypothesis ($H_{02}$) states that the error metric of the data set composed of cash-flow data and semantically redundant accounting accruals will be larger than the error metric of the data set composed of cash-flow data only. The results of the hypothesis test provide strong evidence against this null hypothesis. The rejection of $H_{02}$ indicates that semantically redundant accounting accruals have pragmatic information. Rejection of $H_{02}$ also provides direct support for the theory of the pragmatic information of accounting accruals and indirect support for the theory of the functions of accounting accruals.
The results provide evidence related to the two secondary null hypotheses associated with \( H_0^2 \). No support is generated for the first secondary null hypothesis \( (H_{0a}^2) \) which focuses on the pragmatic information of semantically redundant accounting accruals in annual financial statements. The results provide no evidence against the second secondary null hypothesis \( (H_{0b}^2) \) which is concerned with the pragmatic information of semantically redundant accounting accruals in quarterly financial statements. Thus the results indicate that semantically redundant accounting accruals have pragmatic information in annual financial statements and contain no pragmatic information in quarterly financial statements. The implication of these results is that annual semantically redundant accounting accruals have value in external financial reporting and that quarterly semantically redundant accounting accruals do not have value in external financial reporting.

The results of testing primary research hypotheses 1 and 2 suggest that accounting accruals have pragmatic information in the context of this study. These results are contrary to the position espoused by some business theorists\(^{13}\). Accounting accruals, however, may not have value in all business situations. Research examining the pragmatic information of accounting accruals in various settings may enable the refinement of the theory of the pragmatic information of accounting accruals\(^{14}\).

\(^{13}\) For example, see Copeland and Weston (1988).

\(^{14}\) For example, research may investigate the value of accounting accruals in managerial, tax, and auditing settings. The results of these studies would enable researchers to refine the theory of the pragmatic information of accounting accruals.
Two propositions could explain the difference in pragmatic information possessed by annual accounting accruals and quarterly accounting accruals. The first proposition is that managers use accounting accruals to smooth income\textsuperscript{15}. If accruals are manipulated to induce changes in income, the relationship between accruals and future cash flows would be distorted. The fact that quarterly financial statements are not audited while annual financial statements are is the basis for the second proposition. Audited financial statements presumably are significantly more reliable than unaudited financial statements\textsuperscript{16}. Consequently, a substantial difference may exist between the strength of the relationship between annual accounting accruals and future cash flows and the strength of the relationship between quarterly accounting accruals and future cash flows. The fact that quarterly financial statements are unaudited may induce managers to engage in income smoothing behavior for quarterly reports. Further research investigating the lack of pragmatic information in quarterly accounting accruals is needed.

The third primary null hypothesis (H\textsubscript{03}) states that no difference will exist between the error metrics of the data set composed of cash-flow data plus syntactically redundant accounting accruals and the error metric of the data set composed of cash-flow data plus semantically redundant accounting accruals. The hypothesis test generates no evidence against H\textsubscript{03}. This result indicates that no difference exists between the amount of

\textsuperscript{15} See Watts and Zimmerman (1986, 144-146) for a discussion of the smoothing hypothesis.

\textsuperscript{16} This research does not examine the pragmatic information or reliability of audited financial statements or unaudited financial statements. Future research could focus on these areas.
pragmatic information of syntactically redundant accounting accruals and semantically redundant accounting accruals.

The two secondary null hypotheses associated with null hypothesis 3 \( (H_0^3_a \text{ and } H_0^3_b) \) focus on differences between the pragmatic information of syntactically redundant and semantically redundant accounting accruals in annual financial statements \( (H_0^3_a) \) and quarterly financial statements \( (H_0^3_b) \). The results of the hypothesis tests provide no evidence against these null hypotheses indicating that the pragmatic information in syntactically redundant accounting accruals corresponds in measure to the pragmatic information in semantically redundant accounting accruals in both annual and quarterly financial statements. Although syntactically redundant accounting accruals and semantically redundant accounting accruals do not display different levels of pragmatic information in this study, research examining different measures of pragmatic information may reach different conclusions.

Primary null hypothesis 4 \( (H_0^4) \) states that the difference between the accrual accounting data error metrics and the cash-flow data error metric will either increase or remain unchanged as the financial statement time frame increases. No evidence refuting this null hypothesis is generated. The results support the concept that accounting accruals contain more pragmatic information in annual financial statements than in quarterly financial statements\(^{17}\).

\(^{17}\) If the expected relationship between the pragmatic information content of accounting accruals and the financial statement time frame is positive, then the test results of \( H_0^4 \) and \( H_0^4_a \) are significant at overall \( \alpha=.01 \). These results indicate that annual accounting accruals contain a greater amount of pragmatic information than quarterly accounting accruals.
The results also fail to reject secondary null hypotheses $H_{04a}$ and $H_{04b}$. Failure to reject $H_{04a}$ implies that the pragmatic information of syntactically redundant accounting accruals does not decrease as the financial statement time frame increases. The lack of evidence against $H_{04b}$ suggests that a negative relationship between the pragmatic information of semantically redundant accounting accruals and the financial statement time frame does not exist.

A possible explanation of the results of testing $H_{04}$, $H_{04a}$, and $H_{04b}$ is related to the discussion of income smoothing and audited versus unaudited financial statements. If accounting accruals are used to smooth quarterly income and are less reliable than their audited (annual) counterparts, then quarterly accounting accruals may be a source of syntactic noise in the cash-flow forecasting process. An explanation for the difference in pragmatic information contained in annual and quarterly accruals is that audits reduce the noise in accounting data and thus increase the reliability of the data. The relationship between audits and the reliability of accounting data should be investigated further. The possibility exists that quarterly accounting accruals induce syntactic noise into the cash-flow forecasting process at a level higher than the noise level tolerated by neural networks. An inordinately high level of syntactic noise in quarterly accrual data would explain the results obtained for $H_{04}$, $H_{04a}$, and $H_{04b}$. Ascertaining the amount of syntactic noise in quarterly accounting accruals involves discovering true values of quarterly accruals. The difficulty in determining true values of quarterly accounting accruals for the sample firms would be extreme. Further research in this area may clarify the situation. Table 20 summarizes the results of all hypothesis tests.
<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0^1$: Syntactically Redundant Accruals—All Data</td>
<td>Reject</td>
</tr>
<tr>
<td>$H_0^1_a$: Syntactically Redundant Accruals—Annual Data</td>
<td>Reject</td>
</tr>
<tr>
<td>$H_0^1_b$: Syntactically Redundant Accruals—Quarterly Data</td>
<td>Fail to Reject</td>
</tr>
<tr>
<td>$H_0^2$: Semantically Redundant Accruals—All Data</td>
<td>Reject</td>
</tr>
<tr>
<td>$H_0^2_a$: Semantically Redundant Accruals—Annual Data</td>
<td>Reject</td>
</tr>
<tr>
<td>$H_0^2_b$: Semantically Redundant Accruals—Quarterly Data</td>
<td>Fail to Reject</td>
</tr>
<tr>
<td>$H_0^3$: Syntactically Redundant Accruals Versus Semantically Redundant Accruals—All Data</td>
<td>Fail to Reject</td>
</tr>
<tr>
<td>$H_0^3_a$: Syntactically Redundant Accruals Versus Semantically Redundant Accruals—Annual Data</td>
<td>Fail to Reject</td>
</tr>
<tr>
<td>$H_0^3_b$: Syntactically Redundant Accruals Versus Semantically Redundant Accruals—Quarterly Data</td>
<td>Fail to Reject</td>
</tr>
<tr>
<td>$H_0^4$: Annual Time Frame Versus Quarterly Time Frame—All Data</td>
<td>Fail to Reject</td>
</tr>
<tr>
<td>$H_0^4_a$: Annual Time Frame Versus Quarterly Time Frame—Syntactically Redundant Accruals</td>
<td>Fail to Reject</td>
</tr>
<tr>
<td>$H_0^4_b$: Annual Time Frame Versus Quarterly Time Frame—Semantically Redundant Accruals</td>
<td>Fail to Reject</td>
</tr>
</tbody>
</table>
CHAPTER 5
SUMMARY AND CONCLUSIONS

This chapter summarizes the study and discusses the contributions and limitations of the study. Avenues of future research are also considered. The first section contains a review of the study. The second section focuses on implications of the study. The third section delineates limitations of the study. The final section suggests areas of future research.

Summary of the Study

The theory of the functions of accounting accruals and the theory of the pragmatic information of accounting accruals form the core of this study. The theory of the functions of accounting accruals states that accounting accruals fall into one of two general categories: syntactically redundant accounting accruals or semantically redundant accounting accruals. The theory of the pragmatic information of accounting accruals states that accounting accruals have pragmatic information because of their functions. A hypothesis associated with these theories asserts that the pragmatic information of accounting accruals decreases as the financial statement time frame increases. Four primary and eight secondary research hypotheses are derived from these theories and the related hypothesis.

94
Cash-flow and accrual data were obtained from COMPUSTAT for a sample of firms and were provided to a neural network to generate forecasts of cash flows from operations. The forecasted cash flows were then used to calculate an error metric that functioned as a measure of pragmatic information. A variant of the Bonferroni procedure was used to detect differences between means of error metrics and to test the research hypotheses. The results of the hypothesis tests indicated that syntactically and semantically redundant accounting accruals contain pragmatic information. Tests of secondary research hypotheses, however, provided evidence that only annual accounting accruals contain pragmatic information. Hypothesis tests also indicated that no difference exists between the pragmatic information contained in syntactically redundant accounting accruals and the pragmatic information contained in semantically redundant accounting accruals. Finally, the results of hypothesis tests presented no evidence of a negative relationship between the pragmatic information of accounting accruals and the financial statement time frame.

Implications

Accounting accruals have been used for at least 150 years. Throughout the history of accounting accruals, no theories of the functions of accounting accruals or the value of accounting accruals have been proposed. Bedford (1965, 29) espoused the need for theories of accruals:

\[ \text{Footnote 1} \text{ For example, see Watts and Zimmerman (1986) for a discussion of the introduction of depreciation to the railroad industry in the 1840s and 1850s.} \]
there appears to be no satisfactory explanation of the nature of the accrual concept, nor are sufficient reasons set forth explaining why it should be used, and there exists no overall authoritative statement indicating where the concept should be used, or how the accrual is to be computed. . . .

Further, there should be some generalized statement of its [the accrual's] nature to simplify the concept and thus contribute to an understanding of accounting measurements. Until the concept is stated in generalized terms, confusion will exist in the minds of many.

The primary contribution of this study is the provision of two general theories of accounting accruals: the theory of the functions of accounting accruals and the theory of the pragmatic information of accounting accruals. The first theory delineates accounting accruals into two broad categories while the second theory defines the value of accounting accruals in terms of their functions. This study provides direct evidence supporting the theory of the pragmatic information of accounting accruals and indirect evidence supporting the theory of the functions of accounting accruals. These theories can be used as a framework in future research investigating accounting accruals. The theories discussed in this study can assist other researchers in supporting the results of their studies.

This study provides evidence of pragmatic information of annual accounting accruals while no evidence is generated supporting pragmatic information of quarterly accounting accruals. This result implies that annual accrual accounting data is a better predictor of future cash flows from operations than cash-flow data while quarterly accrual accounting data is not. An explanation is that quarterly accrual accounting data is significantly noisier than annual accrual accounting data. Quarterly financial statements never are audited while annual financial statements are. Audits may enhance the reliability of annual accounting data.
compared to the reliability of quarterly accounting data. Research investigating the value of audits is needed to clarify this situation. If the reliability of accounting data is enhanced by audits, the need for audits of quarterly financial statements is suggested (subject to the cost-benefit principle).

Another explanation for the difference between the pragmatic information of annual and quarterly accounting accruals is that managers use accruals to smooth quarterly income. Managers probably do not smooth annual income to the extent that they smooth quarterly income because of audits. Evidence indicating restriction of managerial income smoothing would further support the value of audits. Additional research is needed to investigate this area.

Many accounting studies have focused on the information content of accounting accruals and the incremental information content of accounting accruals over cash flows. Results of these studies have been inconsistent. A possible explanation of this inconsistency is the absence of a solid theory, resulting in restricted generalizability across variables. The absence of a theory to direct research also reduces the ability of researchers to compare and contrast the results of the various studies. The theory of the function of accounting accruals and the theory of the pragmatic information of accounting accruals would enhance the consistency of variables examined in accounting information studies and promote greater comparability across studies.

Information overload is a significant problem in the data-profuse environment of business. This study demonstrates a method that can be used to determine the value of accounting accruals in a specific decision
situation. Modifying the measure of pragmatic information will facilitate the use of this method in various decision environments.

Past accounting research using semiotic concepts focused on singular semiotic elements, e.g., syntactic information, semantic information, or pragmatic information. This study acknowledges the linkage between semiotic elements and defines the functions of accounting accruals in terms of syntactic and semantic information in order to measure the pragmatic information contained in accounting accruals. Recognizing that information does not consist exclusively of surprise value, meaning, or value will enable future accounting research to define more precisely the informational aspects of accounting data.

This study used neural networks as a methodological tool. Research in other disciplines has shown that neural networks are a useful data analysis technique. This study is one of the first to use a neural network in accounting research. A backpropagation neural network was used in this research to forecast cash flows from operations. This type of neural network can be used to forecast other accounting data. Other neural network paradigms can be used in numerous other pattern recognition situations. For example, a counterpropagation neural network can be used to classify firms as either bankrupt or nonbankrupt, a Kohonen self-organizing map into backpropagation neural network can be used for stock market forecasts, and a probabilistic neural network can be used to detect fraud. Neural networks comprise a new and powerful data analysis technique for accounting research and may enhance the ability to investigate complex research questions.
Limitations

The first limitation of this study is the lack of generalizability of the results of the study. The limited generalizability can be attributed to several factors. Because the firms included in the study were exclusively manufacturing firms, the results may not be generalizable to nonmanufacturing firms. Data restrictions placed on the sample also result in lack of generalizability. Only firms without certain missing data items were included in the sample. Firms with more complete reporting practices may differ from firms with less complete reporting methods. Consequently, the results may not be generalizable to firms excluded from the sample because of missing data items.

This is one of the few studies that used semiotic concepts in an accounting setting. The theory needs to be tested in several other accounting settings, using different methodologies, before we can validate its constructs. This study examined the use of accounting accruals in the forecasting of cash flows from operations. External financial statements were the only financial reports considered in this study. Accounting accruals are used in numerous situations, e.g., predicting bankruptcy, evaluating managerial performance, and detecting fraud, and appear in many forms of reports. The results of this study may not extend to other circumstances or financial reports.

Another constraint to the generalizability of the results is the time period investigated. Only data from the time period beginning with the first quarter of 1986 and ending with the fourth quarter of 1988 were used in the study. The results of this study may not be applicable to other time periods.
The measure of pragmatic information used in this study may be another limitation. The study relies upon the Financial Accounting Standards Board's position in Statement of Financial Accounting Concept No. 1 which asserts that data that assists financial statement users in forecasting cash flows has value (pragmatic information). If the ability to forecast cash flows does not have value to external financial statement users, then the conclusions of this study are tentative.

Missing data restrictions necessitated the use of training data that encompassed short time periods. One year of data was used to train a neural network\(^2\) in the forecasting of annual cash flows and a range extending from four to eleven quarters of data was used to train a neural network in the forecasting of quarterly cash flows. The use of training data from longer periods of time may result in different conclusions.

Quarterly COMPUSTAT data is highly aggregated. Annual data was aggregated to the same level as quarterly data to prevent different levels of data aggregation from influencing the outcome. However, the level of aggregation used may have affected the results. Less aggregated data may contain levels of pragmatic information different from the level of pragmatic information of the data used. Diverse levels of data aggregation, however, were not used in this study.

**Suggestions for Future Research**

This study examined the ability of accounting accruals to enhance the accuracy of forecasts of cash flows from operations. Investigation of

\(^2\) Neural networks are a developing methodological tool. The appropriateness of using neural networks to predict cash flows must be viewed with caution.
accounting accruals in other situations, e.g., predicting bankruptcy, evaluating managerial performance, and detecting fraud, would increase knowledge of the value of accounting accruals. Examination of the pragmatic information of accounting accruals in various situations would require the use of different measures of pragmatic information.

The time period of interest in this study ranges from 1986 to 1988. Extending the time period would provide additional evidence on the relationship between the pragmatic information of accounting accruals and the financial statement time frame. Future research also should explore fluctuations in the pragmatic information of accounting accruals over a period of time.

Behavioral experiments could be used to indicate the value of accounting accruals in real world situations. For example, financial analysts could be required to forecast cash flows with both cash-flow data and accrual accounting data. Designing behavioral experiments to minimize the influence of the syntactic and semantic qualities of accounting accruals on the pragmatic information of accounting accruals may be difficult.

This study examined only the pragmatic information of accounting accruals. Future research should isolate and investigate the syntactic and semantic information of accounting accruals. Results from these studies could be used in conjunction with results from studies on the pragmatic information of accounting accruals to construct comprehensive theories of information contained in accounting accruals.

The pragmatic information of individual accruals should be studied in the future. These studies would allow the determination of which accruals are most useful in specific situations.
Semiotics should be used to examine other aspects of accounting. Research could be conducted on the syntactic, semantic, and pragmatic information contained in aggregated and disaggregated accounting data. Comparisons of the information generated by different accounting methods also could be made.
REFERENCES


GLOSSARY 1

SEMIOTICS

Channel. A communication medium.

Channel Capacity. The maximum amount of data that can be transmitted on a communication medium.

Communication. "The establishment of a social unit from individuals, by the shared usage of language or signs" (Cherry 1978, 339).

Conditional Entropy. The average amount of uncertainty remaining after the receipt of a noisy signal.

Data. "...facts that are obtained through empirical observation" (Caspari 1968, 8).

Entropy. The expected syntactic information of a message.

Equivocation. The average amount of syntactic information that can be communicated on a noisy channel.

Information. "...the resultant of the coordination of data with knowledge when the data are screened, edited and evaluated for use..." (Caspari 1968, 8).

Knowledge. "...a group of well-confirmed law-like generalizations which relate data to their environment" (Caspari 1968, 8).

Message. "An ordered selection from an agreed set of signs...intended to communicate information" (Cherry 1978, 341).

Meta-Language. A language used for describing object-languages.

Object-Language. A language used for communication. Not used for the description of other object-languages.

Pragmatic Information. Indicates that a data item must be of value to the receiver.

Pragmatic Information Theory. The area of semiotics that entails the study of the effects of information upon individuals.
Receiver. An individual who receives a message. In some cases, such as the transmission of messages between computers, the receiver of a message may be an object.

Referent. That which is represented by a sign.

Semantic Information. Concerned with the relationship between signs or symbols and the objects or ideas that they represent. In order for a data item to contain semantic information, it must have meaning to the receiver.

Semantic Information Theory. The branch of semiotics that involves the study of the meaning of data.

Semantic Noise. Distortion of a transmitted message.

Semantic Redundancy. The existence of data beyond the original signal that presents the original message in a different format.

Semiotics. The study of signs. Semiotics encompasses syntactic information theory, semantic information theory, and pragmatic information theory.

Sign. "A transmission, or construct, by which one organism affects the behavior or state of another, in a communication situation" (Cherry 1978, 342).

Signal. "The physical embodiment of a message" (Cherry 1978, 342).

Syntactic Information. Indicates that the content of a message or data item should be unknown to the receiver before the message is received. This characteristic also is known as surprise value.

Syntactic Information Theory. The area of semiotics that concerns the study of the probabilities of occurrence of events.

Syntactic Noise. Distortion of communication signals.

Syntactic Redundancy. The existence of data beyond that needed to convey a message.
GLOSSARY 2

NEURAL NETWORKS

Adaptive Coefficients. The vectors of weights of previous computations that are stored in local memory of processing elements and are used to modify succeeding computations (Hecht-Nielsen 1988).

Backpropagation. A type of learning law that propagates error backward through a neural network in the process of adjusting the adaptive coefficients.

Connection. The electronic counterpart of dendrites, axons, and synapses. Its purpose is to provide a signal transmission pathway between processing elements (Hecht-Nielsen 1988).

Input Layer. The layer of processing elements that is initially presented data.

Learning Law. An equation that determines the revision of the adaptive coefficients as a function of the input values and (sometimes) of a feedback value supplied by the transfer function. This enables the system to modify itself in order to provide output values more closely related to the desired output (feedback value) and to structure information within itself, and thereby learn (Hecht-Nielsen 1988).

Middle Layer. The layer(s) of processing elements between the input layer and output layer.

Output Layer. The final layer of processing elements that produces an output.

Processing Element. The electronic counterpart of a neuron.

Threshold. A constant that screens out background noise (insignificant stimuli) and whose value is dependent upon the expected or actual level of the noise (Caudill November 1988).

Transfer Function. A function that determines the output of a processing element in response to the inputs and the adaptive coefficients.
Weight. A single adaptive coefficient (vector element) within a processing element that is associated with a single input connection. The weight determines the magnitude of the connection (Hecht-Nielsen 1988). A weight may be either excitatory or inhibitory, depending on the sign (+,−) associated with the weight. A weight is the counterpart of the strength (weight) of a synapse in the brain.
APPENDIX 1
SYNTACTIC REDUNDANCY OF ACCRUAL FINANCIAL STATEMENTS

Because conditional entropy is equal to the remaining uncertainty after receiving a noisy signal, e.g., cash-flow reports, the amount of syntactic information conveyed by accounting accruals must be less than or equal to this amount (some amount of uncertainty is always associated with financial statements). Therefore, the amount of syntactic information transmitted by accrual financial statements can be represented by:

\[ I = A - N + (CE - U) \]  

where \( I \) is the total syntactic information conveyed, \( A \) is the entropy (expected syntactic information from a noiseless channel), \( N \) is the syntactic noise induced by periodic reporting under a system of cash receipts and cash disbursements and other sources, and \( CE - U \) is the conditional entropy of periodic cash-flow reports minus any uncertainty after accrual adjustments (this final term represents the amount of syntactic information conveyed by the accruals). The amount of syntactic noise and remaining uncertainty vary depending on specific circumstances, such as the number

1 Note that in this equation \( N = CE \). These terms will be dropped later. Also note that \( U \) can also be thought of as conditional entropy. Recognize that in this study, a distinction is made between the conditional entropy of periodic cash-flow reports (represented in equation A1.2 by upper case letters) and the conditional entropy of periodic financial statements on an accrual basis (represented here by \( U \) and later in equations A1.4 and A1.5 by lower case letters).
of incomplete cash transactions at the financial reporting cut-off date and
the performance (or non-performance) and extent of an audit.

Equations 2.1 and 2.3 can be used to extend equation A1.1 to:

\[ (A1.2) \quad I = \sum_{i=1}^{n} p_i \log \frac{1}{p_i} + \sum_{ij} P(X_i|Y_j) \log \frac{1}{P(X_i|Y_j)} - N - U \]

where the first term on the right represents expected syntactic information
from a noiseless channel (entropy—equation 2.1), the second term on the
right represents the maximum amount of uncertainty that could be removed
by accounting accruals (conditional entropy—equation 2.3), the third term
on the right represents the amount of syntactic noise that is induced by
periodic reporting under a system of cash receipts and cash disbursements
and other sources, and the final term on the right represents any
remaining uncertainty not removed by accounting accruals. Since the
second and third term are equivalent, the equation simplifies to:

\[ (A1.3) \quad I = \sum_{i=1}^{n} p_i \log \frac{1}{p_i} - U \]

Since the second term on the right of the equation can be defined as the
conditional entropy remaining after accounting accruals, equations 2.3 and
A1.3 can be used to derive the final form of the equation:

\[ (A1.4) \quad I = \sum_{i=1}^{n} p_i \log \frac{1}{p_i} - \sum_{ij} p(x_i|y_j) \log \frac{1}{p(x_i|y_j)} \]

which was previously defined as equivocation.

Based on this definition of the syntactic information of accrual
financial statements, equations 2.4 and A1.4 can be used to define the
amount of syntactic redundancy (from all sources, including accounting
accruals) in those statements:
\[(A1.5) \quad R = 1 - \left[ \sum_{i=1}^{n} p_i \log \frac{1}{p_i} - \sum_{i,j} p(x_i,y_j) \log \frac{1}{p(x_i|y_j)} \right] + C_{\text{max}} \]

where $C_{\text{max}}$ is the channel capacity of the financial statements\(^2\).

\(^2\) It should be noted that the channel capacity of the method of communicating financial statements varies depending upon the communication medium used.
Neural networks are computers that are patterned after the cerebral cortex and consequently, are composed of a complex net of many relatively simple processing elements and process information in the same manner as the human brain, i.e., in parallel. Thus, many of the theories of the information processing capabilities of the brain also apply to neural networks.

Syntactic information theory can be used to illustrate one of the most interesting of a neural network's characteristics: the ability to learn. Syntactic information theory states that the information of a message about the occurrence of an event is \( \log \frac{1}{p} \), where \( p \) is the probability of the event occurring. Whenever there exists a series of possible events, the expected syntactic information of a message related to this series of events is called entropy and is defined as:

\[
\sum_{i=1}^{n} p_i \log \frac{1}{p_i}
\]

for a definite message (a message indicating which of the events occurred) where \( p_i \) is the probability of the occurrence of the event. The expected information of a nondefinite message (a message that indicates changes in the probabilities of events) is:
\((A2.2)\) \[ \sum_{i=1}^{n} q_i \log \frac{q_i}{p_i} \]

where \(q_i\) is the revised probability and \(p_i\) is the original probability. Thus, syntactic information theory describes the human brain's processing of syntactic information in terms of probabilities. Probabilities or weights also can be used to describe the information processing characteristics of neural networks.

There are other characteristics that distinguish neural networks from other technology. Neural networks may be viewed as a type of associative memory that is capable of adaptive pattern recognition (Pao 1989). This means that the system associates input patterns with the desired output patterns and modifies the mapping from input to output as environmental features change. Thus, when new patterns are input, the system outputs the desired output even if the relationship between inputs and outputs has shifted (Wasserman and Schwartz 1988). The system also can produce the desired output from a fragmented or partially incorrect input. Therefore, a neural network can be more accurately described as a content-addressable associative memory (the same as the brain) (Allman 1987). All processing in a neural network is performed with rather simple processing elements, similar to neurons in the brain. What the neural network paradigm attempts to demonstrate is how relatively slow, simple processing elements can be used to solve complex problems rapidly and reliably (Roberts 1988).

Robert Hecht-Nielsen (1988, 37) summarizes the important elements of neural networks in the following statement which describes the field of neurocomputing as:
... the engineering discipline concerned with nonprogrammed adaptive information-processing systems - neural networks - that develop associations (transformations or mappings) between objects in response to their environment. Instead of being given a step-by-step procedure for carrying out the desired transformation, the neural network itself generates its own internal rules governing the association, and refines those rules by comparing its results to the examples. Through trial and error, the network literally teaches itself how to do the task.

Neurocomputing is a fundamentally new and different information-processing paradigm - the first alternative to algorithmic processing.

Neural networks may be classified as either unsupervised learning networks or supervised learning networks. Unsupervised learning entails providing data to the network with no feedback as to how the network should classify the patterns present in the data—the network establishes relationships among the data items, discriminates patterns, and creates its own categorization\(^1\). In supervised learning, a feedback value informing

\[\text{Pattern recognition and classification are not the only capabilities of neural networks. Several types of network paradigms have been developed, each for a specific type of problem. A network paradigm is the consequence of the combination of processing element neuro-dynamics and network architecture (Klimasauskas 1988). Processing element neuro-dynamics are the result of the combined effects of the way that inputs are combined, how the activation level is used to propagate an output, and the way the learning law changes the weights associated with the processing elements (Klimasauskas, 1988). Network architecture encompasses the pattern and structure of the connections, the use or nonuse of feedback, the use or nonuse of competition among processing elements, and the flow of information through the network (Klimasauskas 1988). Thus, the kinds of problems that a particular network can solve usually depend on what network paradigm is used (Klimasauskas 1988). Although each paradigm is different from all others, network paradigms can be classified by function into four general types (Bayle 1988):

Pattern Associateor. Pairs of patterns are used to train this category of neural network which learns to associate input patterns with corresponding output patterns. Consequently, the network responds with the appropriate output pattern when presented with an input pattern.

Auto Associateor. This class of neural network is trained with complete data patterns and has the ability to output a complete pattern when presented with an incomplete portion of an original pattern or a complete pattern similar to an original pattern.}
the network how to classify the data pattern is given to the network in conjunction with the associated data patterns. The network then establishes relationships among the data items in a manner consistent with the feedback.

Neural networks are trained through the use of a training set. A training set consists of historical data that is relevant to the classification decision that the network will be making. Training data should consist of as many relevant data patterns as possible in order to assure that the network will have no difficulty in making classifications once training is complete. Each data pattern should consist of as many relevant data items as possible in order to assure that no information useful in discriminating among classes is missing. Once training is complete, new data patterns are compared to the data patterns previously input, the network determines which pattern the new data most closely fits, and a classification decision is made based on association of the new data pattern with the training set data patterns. Any information provided by the new data is incorporated into the neural network and is used in determining the classification of future data sets.

Classification Builder. Data patterns and their corresponding categories are used to train this type of neural network. After training is complete, the network is able to categorize data patterns similar to those used in training.

Regularity Detector. This category of neural network operates in the unsupervised mode (no feedback) and attempts to locate prominent relationships in data sets. Because no feedback is used, no training takes place with this type of network.
Neural Network Components

The architecture of neural networks is patterned after the human brain. Comparing and contrasting the architecture of neural networks with that of the human brain can describe the information processing characteristics of neural networks and their limitations.

The brain is composed of a vast network of interconnected processing elements called neurons. Individual neurons by themselves are not intelligent, but when connected to each other, they become very intelligent (Allman 1987). Each neuron is composed of (1) dendrites, a branching set of fibers that receive signals from other neurons, (2) a cell body, and (3) axons, another set of fibers that transmit signals to other neurons. Between the axons of one neuron and the dendrites of another are minute, adjustable connections called synapses, which can be either excitatory or inhibitory in nature. When a neuron receives input signals at its dendrites, a pulse is produced that travels through its axons. If the synapse at the end of the axon is excitatory, the pulse rate is increased. If the synapse is inhibitory, the pulse rate is decreased. The pulse rate output by the neuron also depends on the strength of the input signal and the strength (weight) of the synapse (Howard, Jackel, and Graf 1988).

It is estimated that the human brain contains between ten billion and one trillion neurons, each of which can be connected to between one thousand and one hundred thousand others (Allman 1987). These connections between neurons degenerate and regenerate over time. Thus, old connections deteriorate and new connections form with different neurons over a period of time.
Neural networks contain elements that correspond to many of the brain's components although these components are simplified compared to their brain counterparts. These simplified elements, however, are sufficient to provide a practical simulation of some of the brain's mechanisms. The following items are the major components of neural networks:

1. **Processing element**: The electronic counterpart of a neuron, sometimes referred to as a neurode. Processing elements have no control over what input is presented to them and thus react to, rather than create, their environment (Caudill February 1988). Processing elements consist of small amounts of local memory and processing power (Hecht-Nielsen 1988). The electronic component used as a processing element is traditionally a resistor, although many other forms of processing elements recently have appeared, e.g., silicon hydride compounds, "slices" of time, groups of digital transistors, and digital gates at the crosspoint of conductors (Hecht-Nielsen 1988).

2. **Connection**: The electronic counterpart of dendrites, axons, and synapses. Its purpose is to provide a signal transmission pathway between processing elements (Hecht-Nielsen 1988). Connections do not change, as do their brain counterparts. Therefore, the degree of connectivity between processing elements in neural networks is generally quite high in order to allow for simulated degeneration and regeneration. The degeneration and regeneration of connections can be simulated with learning laws pursuant to which the weights between processing elements fluctuate; a decrease in the absolute value of a weight to 0 represents a degeneration and an increase in the absolute value of a weight from 0 represents a regeneration. Although this is not biologically accurate, it is much easier
to construct this type of system than a true-to-biology model (Caudill February 1989).

(3) **Adaptive coefficients:** The vectors of weights of previous computations that are stored in local memory and are used to modify succeeding computations (Hecht-Nielsen 1988). The variation of the adaptive coefficients in response to input data is what provides the capability to learn; this is similar to the syntactic information theory concept of changes in probabilities resulting in information. Capacitors are often used to implement these weight vectors in neural networks.

(4) **Weight:** A single adaptive coefficient (vector element) within a processing element that is associated with a single input connection. The weight determines the magnitude of the connection (Hecht-Nielsen 1988). A weight may be either excitatory or inhibitory, depending on the sign (+ or -) associated with the weight. This is the counterpart of the strength (weight) of a synapse in the brain.

(5) **Transfer function:** A function that determines the output of a processing element in response to the inputs and the adaptive coefficients. It is also known as the activation function and in many cases is a sigmoidal function (Caudill June 1988), a hard limiter function, or a threshold logic function (Bayle 1988). The transfer function includes the system's learning law (Hecht-Nielsen 1988) and is normally fixed at the time a network is constructed (Caudill February 1988).

(6) **Learning law:** An equation that determines the revision of the adaptive coefficients as a function of the input values and (sometimes) of a feedback value supplied by the transfer function. This enables the system to modify itself in order to provide output values more closely related to the
desired output and to structure information within itself, and thereby learn (Hecht-Nielsen 1988).

Learning within a network can be either supervised or unsupervised. Supervised learning implies that feedback is presented to the network during training to indicate the desired output for a given set of inputs. Unsupervised learning indicates that the network has no feedback of desired outputs and thus has to structure its output based only upon the input training sets (Caudill February 1988).

(7) **Threshold:** A constant that screens out background noise (insignificant stimuli) and whose value is dependent upon the expected or actual level of the noise (Caudill November 1988).

These items capture the essence of the brain's functioning; however, since researchers do not totally understand how the brain works, all of the functions and elements of the brain cannot be incorporated into a neural network. A neural network is not a perfect replica of a brain, although it may reasonably simulate its information processing mechanisms.

**Neural Networks Versus Conventional Computers**

Neural Networks differ from conventional computers in several ways. First, neural networks differ from conventional computers in their architecture. The nucleus of a conventional computer is generally one (or more) extremely complex processing element which addresses an array of memory locations where data and instructions are separately stored. It then fetches an instruction and any data required by that instruction, executes the instruction, and then stores the result back in memory. Thus, conventional computers are basically sequential: everything occurs in a deterministic sequence of operations (Caudill December 1987).
Alternatively, neural networks are neither sequential nor deterministic. They do not execute instructions as much as they behave or react to input. They are composed of a large number of extremely simple processing elements and do not have a separate memory for storing data. These processing elements do not execute a series of instructions, but respond in parallel to inputs. The final result is not stored in a memory location, but consists of the overall state of the system after it has reached an equilibrium condition (Caudill December 1987). Thus, information in a neural network is stored both in the way the processing elements are connected and in the weights associated with each processing element. Neural network architectures may be entirely electronic, electro-optical or totally optical (Hecht-Nielsen 1988).

Although the architectures and operations of neural networks and conventional computers are quite different, neural networks can perform the same types of tasks normally associated with conventional computers. Conventional computers can also simulate neural networks. Since a conventional computer can be constructed from exclusive-or gates (gates that are on when exactly one of two inputs is on--off otherwise), and exclusive-or gates can be constructed inside neural networks, it follows that a conventional computer can be built inside a (large) neural network. Thus, neural networks can perform the types of tasks generally executed by conventional computers (Roberts 1988). And although conventional computers can simulate a neural network, the processing occurs sequentially, as opposed to parallel, and consequently, the simulation will be slower than a true neural network (Roberts 1988).
Neural networks also are generally more efficient in the performance of a task than conventional computers. This phenomenon is the result of the spectacular speed at which neural networks operate. Because neural networks operate much faster than conventional computers, the number of gate-seconds needed by a neural network to perform a task is much lower than the number of gate-seconds required by a conventional computer (Roberts 1988). However, many problems are subject to the von Neumann bottleneck—they are not naturally parallel. Thus, if these types of problems are solved in a neural network, the system often runs only as fast (and thus will be only as efficient) as a conventional, serial computer (Wasserman and Schwartz 1987).

Neural networks also are much more tolerant of noise in the input data than are conventional computers. While conventional computers require precise inputs to produce the desired outputs, it has been estimated that neural networks can effectively operate with input noise levels as high as 30 or 40 percent. Thus, neural networks can extract ideal output from less than ideal input (Wasserman and Schwartz 1988).

Neural networks can operate with damage to as many as 15 percent of their processing elements. This much damage to a conventional computer would be disastrous. In fact, damage to as little as one percent of a conventional computer's circuits will probably prevent its operation (Allman 1987).

Neural networks are best at solving types of problems that are difficult for conventional computers, and conventional computers are best at solving types of problems that are difficult for neural networks. Generally, neural networks are good at solving problems at which humans
excel, such as understanding continuous language or recognizing patterns. Conventional computers are not very good at performing these types of tasks. However, conventional computers excel at arithmetic and logic, areas in which neural networks (and humans) are notably deficient (Allman 1987).

Programming a neural computer differs from programming a conventional computer. Traditional, automated information processing is based upon John von Neumann's "glorified adding machine" concept (Hecht-Nielsen 1988). Programming a conventional computer entails the detailing of specific step-by-step instructions for the computer to execute. If the problem is not thoroughly understood, no program can be written, and the problem cannot be solved with a conventional computer.

The "programming" of neural networks, however, entails providing the network inputs and either telling it what the output should be or letting it determine that for itself. The network determines the relationship among the inputs by itself—step-by-step instructions are not needed. The only algorithms that need to be provided are those involving transfer functions, scheduling functions, learning laws, etc. Only a few programming languages, including PS, Panspec, Anspec and Axon, exist to alter these algorithms (as compared to the large number available for conventional computers) (Hecht-Nielsen 1988).

Neural Networks Versus Expert Systems

Traditional expert systems are based on the concept that all human thought processes involve symbol manipulations. "Theorists believed they could simply translate the world into symbols, manipulate the symbols, and translate the results back into the language of the real world" (Allman 1987, 60). However, just because some types of thinking, such a formal
logic and arithmetic, involve symbol manipulation, it is not reasonable to conclude that all other types of thinking do also. Scientists have identified over 100 different types of neurons (Klimasauskas 1988); it is possible that they do not all process information in the same manner. Perhaps the situation is that the majority of thinking that we do does not involve symbol manipulation. If this is the case, then traditional expert systems that use symbol manipulations to solve problems should be less than adequate in certain domains. This is certainly the case. This problem of expert systems is not solved by simply adding more structures (rules). According to Thompson and Thompson (1987, 27) "Unfortunately we don't know how to make these structures flexible enough to provide the kind of robust, general-purpose intelligence most people possess." A good example of how expert systems fail at certain relatively simple tasks (simple at least for humans) is related to the Strategic Defense Initiative (SDI). SDI relies on the assumption that space-borne computers should be able to distinguish ICBMs from dummy warheads—from any angle. Unfortunately, traditional expert systems are not very good at pattern recognition and thus, cannot perform this task adequately. Neural networks on the other hand, are quite good at pattern recognition. They are also very good at the classification of items and, thus, are superior to expert systems in problems involving classification: pattern recognition, continuous speech recognition, forecasting and modeling, diagnosis, and so forth.

Neural networks also have other advantages compared to expert systems. For example, expert systems have several disadvantages including the time-consuming nature of knowledge extraction [this problem is known as the Feigenbaum bottleneck--not only must programs be written, but
knowledge also must be extracted from experts (Wasserman and Schwartz 1987), the inability to learn, and unpredictable behavior outside of their domain (Bailey, Thompson and Feinstein 1988). Neural networks, however, do not require knowledge acquisition from outside sources: they obtain their knowledge from the data itself and can adapt to changes in the data (learn). The neural network learning process does not mean that the need for knowledge engineers and experts will be eliminated. Experts will still be needed to indicate relevant data and knowledge engineers will be needed to determine how the data should be preprocessed (neural networks generally work better with data that has been preprocessed to emphasize important relationships). Thus, neural networks simplify the knowledge acquisition process. They also have the ability to generalize from specific examples and have the ability to discover complex relationships among data items. Neural networks can also operate effectively with noisy input and have controlled degradation in problems outside their range of experience (Bailey, Thompson and Feinstein 1988).
APPENDIX 3
BACKPROPAGATION NEURAL NETWORKS

Backpropagation is one of the most successful neural network paradigms (Caudill and Butler 1990). Backpropagation networks must contain at least three layers of processing elements: the input layer, a middle layer, and an output layer (Caudill June 1988). The number of processing elements in the input layer must be greater than or equal to the number of inputs, and the output layer should have a sufficient number of processing elements [to enable the output of the desired response] (Caudill June 1988). Care must be taken, however, in deciding how many processing elements to include in the middle layer. If too many processing elements are included, the network will tend to memorize (grandmother) the input patterns rather than generalize the characteristics of the input. Grandmothering reduces the network's ability to adapt to changing conditions after completion of training. On the other hand, if too few processing elements are included, the number of iterations required to train the network will dramatically increase, and the accuracy of recall probably will decrease (Caudill June 1988). Each layer in the network is connected to the immediately preceding and subsequent layers. However, there are no connections between processing elements within the same layer.

Training of the network is begun by randomizing all of the weights in the network, except those in the input layer which are set to +1.
Training data is then presented to the network and the weights are adjusted using the Delta Rule:

\[(A3.1) \quad W_{new} = W_{old} + \beta \frac{EX}{|X|^2}\]

where \(W_{new}\) is the weight vector after adjustment, \(W_{old}\) is the weight vector before adjustment, \(E\) is the error value vector (values representing the difference between the desired response and the actual output), \(X\) is the input vector, and \(\beta\) is a learning constant that must be positive and less than 1 (Caudill February 1988). The processing elements in the middle layer compute activation functions based on the data received from the input layer. The activation functions determine the activity, or excitation level, generated in each processing element as a result of input signals of a particular size (Caudill June 1988). An activation function is computed as follows:

\[(A3.2) \quad A = f(\sum W_i \times X_i)\]

where \(W_i\) is the weights of the \(i\)th input processing element, \(X_i\) is the input for the \(i\)th input processing element, and \(A\) is the summed input. The function \(f(x)\) is the activation function of the processing element (Caudill June 1988).

The activation function should always be sigmoidal for a backpropagation network. Generally, the upper limit of the sigmoid is set to +1 and the lower limit is set to either 0 or -1. The exact details of the sigmoid, however, are not important as the general S-shape (Caudill June 1988). The sigmoidal curve can be expressed as:

\[1.\] Appendix 4 contains a detailed discussion of the Delta rule.
where $T$ is the threshold, $X$ is the input and $e$ is the mathematical exponential constant (Caudill June 1986).

The value (activity) generated by the activation function is the output for the middle-layer processing elements and generally will be between 0 and 1 for each processing element. These outputs then become the inputs for the output-layer processing elements. Each output-layer processing element then computes an activation in exactly the same fashion as the middle-layer processing elements. These then become the output for the system (Caudill June 1988).

The weights for the processing elements are then adjusted using the Delta Rule. However, the error value for the middle layer is uncertain. This, in turn, means that the percentage of the error in the output layer caused by incorrect processing of the output layer and the percentage caused by incorrect output from the middle layer cannot be determined. To distinguish these two potential causes of output-layer error, the errors of each output-layer processing element are transmitted back to the middle-layer processing elements using the same connections and weights as the middle layer used to transmit its outputs to the output layer (Caudill June 1988). An error for each middle-layer processing element is then computed based on its responsibility for the output layer's error. This is computed as follows:

$$(43.4) \quad e_i = f'(o) \times [\sum W_{ij} \times E_j]$$

where $e_i$ is the error of the $i$th middle-layer processing element and the
sum is taken over \( j \), where \( j \) indicates the \( j \)th output-layer neurode. The term \( f'(I) \) is the first derivative of the activation function of the middle-layer processing element (Caudill June 1988). Based on the sigmoidal activation function listed earlier, this derivative is computed as:

\[
(A3.5) \quad f'(X) = f(X)(1-f(X))
\]

and serves two purposes: (1) it ensures that, as the outputs approach 0 and 1, only small changes occur and thus, contributes to network stability; and (2) it moderates the blame assigned to the middle-layer processing elements and thus ensures that only small to moderate changes are made to the middle-layer processing elements' weights (Caudill June 1988). After the errors are assigned to the middle-layer processing elements, the Delta Rule is applied and the weights are adjusted. The middle-layer errors are then transmitted back to the input layer and the process continues as discussed above.

A problem with the backpropagation learning law is that it can become stuck in a local minimum. A solution to this problem is to add a momentum term to the Delta Rule. The momentum term, \( \alpha \), is a constant that is multiplied by the change in the weight vector of the processing element from the previous computation, \( (W_{new} - W_{old}) \). Thus, the Delta Rule becomes (Caudill June 1988):

\[
(A3.6) \quad W_{new} = W_{old} + \frac{\beta EX}{X^2} + \alpha(W_{new} - W_{old})_{prev}
\]
APPENDIX 4

THE DELTA RULE

The most widely used neural network learning law is the Delta Rule (or Least Mean Squared Training Law) developed by Bernard Widrow and Ted Hoff at Stanford University (Caudill February 1988). This learning law operates in the supervised mode which means that it provides feedback to the processing elements regarding the desired output for a given set of inputs. This feedback value, sometimes called the mentor input, should have a weight of +1 permanently attached to it (Caudill February 1988).

When a processing element is presented a set of input values, it computes an error value, $E$, for that input set by comparing its actual output, $y$, to the desired output, $I_d$. Thus, $E = I_d - y$. The adjustments that are made to the weights based on this error value are computed using the Delta Rule:

$$W_{\text{new}} = W_{\text{old}} + \frac{\beta E X}{|X|^2}$$

(44.1)

where $W$ is the weight vector, $X$ is the input vector and $\beta$ is a learning constant that must be positive and less than 1 (Caudill February 1988).

The learning constant, which is the rate at which the current weight vector shifts to the ideal weight vector, must be less than 1, otherwise the Delta Rule will tend to overshoot this ideal weight vector as this position is approached (Caudill February 1988). The ideal value of the learning constant also depends on the distribution of the training data. If the
training are grouped tightly together, then a larger value of \( \beta \) (e.g., .8 to .9) can be used than if the training data are relatively scattered (e.g., .1 to .2) (Caudill February 1988).

The Delta Rule attempts to ensure that the total mean squared error is minimized in the network and results in a vector that is parallel to the input vector and that has a magnitude as defined in the Delta Rule equation. The weight vector is then adjusted by adding the Delta Rule vector to the current weight vector (Caudill, February 1988).

This process of adjusting the weights can be explained through the use of simple geometric terms. According to Caudill (February 1988) the aggregate mean squared error is a quadratic function of the weight vector. Consequently, plotting the mean squared error against all possible weight vectors yields an n-dimensional hyperparaboloid, the bottom of which represents the minimum mean squared error. The weight vector associated with the minimum mean squared error is the ideal weight vector.

The Delta rule moves the current weight vector from its position on the hyperparaboloid toward the ideal weight vector in a negative gradient fashion. Because the negative gradient is the most direct route to the bottom of the hyperparaboloid, the Delta rule minimizes the mean squared error in the most efficient manner possible.
APPENDIX 5

DECISION MODEL FOR DATA ITEM SELECTION

Testing the theories of this study required the use of cash-flow and accrual accounting data related to operating activities. COMPUSTAT contains numerous accounting data items associated not only with operating activities but also financing and investing activities. Four decision rules were used to select cash-flow and accrual accounting data items related to cash flows from operating activities.

The first decision rule was used to choose data items representing accruals from the income statement.

1. The data item is used on the statement of cash flows (indirect method) to reconcile net income to cash flows from operations.

This decision rule resulted in selection of the following accrual items related to operating flows:

a. depreciation and amortization,
b. provision for losses on accounts receivable,
c. gain on sale of facility,
d. undistributed earnings of affiliate,
e. change in accounts receivable,
f. change in inventory,
g. change in prepaid expenses
h. change in other current assets,
i. change in accounts payable and accrued expenses,
j. change in income taxes payable, and
k. change in other current liabilities.

1 According to Ijiri (1986, 746):

The role of flow accounts . . . is to explain or to "account for" changes in the net balance in stock accounts based on the underlying causes that are considered to be responsible for the change.
These accruals were matched with data items on COMPUSTAT. Provision for losses on accounts receivable was eliminated because it is not a separate data item on COMPUSTAT. Change in prepaid expenses could not be calculated because prepaid expenses is not an individual data item on COMPUSTAT. Change in other current assets and change in other current liabilities each contained a plethora of accruals, both syntactically and semantically redundant, and consequently were not used. The COMPUSTAT data items that correspond to the remaining accruals are:

a. depreciation and amortization,
b. sale of property, plant and equipment and sale of investments—loss (gain),
c. equity in net loss (earnings),
d. accounts receivable—decrease (increase),
e. inventories—decrease (increase),
f. accounts payable and accrued liabilities—increase (decrease), and
g. income taxes payable—increase (decrease).

The second decision rule was used to select data items representing accruals from the balance sheet.

2. The data item represents potential cash inflows or outflows from operating activities.

This decision rule resulted in the selection of the following accrual items related to operating stock:

---

2 Prepaid expenses appears as a separate data item on annual COMPUSTAT tapes but does not appear on quarterly COMPUSTAT tapes because of data aggregation. The level of data aggregation used in this study is equal to that of quarterly COMPUSTAT tapes and, consequently, prepaid expenses was not available for use in this study.

3 Accounts receivable—decrease (increase), inventories—decrease (increase), accounts payable and accrued liabilities—increase (decrease), and income taxes payable—increase (decrease) were not directly available on COMPUSTAT but were calculated by subtracting the previous period's balance from the current period's balance.

4 Stock accounts are used to report assets and liabilities at the end of a fiscal period (Ijiri 1986).
a. accounts receivable,
b. inventories,
c. accounts payable and accrued liabilities,
d. income taxes payable,
e. current portion of long-term debt, and
f. deferred taxes.

The corresponding data items on COMPUSTAT are:

a. receivables--total,
b. inventories--total,
c. accounts payable and accrued liabilities,
d. income taxes payable,
e. debt in current liabilities, and
f. deferred taxes.

The third decision rule was used to select a cash-flow data item and
reflects the fact that cash flows from operations may follow a random walk.

3. The data item can be used to construct a random walk model of
cash flows from operations.

A random walk process generates expectations of future cash flows that
depend only on the prior period's cash flows. Consequently the data item
identified by the third decision rule is:

a. cash flows from operations.

The COMPUSTAT data item that corresponds to this cash-flow data item is:

a. operating activities--net cash flow.

The fourth decision rule recognizes the fact that future cash flows
can be influenced by cash presently on hand. For example, if a shortage
of cash exists, payments to creditors may be delayed. Shortages of cash
may also result in increased effort in the collection of amounts due from
customers.

---

5 Although this data item is not directly related to operating
activities, a sizable amount of debt due within a next year could influence
operating activities, e.g., expenses could be reduced to build up cash or
the effort expended to collect overdue accounts receivable could be
increased.
4. The data item represents cash on hand.

The data item identified by this decision rule is:

- cash on hand.

The COMPUSTAT data item that corresponds to this cash-basis data item is:

- cash and short-term investments.

Figure 3 summarizes the decision model used to select data items.
APPENDIX 6

DECISION MODEL FOR CLASSIFYING DATA ITEMS INTO DATA TYPES

The classification of data items into data type categories followed a decision model composed of five decision rules. The first decision rule identifies cash-flow accounting data.

1. The data item is not an accounting accrual.

Cash-flow accounting data in this study included:

a. cash and short-term investments\(^1\), and

b. operating activities—net cash flow.

The second decision rule classifies balance sheet accruals as syntactically redundant.

2. The data item signals a message that is not reported under a system of cash receipts and cash disbursements.

The following accruals were classified as syntactically redundant by decision rule 2:

a. receivables—total: signals an amount of cash that is owed to the firm,

b. accounts payable and accrued liabilities: signals an amount of cash that is owed by the firm to various parties,

c. income taxes payable: signals an amount of cash that is owed by the firm to the government,

d. equity in net loss (earnings): signals the amount of a subsidiary's net income or loss that is associated with stock held by the firm, and

\(^{1}\) Short-term investments, an accrual, is combined with cash because of COMPUSTAT data aggregation.
e. deferred taxes: signals the difference in the amount of income taxes owed to the government and the amount of income taxes that has been expensed.

Decision rule 3 identifies balance sheet accruals that are semantically redundant.

3. The data item restates a message that is signaled under a system of cash receipts and cash disbursements.

This decision rule classified the following accruals as semantically redundant:

a. Inventories—total: Cash-flow accounting reports inventory cost as a disbursement. This accrual initially signals inventory cost as an unexpired cost (asset).

b. Debt in current liabilities: Cash-flow accounting reports this item as a receipt. This accrual signals this item as a liability.

c. Depreciation and amortization: Cash-flow accounting reports the cost of property, plant, and equipment and the cost of intangible assets, e.g., research and development costs, as disbursements when incurred. This accrual indicates that the expiration of the cost occurs over a period of time and is associated with revenues generated by the costs.

d. Sale of property, plant and equipment and sale of investments—loss (gain): Cash-flow accounting reports this item as receipt. This accrual indicates that fact that unexpired costs are associated with the cash inflows and that the difference between the unexpired costs and cash inflows represent revenue.

The fourth decision rule identifies syntactically redundant accruals related to flows.

4. The data item represents a change in a syntactically redundant balance sheet accrual.

The following syntactically redundant accruals were identified by decision rule 4:

2 According to IJiri (1986, 746):

The role of flow accounts . . . is to explain or to "account for" changes in the net balance in stock accounts based on the underlying causes that are considered to be responsible for the change.
a. accounts receivable—decrease (increase),

b. accounts payable and accrued liabilities—increase (decrease), and
c. income taxes payable—increase (decrease).

Decision rule 5 identifies semantically redundant accruals related to flows.

5. The data item represents a change in a semantically redundant balance sheet accrual.

One semantically redundant accrual was identified by decision rule 5:

a. inventories—decrease (increase).

Figure 4 summarizes the decision model used to categorize data items into data types.
Fig. 1. Information Function. Adapted From: Lev 1969a, fig.1
Fig. 2. Entropy Function. Adapted From: Lev 1969a, fig. 2
Separate data item on COMPUSTAT?

YES

Single data item on COMPUSTAT?

YES

Is data item used to reconcile net income to cash flows from operations?

YES

Include in data set as income statement accrual.

NO

Fig. 3.—Decision Model for Data Item Selection
Data item represents potential operating cash inflow or cash outflow?

Include in data set as balance sheet accrual.

Data item can be used to construct random walk model of operating cash flows?

Include in data set as cash-flow data item.

Data item represents cash on hand?

Include in data set as cash-flow data item.

Fig. 3.—Continued
Fig. 3.—Continued
Fig. 4.—Decision Model for Classifying Data Items
Fig. 4.—Continued
APPENDIX 8

LETTERS OF PERMISSION
Harian Etheridge  
3720 West Alabama, Apt. # 5217  
Houston, TX  77021

Pau Gerhardt  
American Accounting Association  
6111 Bessie Drive  
Sarasota, FL  34231

Dear Mr. Gerhardt:

I am requesting permission from the American Accounting Association to use two figures derived from Baruch Lev's monograph, Accounting and Information Theory, Studies in Accounting Research # 2, in my dissertation. I have enclosed copies of the figures from my dissertation for your examination. Please send your reply to:

Harian Etheridge  
3720 West Alabama, Apt. # 5217  
Houston, TX  77021

Thank you for your assistance.

Sincerely,

Harian Etheridge

enclosures

Figures, courtesy of the American Accounting Association - Reprint Date:

[Signature]

12-26-91
VITA

Harlan Lynn Etheridge was born on August 29, 1960 in Lake Charles, Louisiana. He attended Lake Charles public schools and in 1978 entered McNeese State University. In July 1981, he graduated from McNeese with a Bachelor of Science in accounting.

He worked for six years in public accounting and industry before deciding to pursue a Ph.D. in accounting at Louisiana State University. Four years and numerous gray hairs later, he completed his Ph.D. He is currently an assistant professor of accounting at the University of Houston, University Park specializing in accounting information systems.

His research interests include the application of artificial intelligence to accounting and auditing situations, examination of the intrinsic qualities of accounting data, and the application of chaos theory to accounting and auditing. His teaching interests include accounting information systems, EDP auditing, and managerial accounting.
DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: Harlan Lynn Etheridge

Major Field: Accounting

Title of Dissertation: An Examination of Semiotic Theories of Accounting Accruals

Approved:

[Signatures]

Major Professor and Chairman
Dean of the Graduate School

EXAMINING COMMITTEE:

[Signatures]

Date of Examination: 8/21/91